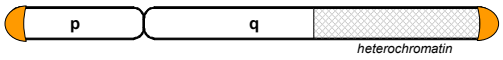


State of the Y Chromosome:

New Advances and State of the Science for Y Chromosome DNA Testing

John M. Butler, Ph.D.

Canadian Forensic DNA Technology Workshop
Toronto, Ontario
June 8, 2005

Presentation Outline

- Advantages of the Y-Chromosome
- Characteristics of Core Y-STR Loci
- Available Y-STR Kits
- Available Y-STR Databases
- Y-STR Population Studies
- Statistics with Y-STR Haplotypes
- New Y-STRs
- Locus Duplication and Deletion
- Y-SNPs

"State of the Y STR Assay" in June 2000

From J.M. Butler talk June 1, 2000 at CHI "DNA Forensics" meeting (Springfield, VA)

- A number of multiplex reactions have been reported in the literature but Y STR multiplexes have not reached their potential…
- Very little PCR optimization to-date (most work has been done with the original PCR primer sequences)
- No commercial Y STR kit exists yet (therefore these markers remain inaccessible to the general forensic DNA community)
- New Y STR markers are becoming available which will greatly improve the power of discrimination between unrelated individuals (e.g., DYS385) and these will need to be incorporated into future multiplex sets

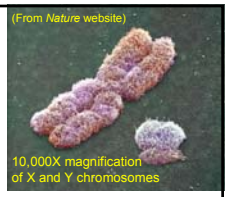
What has happened in the past 5 years

- "Full" Y-chromosome sequence became available in June 2003; over 200 Y-STR loci identified (only ~20 in 2000)
- Selection of core Y-STR loci (SWGDAM Jan 2003)
- Multiple commercial Y-STR kits released
  - Y-PLEX 6,5,12 (2001-03), PowerPlex Y (9/03), Yfiler (12/04)
- Many population studies performed and databases generated with thousands of Y-STR haplotypes
- Forensic casework demonstration of value of Y-STR testing along with court acceptance

THE HUMAN Y CHROMOSOME: AN EVOLUTIONARY MARKER COMES OF AGE

Mark A. Jobling & Chris Tyler-Smith

Nature Reviews Genetics (2003) 4, 598-612

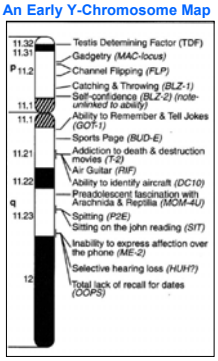


Abstract

- **Until recently, the Y chromosome seemed to fulfill the role of juvenile delinquent among human chromosomes — rich in junk, poor in useful attributes, reluctant to socialize with its neighbors and with an inescapable tendency to degenerate. The availability of the near-complete chromosome sequence, plus many new polymorphisms, a highly resolved phylogeny and insights into its mutation processes, now provide new avenues for investigating human evolution. Y-chromosome research is growing up.**

Traits found on the Y - Chromosome

An Early Y-Chromosome Map



Science (1993) 261:679

- spitting
- incessant use of TV remote buttons
- if lost, cannot stop and ask for directions
- ability to recall facts about baseball/basketball/hockey/golf/etc.
- male pattern baldness
- congregates with other Y-chromosome bearers to do "guy things"
- Source of "Testosterone poisoning"

Value of Y-Chromosome Markers

J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition; Table 9.1

Application	Advantage
Forensic casework on sexual assault evidence	**Male-specific amplification** (can avoid differential extraction to separate sperm and epithelial cells)
Paternity testing	Male children can be tied to fathers in motherless paternity cases
Missing persons investigations	**Patrilineal male relatives may be used for reference samples**
Human migration and evolutionary studies	Lack of recombination enables comparison of male individuals separated by large periods of time
Historical and genealogical research	Surnames usually retained by males; can make links where paper trail is limited

Disadvantages of the Y-Chromosome

- Loci are not independent of one another and therefore rare random match probabilities cannot be generated with the product rule; must use haplotypes (combination of alleles observed at all tested loci)
- Paternal lineages possess the same Y-STR haplotype (barring mutation) and thus fathers, sons, brothers, uncles, and paternal cousins cannot be distinguished from one another
- Not as informative as autosomal STR results

Forensic Advantages of Y-STRs

- **Male-specific amplification** extends range of cases accessible to obtaining probative DNA results (e.g., fingernail scrapings, sexual assault without sperm)
- **Technical simplicity due to single allele profile**; can potentially recover results with lower levels of male perpetrator DNA because there is not a concern about heterozygote allele loss via stochastic PCR amplification; number of male contributors can be determined
- **Courts have already widely accepted STR typing**, instrumentation, and software for analysis (Y-STR markers just have different PCR primers)
- **Acceptance of statistical reports using the counting method** due to previous experience with mtDNA

Scenarios Where Y-STRs Can Aid Forensic Casework

- Sexual assaults by vasectomized or azoospermic males (no sperm left behind for differential extraction)
- Extending length of time after assault for recovery of perpetrator's DNA profile (greater than 48 hours)
- Fingernail scrapings from sexual assault victims
- Male-male mixtures
- Other bodily fluid mixtures (blood-blood, skin-saliva)
- Gang rape situation to include or exclude potential contributors

Some Reported Casework Examples

J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition; Table 9.7

Kit/Loci Used	Reference	Comments
In-house assay with DYS19, DYS390, DYS389I/II	Prinz *et al.* (2001) *Forensic Sci. Int.* 120: 177-188	In one year at the New York City Office of the Chief Medical Examiner, Y-STR testing was performed in more than 500 cases with over 1000 evidence and reference samples examined. A full or partial profile was obtained on 81% of all tested evidence samples (740 worked/915 samples tested). Mixtures of at least two males were observed in 97 instances. In male/female mixtures of up to 1:4000, the male component could be cleanly detected.
In-house assay with 9 Y-STR loci amplified in 3 PCR reactions	Dekairelle and Hoste (2001) *Forensic Sci. Int.* 118: 122-125	Y-STR typing was attempted on 166 semen traces from 89 cases that failed to yield a detectable male autosomal profile following differential extraction. About half of the cases had sufficient DNA to produce a Y-STR profile.
In-house assay with DYS393, DYS389I/II	Sibille *et al.* (2002) *Forensic Sci. Int.* 125: 212-216	**Y-STR results could still be obtained more than 48 hours after the sexual assault in 30% of the cases examined.** In 104 swabs collected with no evidence of sperm, Y-STRs could be detected in ~29% of the samples tested.
In-house assay with DYS19, DYS390, DYS389I/II	Prinz (2003) *Forensic Sci. Rev.* 15: 189-196	Six case studies are reviewed along with advantages and disadvantages of Y-STR testing in each case: (1) different semen donors on vaginal swab and underwear; (2) possible oligospermic perpetrator gave a nice Y-STR profile but failed to have a "male" fraction with differential extraction; (3) oral intercourse with no autosomal results—not possible to enrich male cell fraction with differential extraction in cases involving saliva; (4) presence of multiple semen donors created a complex autosomal mixture that could be sorted out with Y-STR results; (5) sperm cell fraction lacked amelogenin Y-specific peak due to known deletion—Y-STR results confirmed that the sperm cell fraction DNA was of male origin; and (6) Y-STR testing was used to rapidly screen 18 semen stains for comparison to 5 suspects and thus save the time of performing the differential extraction
Y-PLEX 6 and Y-PLEX 5 kits	Sinha (2003) *Forensic Sci. Rev.* 15: 197-201	Five cases are reviewed: (1) criminal paternity case with a male fetus where the alleged father could not be excluded as the biological father; (2) autosomal STR test resulted in an uninterpretable mixture—suspect was excluded at 3 of the 7 Y-STR loci tested; (3) Y-PLEX 6 STR profile matched suspect with sweat stains on cloth found at crime scene; (4) fingernail cuttings from a victim matched a suspect at 11 Y-STR loci while another suspect was excluded at 2 loci; (5) semen positive stain with no sperm cells produced a Y-PLEX 6 profile consistent with the male suspect
Y-PLEX 6 and Y-PLEX 5 kits	Sinha *et al.* (2004) *J. Forensic Sci.* 49: 691-700	Seven cases are reviewed (some are the same as Sinha 2003) and a list of cases where Y-STR results have been accepted in U.S. courts is provided.

Y-STRs in Casework

Sinha *et al.* (2004) *J. Forensic Sci.* 49: 691-700

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

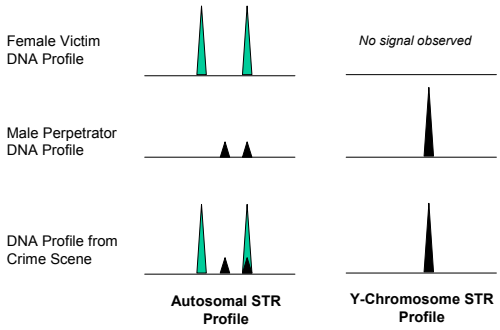
Sudhir K. Sinha,[1] Ph.D.; Bruce Budowle,[2] Ph.D.; Ranajit Chakraborty,[3] Ph.D.; Ana Paunovic,[1] B.S.; Robin DeVille Guidry,[1] B.S.; Chris Larsen,[1] M.S.; Amrita Lal,[1] M.S.F.S.; Megan Shaffer,[1] Ph.D.; Gina Pineda,[1] M.S.; Siddhartha K. Sinha,[1] B.S.; Elaine Schneida,[1] B.S.; Huma Nasir,[1]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

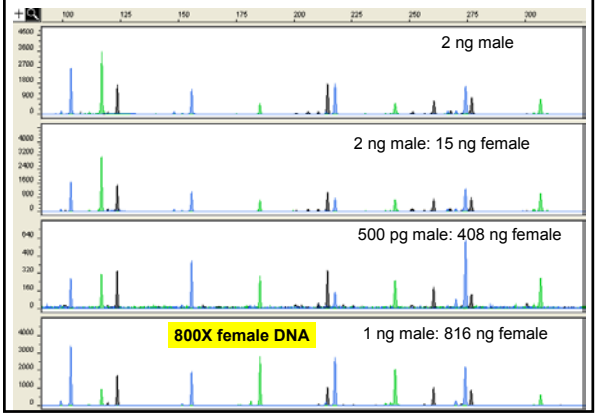
Case	Date	Jurisdiction	Docket No.	Notes
State of LA vs. Samuel Williams	10/23/01	Orleans Parish	416-355	Criminal paternity case
State of MS vs. Leon Felder	6/26/01	Pike County	00-557-KA	Sexual assault case-also had other STRs, Y-STR produced no result
State of GA vs. Ali R. Shabazz	7/31/02	Dekalb County	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

Y-STRs can permit simplification of male DNA identification in sexual assault cases

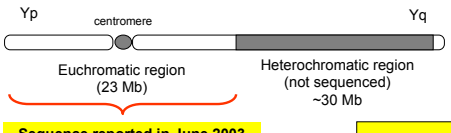


Butler, J.M. (2005) *Forensic DNA Typing, 2nd Edition*, Figure 9.2, ©Elsevier Science/Academic Press

PowerPlex Y Performance in Our Hands



"Full" Y-Chromosome Sequence



Sequence reported in June 2003 as part of Human Genome Project

Sequence analysis of the Y-chromosome has revealed that it is much more complex than previously thought…

Butler, J.M. (2005) *Forensic DNA Typing, 2nd Edition*, Figure 9.4(B), ©Elsevier Science/Academic Press

Y-Chromosome Sequence Published

Skaletsky *et al.* (2003) *Nature* 423: 825 - 837 (19 June 2003)

articles

The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes

Helen Skaletsky*, Tomoko Kuroda-Kawaguchi*, Patrick J. Minx†, Holland S. Cordum†, LaDeana Hillier†, Laura G. Brown*, Sjoerd Repping‡, Tatyana Pyntikova*, Johar Ali*, Tamberlyn Bieri†, Asif Chinwalla†, Andrew Delehaunty†, Kim Delehaunty†, Hui Du†, Ginger Fewell†, Lucinda Fulton†, Robert Fulton†, Tina Graves†, Shun-Fang Hou†, Philip Latrielle†, Shawn Leonard†, Elaine Mardis†, Rachel Maupin†, John McPherson†, Tracie Miner†, William Nash†, Christine Nguyen†, Philip Ozersky†, Kymberlie Pepin†, Susan Rock†, Tracy Rohlfing†, Kelsi Scott†, Brian Schultz†, Cindy Strong†, Aye Tin-Wollam†, Shiaw-Pyng Yang†, Robert H. Waterston†, Richard K. Wilson†, Steve Rozen* & David C. Page*

* Howard Hughes Medical Institute, Whitehead Institute, and Department of Biology, Massachusetts Institute of Technology, 9 Cambridge Center, Cambridge, Massachusetts 02142, USA
† Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Boulevard, St Louis, Missouri 63108, USA
‡ Center for Reproductive Medicine, Department of Gynaecology and Obstetrics, Academic Medical Centre, Amsterdam 1105 AZ, the Netherlands

X-Chromosome Sequence Published

Ross *et al.* (2005) *Nature* 434: 325 - 337 (17 March 2005)

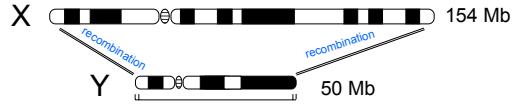
articles

The DNA sequence of the human X chromosome

A list of authors and their affiliations appears at the end of the paper

The human X chromosome has a unique biology that was shaped by its evolution as the sex chromosome shared by males and females. We have determined 99.3% of the euchromatic sequence of the X chromosome. Our analysis illustrates the autosomal origin of the mammalian sex chromosomes, the stepwise process that led to the progressive loss of recombination between X and Y, and the extent of subsequent degradation of the Y chromosome. LINE1 repeat elements cover one-third of the X chromosome, with a distribution that is consistent with their proposed role as way stations in the process of X-chromosome inactivation. We found 1,098 genes in the sequence, of which 99 encode proteins expressed in testis and in various tumour types. A disproportionately high number of mendelian diseases are documented for the X chromosome. Of this number, 168 have been explained by mutations in 113 X-linked genes, which in many cases were characterized with the aid of the DNA sequence.

X- and Y-Chromosomes Recombine at Their Tips



Pseudoautosomal Region 1 (PAR1)

Non-Recombining Portion of Y Chromosome (NRY)

Pseudoautosomal Region 2 (PAR2)

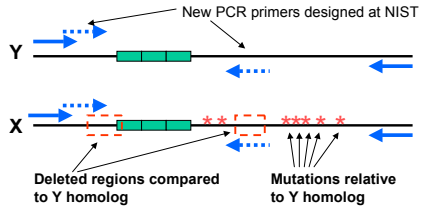
Male-specific region of the Y (MSY)

PCR primers for Y-STRs need to be carefully designed to avoid X-chromosome homology
(e.g., DYS391 in Y-PLEX 6 kit)

Butler, J.M. (2005) *Forensic DNA Typing, 2nd Edition*, Figure 9.4(A), ©Elsevier Science/Academic Press

DYS391 Primer Improvements

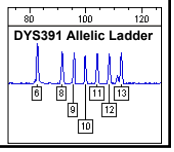
Female artifact problems seen with published and Y-PLEX 6 primers



Significant homology exists between X and Y sequences for the DYS391 locus

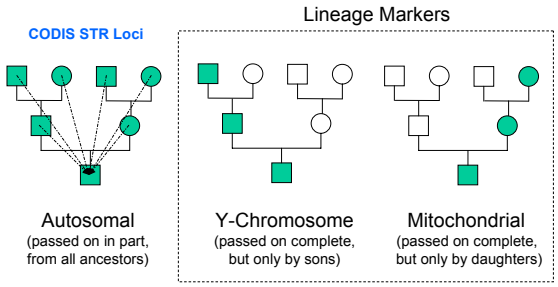
We have designed primers to anneal to regions that only appear on the Y chromosome (targeted X deletion regions)

These primers produce smaller PCR products and were adopted by Promega for their **PowerPlex Y kit**

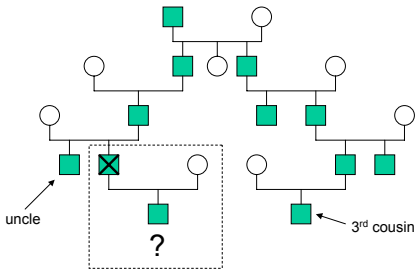


Butler *et al.* (2002) *Forensic Sci. Int.* 129:10-24

Different Inheritance Patterns



Y-STRs permit extension of possible reference samples in missing persons cases



Core Y-STR Loci

History of Y STR Marker Discovery

"Extended Haplotype"

1992 - DYS19 (Roewer et al.)
1994 - YCAI a/b, YCAII a/b, YCAIII a/b, DXYS156 (Mathias et al.)
1996 - DYS389I/II, 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.)

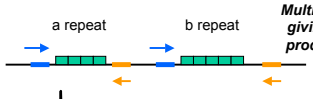
U.S. Haplotype

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)
2003 – DYS597-DYS645 (**50 new Y STRs**; Manfred Kayser GDB entries)

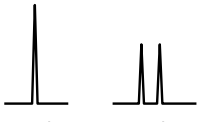
From J.M. Butler (2003) Recent developments in Y-STR and Y-SNP analysis. *Forensic Sci. Rev.* 15:91-111

Y STR Typing of Duplicated Regions

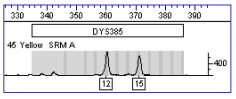
"multi-copy loci"



Multiple primer binding sites occur giving rise to more than one PCR product for a given set of primers

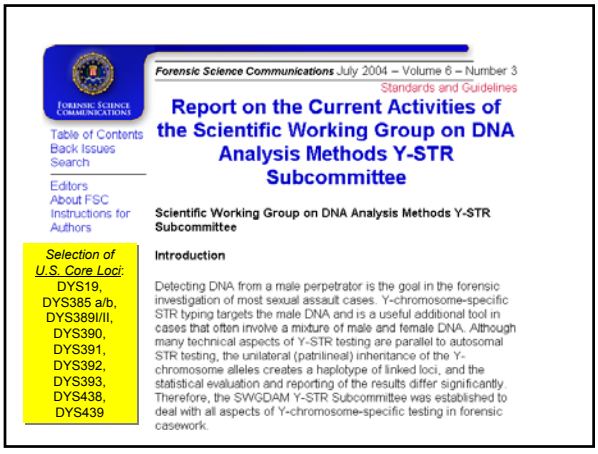


Y-PLEX™ 6 results



DYS385 a/b and YCAII a/b

Y-STR loci are often counted by the number of amplicons rather than the number of PCR primer pairs



SWGDAM Sub-Committee on the Y Chromosome

- Formed in July 2002
- Members
  - Jack Ballantyne (UCF) – chair
  - Mecki Prinz (NYC) – co-chair
  - **John Butler (NIST)**
  - Ann Gross (MN)
  - Jill Smerick (FBI)
  - Sam Baechtel (FBI)
  - John Hartmann (Orange Co.)
  - Jonathan Newman (CFS)
  - Phil Kinsey (OR)
  - Gary Sims (CA DOJ)
  - Bruce Budowle (FBI) – removed in 2004
- **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

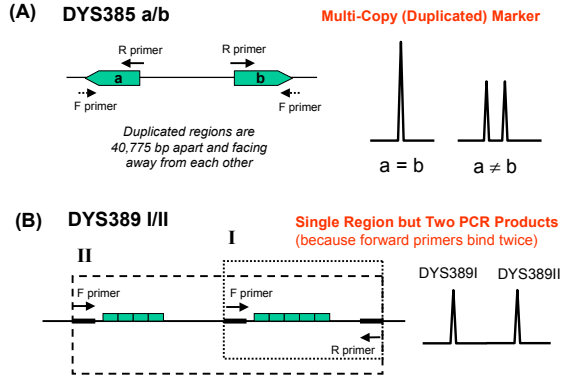
Core Y-STR Characteristics

11 PCR products
9 primer sets

STR Marker	Position (Mb)	Repeat Motif	Allele Range	Mutation Rate
DYS393	3.17	AGAT	8-17	0.05%
DYS19	10.12	TAGA	10-19	0.20%
DYS391	12.54	TCTA	6-14	0.40%
DYS439	12.95	AGAT	8-15	0.38%
DYS389 I/II	13.05	[TCTG] [TCTA]	9-17 / 24-34	0.20%, 0.31%
DYS438	13.38	TTTTC	6-14	0.09%
DYS390	15.71	[TCTA] [TCTG]	17-28	0.32%
DYS385 a/b	19.19, 19.23	GAAA	7-28	0.23%
DYS392	20.97	TAT	6-20	0.05%

Positions in megabases (Mb) along the Y-chromosome were determined with NCBI build 35 (May 2004) using BLAT. Allele ranges represent the full range of alleles reported in the literature. Mutation rates summarized from YHRD (http://www.yhrd.org; accessed 6 Apr 2005).

Butler, J.M. (2005) Genetics and genomics of core STR loci used in human identity testing. *J. Forensic Sci., in press.*

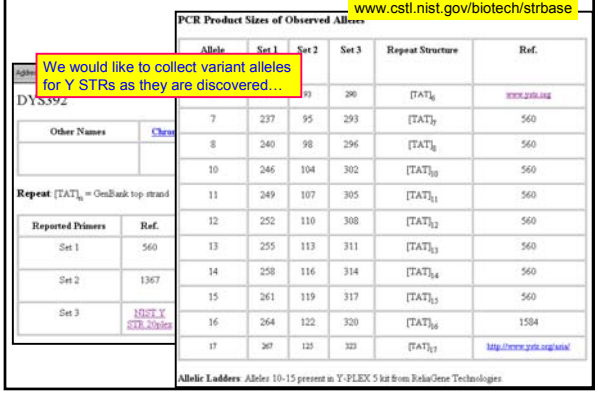


Butler, J.M. (2005) *Forensic DNA Typing, 2nd Edition,* Figure 9.5, ©Elsevier Science/Academic Press

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…



Y-Chromosome Standard NIST SRM 2395

Human Y-Chromosome DNA Profiling Standard

- **5 male samples** + 1 female sample (neg. control)
- **100 ng of each (50 µL at ~2 ng/µL)**
- **22 Y STR markers sequenced**
- **9 additional Y STR markers typed**
- **42 Y SNPs typed** with Marligen kit

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

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

Y-filer - adds DYS635 (C4); now sequenced

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

Sequence Determination of Y STR Repeat Region for Each Component

DYS392 (forward) A — 13 TAT repeats

DYS392 (forward) B — 11 TAT repeats

DYS392 (forward) C — 11 TAT repeats

DYS392 (forward) D — 11 TAT repeats

DYS392 (forward) E — 12 TAT repeats

Sequencing Performed

DYS19	DYS388
DYS385 a/b	DYS426
DYS389 I/II	
DYS390	DYS435
DYS391	DYS436
DYS392	DYS437
DYS393	DYS438
	DYS439
DYS447	DYS460 (A7.1)
DYS448	DYS461 (A7.2)
Y-GATA-H4	DYS462

Y STR Allele Nomenclatures

SRM 2395 as an international standard should help resolve issues

DYS19

[TAGA] 3 TAGG [TAGA] 7-16 — variable

10-19 total

DYS439 *3 different nomenclatures have been published for DYS439*

Now accepted way

[GATA] 2 [GATA] 3 [GATA] 1 [GATA] 1 [GATA] 9-14 — variable

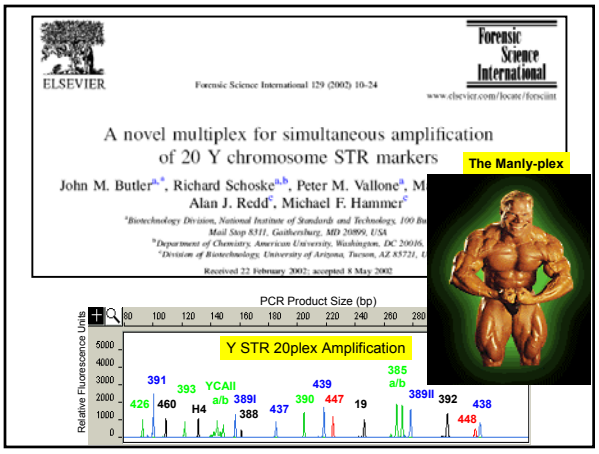
16-21 total

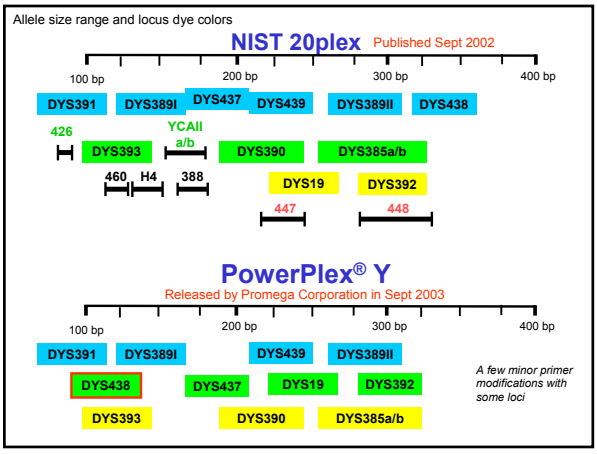
Available Y-STR Kits

Commercial Y-STR Kits

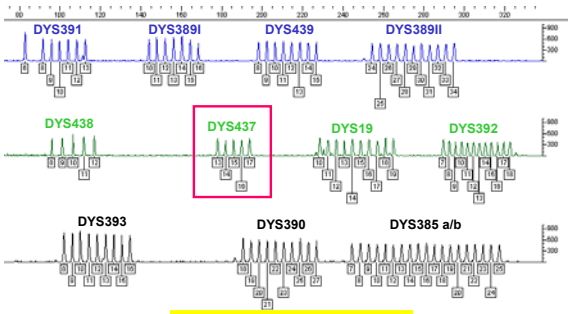
(Minimal/extended haplotype)	(White et al.)	(Ayub et al.)	(Iida et al.)	(Redd et al.)
DYS19	A7.1 (DYS460)	DYS434	DYS441	DYS446
DYS389I/II	A7.2 (DYS461)	DYS435	DYS442	DYS447
DYS390	A10	DYS436	DYS443	DYS448
DYS391	C4	DYS437	DYS444	DYS449
DYS392	H4	DYS438	DYS445	DYS450
DYS393		DYS439		DYS452
DYS385 a/b				DYS453
YCAII a/b	(Bosch et al.) G09411 (DYS462)	43 (51) Y-STRs (217 with Manfred's)		DYS454
DYS388	~~Y-PLEX 6 (ReliaGene)~~			DYS455
DYS425	~~Y-PLEX 5 (ReliaGene)~~			DYS456
DYS426	~~Y-PLEX 12 (ReliaGene)~~			DYS458
YCAIII a/b	PowerPlex Y (Promega)			DYS459 a/b
	Yfiler (Applied Biosystems)			DYS463
				DYS464 a/b/c/d

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





Promega PowerPlex® Y Allelic Ladders



U.S. Core Loci + DYS437

Single amplification; ladders contain 103 alleles

PowerPlex Y Validation Studies

Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Forensic Science International 148 (2005) 1-14
Forensic Science International
www.elsevier.com/locate/forsciint

Validation of a male-specific, 12-locus fluorescent short tandem repeat (STR) multiplex☆

Benjamin E. Krenke[a,*], Lori Viculis[b], Melanie L. Richard[c], Mechthild Prinz[d], Scott C. Milne[e], Carll Ladd[f], Ann Marie Gross[g], Tanis Gornall[d], J. Roger H. Frappier[c], Arthur J. Eisenberg[h], Charles Barna[b], Xavier G. Aranda[h], Michael S. Adamowicz[f], Bruce Budowle[i]

Validation Summary Sheet for PowerPlex Y

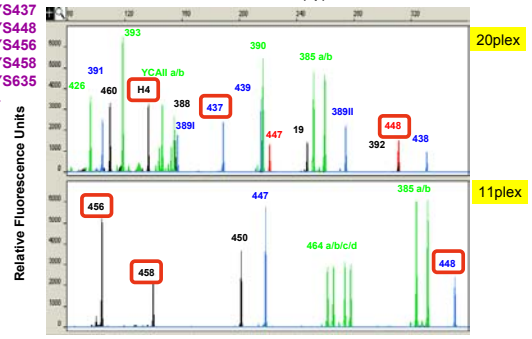
Study Completed (17 studies done)	Description of Samples Tested (performed in 7 labs and Promega)	# Run
Single Source (Concordance)	5 samples x 8 labs	40
Mixture Ratio (male:female)	6 labs x 2 M/F mixture series x 11 ratios (1:0,1:1,1:10,1:100,1:300,1:1000,0.5:300, 0.25:300,0.125:300, 0.0625:300, 0.03:300 ng M:F)	132
Mixture Ratio (male:male)	6 labs x 2 M/M mixtures series x 11 ratios (1:0, 19:1, 9:1, 5:1, 2:1, 1:1, 1:2, 1:5, 1:9, 1:19, 0:1)	132
Sensitivity	7 labs x 2 series x 6 amounts (1/0.5/0.25/0.125/0.06/0.03)	84
Non-Human	24 animals	24
NIST SRM	6 components of SRM 2395	6
Precision (ABI 3100 and ABI 377)	10 ladder replicates + 10 sample replicated + [8 ladders + 8 samples for 377]	36
Non-Probative Cases	65 cases with 102 samples	102
Stutter	412 males used	412
Peak Height Ratio	N/A (except for DYS385 but no studies were noted)	
Cycling Parameters	5 cycles (28/27/26/25/24) x 8 punch sizes x 2 samples	80
Annealing Temperature	5 labs x 5 temperatures (54/58/60/62/64) x 1 sample	25
Reaction volume	5 volumes (50/25/15/12.5/6.25) x [5 amounts + 5 concentrations]	50
Thermal cycler test	4 models (480/2400/9600/9700) x 1 sample + [3 models x 3 sets x 12 samples]	76
Male-specificity	2 females x 1 titration series (0-500 ng female DNA) x 5 amounts each	10
TaqGold polymerase titration	5 amounts (1.38/2.06/2.75/3.44/4.13 U) x 4 quantities (1/0.5/0.25/0.13 ng DNA)	20
Primer pair titration	5 amounts (0.5x/0.75x/1x/1.5x/2x) x 4 quantities (1/0.5/0.25/0.13 ng DNA)	20
Magnesium titration	5 amounts (1/1.25/1.5/1.75/2 mM Mg) x 4 quantities (1/0.5/0.25/0.13 ng DNA)	20

Krenke *et al.* (2005) *Forensic Sci. Int.* 148:1-14 TOTAL SAMPLES EXAMINED **1269**

NIST Multiplexes for High-Throughput Y STR Typing

Laid the Groundwork for Applied Biosystems' **Yfiler Y-STR Kit**

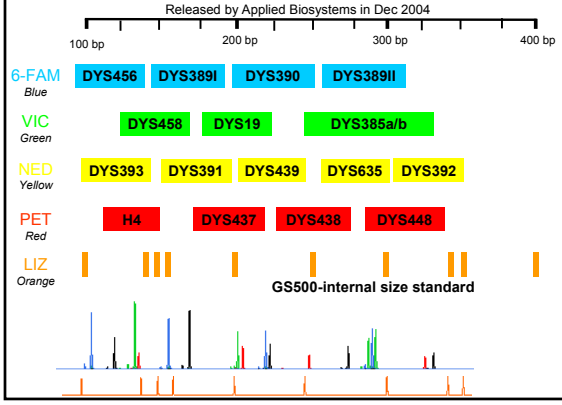
11 U.S. core loci
+ DYS437
+ DYS448
+ DYS456
+ DYS458
+ DYS635
+ H4



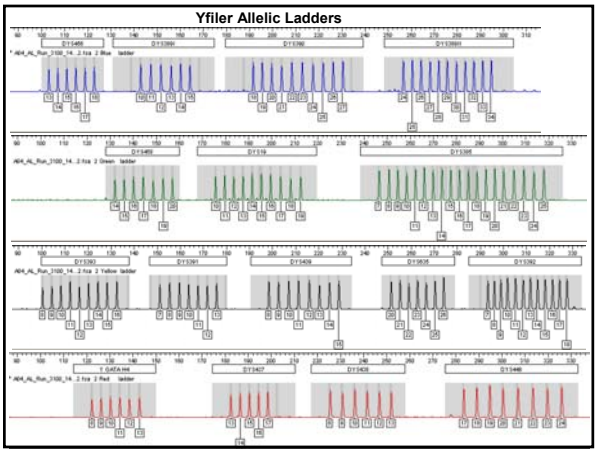
Schoske *et al.* (2004) High-throughput Y-STR typing of U.S. populations…, *Forensic Sci. Int.*, 139:107-121

AmpFℓSTR® Yfiler™

Released by Applied Biosystems in Dec 2004



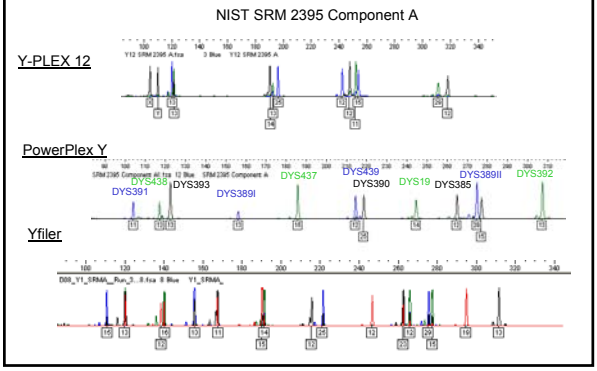
Yfiler Allelic Ladders



Commercial Y-STR Kits Available

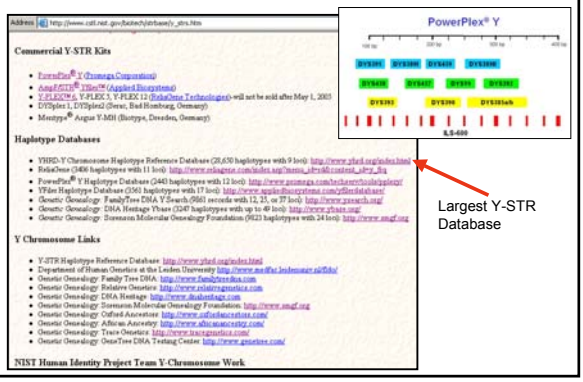
- ReliaGene Technologies (New Orleans, LA)
  - **Y-PLEX™ 6**: DYS19, DYS389II, DYS390, DYS391, DYS393, DYS385 a/b
  - **Y-PLEX™ 5**: DYS389I/II, DYS392, DYS438, DYS439
  - **Y-PLEX™ 12**: DYS19, DYS385 a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS438, DYS439, amelogenin
- Promega Corporation (Madison, WI)
  - **PowerPlex® Y**: DYS19, DYS385 a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS438, DYS439, DYS437, DYS448, DYS456, DYS458, DYS635 (Y-GATA-C4), Y-GATA-H4
- Applied Biosystems (Foster City, CA)
  - **Yfiler™**: DYS19, DYS385 a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS438, DYS439, DYS437, DYS448, DYS456, DYS458, DYS635 (Y-GATA-C4), Y-GATA-H4
- Serac (Bad Homburg, Germany)
  - **genRES® DYSplex-1**: DYS389I/II, DYS390, DYS391, DYS385 a/b, amelogenin
  - **genRES® DYSplex-2**: DYS19, DYS389I/II, DYS392, DYS393
- Biotype (Dresden, Germany)
  - **Mentype® Argus Y-MH**: DYS19, DYS385 a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393

Our Group at NIST Was Involved in Beta-Testing All U.S. Based Y-STR Kits



Available Y-STR Haplotype Databases

New Y-Chromosome Information Resources on STRBase



Y-Chromosome Haplotype Reference Database (YHRD)

http://www.yhrd.org

As of 12/17/04: **28,650 haplotypes**

6,281 haplotypes
with all US required loci

Commercial Y-STR kits exist to amplify all of the core loci in a single reaction (plus a few additional markers)

Run only with minimal haplotype

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

US haplotype requires 2 additional loci:
DYS438
DYS439

Yfiler Haplotype Database

http://www.appliedbiosystems.com/yfilerdatabase/

Population	# Haplotypes
African American	985
Asian	330
Caucasian	1276
Filipino	105
Hispanic	597
Native American	106
Sub-saharan African	59
Vietnamese	103
All	**3561**

#Samples Contributed by NIST

259 African Americans
3 Asians
262 Caucasians

139 Hispanics

Data provided by NIST

663/3561 = **18.6%**

PowerPlex Y Haplotype Database

http://www.promega.com/techserv/tools/pplexy/

595 Caucasians	1311 Caucasians
284 Asians	325 Asians
630 Hispanics	894 Hispanics
577 African Americans	1108 African Americans
357 Native Americans	366 Native Americans
2,443 total	**4,004 total**

March 2005

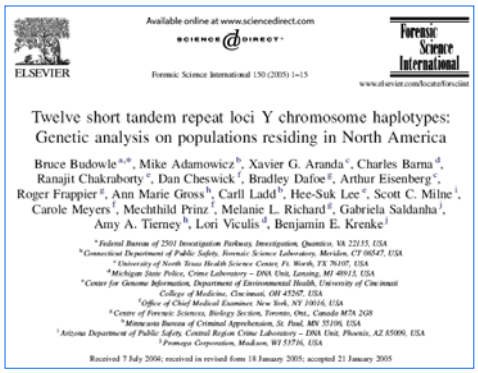
Y-STR Population Studies

A few recent Y-STR population studies

Population	# Samples	# Loci	Reference
5 North American groups	2,443	12	Budowle et al. (2005) FSI 150:1-15
U.S. Caucasians, African Americans, Hispanics	647	22 (27)	Schoske et al. (2004) FSI 139:107-121
Austrian	135	17	Berger et al. (2005) IJLM, in press (Yfiler)
91 European groups	12,700	7	Roewer et al. (2005) Hum Genet 116:279-291

More than 200 Y-STR population studies have been published (most of this data is deposited in the YHRD – Y Chromosome Haplotype Reference Database)

PowerPlex Y Population Study



Available online at www.sciencedirect.com
SCIENCE DIRECT®
Forensic Science International 150 (2005) 1–15
Forensic Science International
www.elsevier.com/locate/forsciint

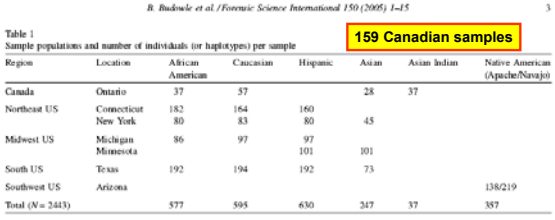
Twelve short tandem repeat loci Y chromosome haplotypes: Genetic analysis on populations residing in North America

Bruce Budowle [a,*], Mike Adamowicz [b], Xavier G. Aranda [c], Charles Barna [d], Ranajit Chakraborty [e], Dan Cheswick [f], Bradley Dafoe [g], Arthur Eisenberg [c], Roger Frappier [g], Ann Marie Gross [h], Carll Ladd [b], Hee-Suk Lee [e], Scott C. Milne [i], Carole Meyers [f], Mechthild Prinz [f], Melanie L. Richard [g], Gabriela Saldanha [j], Amy A. Tierney [h], Lori Viculis [d], Benjamin E. Krenke [j]

[a] Federal Bureau of 2501 Investigation Parkway, Investigation, Quantico, VA 22135, USA
[b] Connecticut Department of Public Safety, Forensic Science Laboratory, Meriden, CT 06547, USA
[c] University of North Texas Health Science Center, Ft. Worth, TX 76107, USA
[d] Michigan State Police, Crime Laboratory – DNA Unit, Lansing, MI 48913, USA
[e] Center for Genome Information, Department of Environmental Health, University of Cincinnati College of Medicine, Cincinnati, OH 45267, USA
[f] Office of Chief Medical Examiner, New York, NY 10016, USA
[g] Centre of Forensic Sciences, Biology Section, Toronto, Ont., Canada M7A 2G8
[h] Minnesota Bureau of Criminal Apprehension, St. Paul, MN 55106, USA
[i] Arizona Department of Public Safety, Central Region Crime Laboratory – DNA Unit, Phoenix, AZ 85009, USA
[j] Promega Corporation, Madison, WI 53716, USA

Received 7 July 2004; received in revised form 18 January 2005; accepted 21 January 2005

Initial 2,443 Samples in PowerPlex Y Haplotype Database



B. Budowle et al. / Forensic Science International 150 (2005) 1–15 3

Table 1
Sample populations and number of individuals (or haplotypes) per sample

159 Canadian samples

Region	Location	African American	Caucasian	Hispanic	Asian	Asian Indian	Native American (Apache/Navajo)
Canada	Ontario	37	57		28	37	
Northeast US	Connecticut	182	164	160			
	New York	80	83	80	45		
Midwest US	Michigan	86	97	97			
	Minnesota			101	101		
South US	Texas	192	194	192	73		
Southwest US	Arizona						138/219
Total (N = 2443)		577	595	630	247	37	357

PowerPlex® Y Haplotype Database
http://www.promega.com/techserv/tools/pplexy/

Compilation of Y STR Population Data

Appendices in Rich Schoske's Ph.D. dissertation; available on STRBase

Locus	Population Ethnicity	Sample Size	Allele Range (Allele Frequencies) 10	11	12	13	14	15	16	17	18	19	Reference
DYS19	EUROPE												
	Innsbruck (Austria)	100				0.14	0.66	0.23	0.07	0.01			Kayser et al. 1997
	Germany	86		0.07			0.512	0.279	0.081	0.046	0.012		Camacedo et al. 2001
	Bavaria (Germany)	151				NL	NL	NL	NL	NL			Anslinger et al. 2000
	Cologne (Germany)	163				0.041	0.5982	0.268	0.0615	0.0615			Hidding et al. 2000
	South Wurtemberg (Germany)	218				0.1009	0.5321	0.2431	0.0826	0.0413			Grave et al. 2000
	Munster (Germany)	272				0.04	0.57	0.23	0.12	0.04			Kayser et al. 1997
	Koln (Germany)	100				0.04	0.62	0.25	0.06	0.03			Kayser et al. 1997
	Jena (Germany)	143				0.06	0.55	0.18	0.15	0.06			Kayser et al. 1997
	Heidelberg (Germany)	113				0.07	0.5	0.29	0.09	0.04			Kayser et al. 1997
	Hanover (Germany)	53				0.07	0.49	0.23	0.13	0.07			Kayser et al. 1997
	Magdeberg (Germany)	210				0.05	0.47	0.25	0.17	0.06			Kayser et al. 1997
	Bradenburg (Germany	233				0.03	0.45	0.27	0.15	0.1			Kayser et al. 1997
	Munich 1 (Germany)	125				0.08	0.46	0.26	0.18	0.03			Kayser et al. 1997
	Munich 2 (Germany)	259				0.16	0.45	0.17	0.19	0.03			Kayser et al. 1997
	Berlin 1(Germany)	233				0.07	0.39	0.27	0.21	0.06			Kayser et al. 1997
	Bremen	49				0.1	0.59	0.16	0.14				Kayser et al. 1997
	Leiden	88				0.04	0.7	0.19	0.03	0.02			Kayser et al. 1997
	Leicester, pooled	339			0.01	0.04	0.46	0.26	0.16	0.07	0.01		Kayser et al. 1997
	British	41				0.05	0.8	0.1	0.02	0.02			Kayser et al. 1997
	Bratislava	57				0.07	0.19	0.21	0.31	0.21			Kayser et al. 1997
	Norway	300				0.027	0.527	0.313	0.1200	0.013			Kayser et al. 1997

- **Source: over 200 published population data papers**
- Helps define observed allele ranges, which aids in multiplex assay development (spacing between loci in the same dye color)
- Information is available to the community through the STRBase website – ***permits analysis of optimal markers for particular population***

Richard Schoske Dissertation

http://www.cstl.nist.gov/biotech/strbase/pub_pres/Schoske2003dis.pdf

Rich Schoske

PhD student from American University
Funded by Air Force

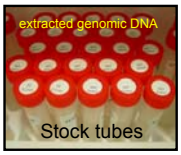
- Worked at NIST from Nov 2000 to May 2003
- **270 page Ph.D. dissertation**
- Entitled "The design, optimization and testing of Y chromosome short tandem repeat megaplexes."
- Available for download on NIST STRBase website

Standard U.S. Population Dataset

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

260 Caucasians, 260 African Americans, 140 Hispanics, 3 Asians = **663 males**

DNA extracted from whole blood (anonymous; self-identified ethnicities) received from Interstate Blood Bank (Memphis, TN) and Millennium Biotech Inc. (Ft. Lauderdale, FL)

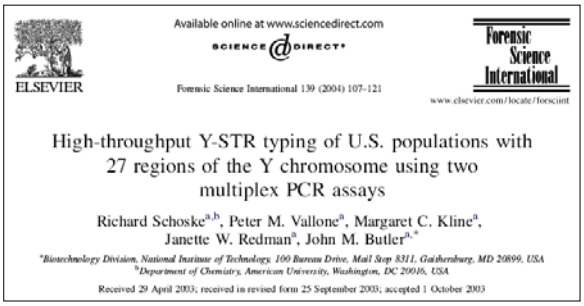


To date: (~95,000 allele calls)

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 STRs 27 new loci (14,535)
Yfiler kit 17 loci (11,237)
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)
mtDNA full control region sequences by AFDIL

Genotypes with various human identity testing markers

U.S. Population Data on 22 Y-STRs



ELSEVIER

Available online at www.sciencedirect.com
SCIENCE DIRECT®

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

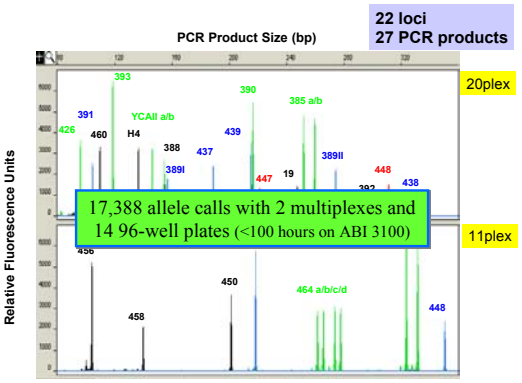
Richard Schoske[a,b], Peter M. Vallone[a], Margaret C. Kline[a], Janette W. Redman[a], John M. Butler[a,*]

[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 Multiplexes for High-Throughput Y STR Typing



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	0.560	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*

Statistical Calculations on Y-STR Data

- **Locus (gene) Diversity** = $(n/n-1)(1 - \Sigma p_i^2)$ where n is the number of samples in the dataset and p_i is the frequency of the i^{th} allele
- **Haplotype Diversity** **(HD) = $(n/n-1)(1 - \Sigma p_i^2)$** where n is the number of samples in the dataset and p_i is the frequency of the i^{th} haplotype
- **Random Match Probability** (RMP) = 1 - HD
- **Discrimination Capacity** (DC) – total number of observed haplotypes divided by the total number of individuals in the dataset
- **Unique Haplotypes** (UH) – number of haplotypes that occur only once in the dataset

Y-STR Marker Combinations	260 African Americans		244 Caucasians		143 Hispanics	
	HD	RMP	HD	RMP	HD	RMP
Y-PLEX 6 kit	0.9974	0.0026	0.9914	0.0086	0.9934	0.0066
"minimal" haplotype	0.9982	0.0018	0.9946	0.0053	0.9957	0.0043
"extended" haplotype	0.9988	0.0012	0.9971	0.0029	0.9975	0.0025
"U.S. haplotype"	0.9993	0.0007	0.9974	0.0026	0.9986	0.0014
Y-STR 11plex	0.9993	0.0007	0.9987	0.0013	0.9992	0.0008
Y-STR 20plex	0.9998	0.0002	0.9998	0.0002	0.9998	0.0002
22 Y-STRs	0.9999	0.0001	0.9999	0.0001	0.9999	0.0001
Top 10 (w/o YCAII a/b)	0.9999	0.0001	0.9999	0.0001	0.9999	0.0001

Y-STR Marker Combinations	260 African Americans		244 Caucasians		143 Hispanics	
	DC	UH	DC	UH	DC	UH
Y-PLEX 6 kit	82.3%	188	68.9%	136	78.3%	97
"minimal" haplotype	88.5%	213	75.8%	161	81.1%	100
"extended" haplotype	91.9%	227	83.6%	184	89.5%	120
"U.S. haplotype"	91.9%	222	82.3%	176	93.3%	121
Y-STR 11plex	93.1%	227	88.5%	198	94.4%	127
Y-STR 20plex	98.5%	252	97.2%	230	98.6%	139
22 Y-STR Markers	98.9%	254	99.6%	242	99.3%	141
Top 10 (w/o YCAII a/b)	96.9%	244	97.5%	232	99.3%	141

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

Statistics with Y-STR Haplotypes

Most labs will probably go with the **counting method** (number of times a haplotype is observed in a database) as is typically done with mtDNA results

Example Y-STR Haplotype

Core US Haplotype

- DYS19 – 14
- DYS389I – 13
- DYS389II – 29
- DYS390 – 24
- DYS391 – 11
- DYS392 – 14
- DYS393 – 13
- DYS385 a/b – 11,15
- DYS438 – 12
- DYS439 – 13

Matches by Databases

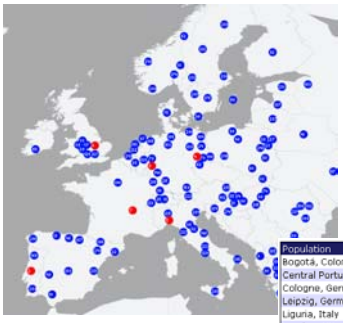
- YHRD (9 loci)
  - 7 matches in 27,773
- YHRD (11 loci)
  - 0 matches in 6,281
- ReliaGene (11 loci)
  - 0 matches in 3,403
- PowerPlex Y (12 loci)
  - 0 matches in 4,004
- Yfiler (17 loci)
  - 0 matches in 3,561

Y-Chromosome Haplotype Reference Database

www.YHRD.org

Release "15" from 2004-12-17 16:11:24

7 matches in 27,773 individuals from 236 worldwide populations



Population	#	Metapopulation
Bogotá, Colombia [European]	1 / 147	Eurasian MP / European MP
Central Portugal	1 / 230	Eurasian MP / European MP
Cologne, Germany	1 / 135	Eurasian MP / European MP
Leipzig, Germany	1 / 661	Eurasian MP / European MP
Liguria, Italy	1 / 81	Eurasian MP / European MP
London, UK	1 / 285	Eurasian MP / European MP
Lyon, France	1 / 125	Eurasian MP / European MP

Minimal Haplotype Result

DYS19 – 14
DYS389I – 13
DYS389II – 29
DYS390 – 24
DYS391 – 11
DYS392 – 14
DYS393 – 13
DYS385 a/b – 11,15

Frequency Estimate Calculations

In cases where a Y-STR profile is observed a particular number of times (X) in a database containing N profiles, its frequency (p) can be calculated as follows:

$$p = X/N$$

7 matches in 27,773

p = 7/27,773 = 0.000252 = **0.025%**

An upper bound confidence interval can be placed on the profile's frequency using:

$$p + 1.96\sqrt{\frac{(p)(1-p)}{N}}$$

$$0.000252 + 1.96\sqrt{\frac{(0.000252)(1-0.000252)}{27{,}773}}$$

= 0.000252 + 0.000187 = 0.000439
= **0.044% (~1 in 2270)**

When there is no match…

In cases where the profile has not been observed in a database, the upper bound on the confidence interval is

$$1-\alpha^{1/N}$$

0 matches in 4,004

where α is the confidence coefficient (0.05 for a 95% confidence interval) and N is the number of individuals in the database.

$1-\alpha^{1/N} = 1-(0.05)^{[1/4,004]} = 0.000748$
= **0.075% (~1 in 1340)**

If using database of 2,443, then the best you can do is 1 in 816

The Meaning of a Y-Chromosome Match

Conservative statement for a match report:

The Y-STR profile of the crime sample matches the Y-STR profiles of the suspect (at *xxx* number of loci examined). Therefore, ***we cannot exclude the suspect*** as being the donor of the crime sample. In addition, we cannot exclude all patrilineal related male relatives and an unknown number of unrelated males as being the donor of the crime sample.

Difficult Questions…

- Which database(s) should be used for Y-STR profile frequency estimate determination?
- Are any of the current forensic Y-STR databases truly adequate for reliable estimations of Y-STR haplotype frequencies?
  - Some individuals share identical Y-STR haplotypes due to recurrent mutations, not relatedness…
  - Is the database a random collection reflecting Y-STR haplotype frequencies of the population?
  - Is the Y-STR haplotype frequency relevant for the population of the suspect?

Issues raised by Peter de Knijff at his Promega meeting presentation (Oct 2004)

Conclusions from Peter de Knijff

From his presentation at the Promega meeting (Oct 2004)

A haplotype frequency taken from any Y-STR database should not be reported or seen as a random match probability

- Because all male relatives have the same haplotype
- Males can share haplotypes without being related

Database estimates are at most qualitative…

What Peter de Knijff Reports with a Y-STR Match

From his presentation at the Promega meeting (Oct 2004)

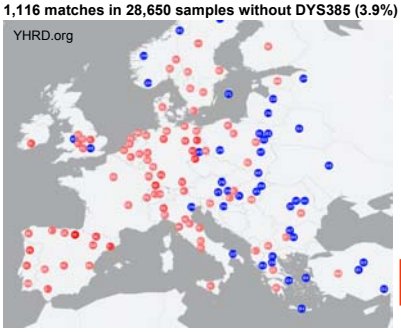
- The Y-STR profile of the stain matches with the suspect.
- Therefore, the suspect cannot be excluded as the donor of the stain.
- On the basis of this DNA evidence, I **can also not exclude all paternally related male relatives of the suspect** as possible donors of this stain.
- In addition, **an unknown number of males from the same region cannot be excluded**. A more accurate answer can only be obtained if (1) we have detailed knowledge of the population structure of the region of interest, (2) the Y-STR frequencies therein are known, and (3) we have knowledge about the family structure of the suspect.

New Y-STRs

More than 150 new Y-STR loci were characterized in June 2004. These new loci need to be studied in common sample sets including U.S. population groups in order to understand their ability to differentiate most common types and closely related individuals.

Most Common Type in Europeans

1,116 matches in 28,650 samples without DYS385 (3.9%)



Most common type

DYS19 – 14
DYS389I – 13
DYS389II – 29
DYS390 – 24
DYS391 – 11
DYS392 – 13
DYS393 – 13
DYS385 a/b – 11,14

Need more Y-STR loci to separate…

606 matches in 27,773 samples (2.1%)

Most Common Type seen in 22 NIST samples (3.7%)
(from all 3 populations)

Samples	Minimal Haplotype
PT84633	11,14-29-24-14-11-13-13-13
PT83902	11,14-29-24-14-11-13-13-13
PT84244	11,14-29-24-14-11-13-13-13
PT83874	11,14-29-2 13-13-13
MT94875	11,14-29-2 13-13-13
PT83535	11,14-29-2 13-13-13
ZT80028	11,14-29-2 13-13-13
MT96356	11,14-29-2 13-13-13
PT84236	11,14-29-2 13-13-13
GT37692	11,14-29-2 13-13-13
PT83863	11,14-29-2 13-13-13
PT84252	11,14-29-24-14-11-13-13-13
OT05562	11,14-29-24-14-11-13-13-13
PT83885	11,14-29-24-14-11-13-13-13
MT97185	11,14-29-24-14-11-13-13-13
OT07753	11,14-29-24-14-11-13-13-13
TT51702	11,14-29-24-14-11-13-13-13
UC10177	11,14-29-24-14-11-13-13-13
WT51359	11,14-29-24-14-11-13-13-13
WT52486	11,14-29-24-14-11-13-13-13
ZT80656	11,14-29-24-14-11-13-13-13
MT97163	11,14-29-24-14-11-13-13-13

1

With SWGDAM US core loci (minimal haplotype+438,439)
most common type breaks into 3 groups

Samples	Minimal Haplotype	438/439
PT84633	11,14-29-24-14-11-13-13-13	-12-11
PT83902	11,14-29-24-14-11-13-13-13	-12-11
PT84244	11,14-29-24-14-11-13-13-13	-12-11
PT83874	11,14-29-24-14-11-13-13-13	-12-11
MT94875	11,14-29-24-14-11-13-13-13	-12-11
PT83535	11,14-29-24-14-11-13-13-13	-12-11
ZT80028	11,14-29-24-14-11-13-13-13	-12-11
MT96356	11,14-29-24-14-11-13-13-13	-12-12
PT84236	11,14-29-24-14-11-13-13-13	
GT37692	11,14-29-24-14-11-13-13-13	
PT83863	11,14-29-24-14-11-13-13-13	
PT84252	11,14-29-24-14-11-13-13-13	
OT05562	11,14-29-24-14-11-13-13-13	
PT83885	11,14-29-24-14-11-13-13-13	
MT97185	11,14-29-24-14-11-13-13-13	
OT07753	11,14-29-24-14-11-13-13-13	
TT51702	11,14-29-24-14-11-13-13-13	-12-12
UC10177	11,14-29-24-14-11-13-13-13	-12-12
WT51359	11,14-29-24-14-11-13-13-13	-12-12
WT52486	11,14-29-24-14-11-13-13-13	-12-13
ZT80656	11,14-29-24-14-11-13-13-13	-12-13
MT97163	11,14-29-24-14-11-13-13-13	-12-13

3

With Promega's loci (minimal haplotype+438,439,437)
most common type breaks into 7 groups

Samples	Minimal Haplotype	438/439	437
PT84633	11,14-29-24-14-11-13-13-13	-12-11	-14
PT83902	11,14-29-24-14-11-13-13-13	-12-11	-14
PT84244	11,14-29-24-14-11-13-13-13	-12-11	-15
PT83874	11,14-29-24-14-11-13-13-13	-12-11	-15
MT94875	11,14-29-24-14-11-13-13-13	-12-11	-15
PT83535	11,14-29-24-14-11-13-13-13	-12-11	-15
ZT80028	11,14-29-24-14-11-13-13-13	-12-11	-16
MT96356	11,14-29-24-14-11-13-13-13	-12-12	-14
PT84236	11,14-29-24-14-11-13-13-13	-12-12	-14
GT37692	11,14-29-24-14-11-13-13-13	-12-12	-15
PT83863	11,14-29-24-14-11-13-13-13	-12-12	-15
PT84252	11,14-29-24-14-11-13-13-13	-12-12	-15
OT05562	11,14-29-24-14-11-13-13-13	-12-	
PT83885	11,14-29-24-14-11-13-13-13	-12-	
MT97185	11,14-29-24-14-11-13-13-13	-12-	
OT07753	11,14-29-24-14-11-13-13-13	-12-	
TT51702	11,14-29-24-14-11-13-13-13	-12-	
UC10177	11,14-29-24-14-11-13-13-13	-12-	
WT51359	11,14-29-24-14-11-13-13-13	-12-	
WT52486	11,14-29-24-14-11-13-13-13	-12-13	-14
ZT80656	11,14-29-24-14-11-13-13-13	-12-13	-15
MT97163	11,14-29-24-14-11-13-13-13	-12-13	-15

7

With NIST 27 Y STRs (2 multiplexes) most common type
breaks into 22 different groups (*all samples differentiated*)

Samples	Minimal Haplotype	438/439	437	YCAIIa/b	464a/b/c/d	388,426,447,448,450,456,458,460,H4
PT84633	11,14-29-24-14-11-13-13-13	-12-11	-14	-19,23	15,16,17,	-12-12-25-22-9-15-17-11-12
PT83902	11,14-29-24-14-11-13-13-13	-12-11	-14	-19,23	14,16,17,	-12-12-25-21-9-15-17-11-11
PT84244	11,14-29-24-14-11-13-13-13	-12-11	-15	-19,23	14,15,17,	-12-12-25-22-9-15-17-11-12
PT83874	11,14-29-24-14-11-13-13-13	-12-11	-15	-19,23	15,16,17,	-12-12-25-22-9-16-17-11-12
MT94875	11,14-29-24-14-11-13-13-13	-12-11	-15	-19,23	15,16,18,	-12-12-24-22-9-17-17-11-11
PT83535	11,14-29-24-14-11-13-13-13	-12-11	-15	-19,23	16,,,	-12-12-24-21-9-16-17-11-12
ZT80028	11,14-29-24-14-11-13-13-13	-12-11	-16	-19,23	15,17,18,	-12-12-25-22-9-16-17-11-11
MT96356	11,14-29-24-14-11-13-13-13	-12-12	-14	-22,23	15,16,17,	-12-12-25-21-9-15-18-11-11
PT84236	11,14-29-24-14-11-13-13-13	-12-12	-14	-19,23	15,16,,	-12-12-25-22-9-16-17-11-13
GT37692	11,14-29-24-14-11-13-13-13	-12-12	-15	-19,23	14,15,16,18	-12-12-26-22-9-16-16-11-13
PT83863	11,14-29-24-14-11-13-13-13	-12-12	-15	-19,23	14,15,17,	-12-12-25-22-9-15-17-11-11
PT84252	11,14-29-24-14-11-13-13-13	-12-12	-15	-19,23	15,16,17,	-12-12-25-22-9-16-18-11-12
OT05562	11,14-29-24-14-11-13-13-13	-12-12	-15	-19,23	15,16,17,	-12-12-26-23-9-16-17-12-12
PT83885	11,14-29-24-14-11-13-13-13	-12-12	-15	-20,23	15,17,,	-12-12-25-22-9-17-17-12-12
MT97185	11,14-29-24-14-11-13-13-13	-12-12	-15	-23,23	15,16,	-16-17-10-12
OT07753	11,14-29-24-14-11-13-13-13	-12-12	-15	-19,23	15,17,	-15-17-11-12
TT51702	11,14-29-24-14-11-13-13-13	-12-12	-15	-19,23	14,15,1	-16-17-11-12
UC10177	11,14-29-24-14-11-13-13-13	-12-12	-15	-19,23	15,16,1	-15-17-11-12
WT51359	11,14-29-24-14-11-13-13-13	-12-12	-15	-19,23	15,16,1	-15-18-11-12
WT52486	11,14-29-24-14-11-13-13-13	-12-13	-14	-19,23	14,17,	-16-18-11-12
ZT80656	11,14-29-24-14-11-13-13-13	-12-13	-15	-19,24	15,17,	-14-17-11-12
MT97163	11,14-29-24-14-11-13-13-13	-12-13	-15	-19,23	15,17,1	-15-17-11-13

22

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**

NIST Work with New Y-STR Loci

DTD 5
ARTICLE IN PRESS
Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Forensic Science International xxx (2005) xxx–xxx
Forensic Science International
www.elsevier.com/locate/forsciint

Announcement of population data

Allele frequencies for 27 Y-STR loci with U.S. Caucasian, African American, and Hispanic samples

John M. Butler*, Amy E. Decker, Peter M. Vallone, Margaret C. Kline

Biotechnology Division, National Institute of Standards and Technology, Gaithersburg, MD 20899-8311, USA

Received 26 January 2005; received in revised form 22 February 2005; accepted 22 February 2005

Abstract

A total of 263 U.S. Caucasians, 260 African Americans and 140 U.S. Hispanics or a subset of 31 Caucasians, 32 African Americans, and 32 Hispanics were typed for 27 Y-chromosome short tandem repeat (Y-STR) markers: DYS444, DYS446, DYS449, DYS463, DYS485, DYS490, DYS495, DYS504, DYS505, DYS508, DYS520, DYS522, DYS525, DYS532, DYS533, DYS534, DYS540, DYS556, DYS557, DYS570, DYS575, DYS576, DYS594, DYS632, DYS635, DYS641, and DYS643. Allele frequencies for each locus are reported along with nomenclature based on sequence analysis.

Number of Alleles Seen with Various Y-STR Loci in Same Set of 95 U.S. Samples

Set	Locus ID	Dye	Length Range (bp)	# of alleles	
B	DYS632	VIC	153-157	2	←
G	DYS575	FAM	221-225	2	←
E	DYS556	FAM	214-222	3	
E	DYS641	NED	222-230	**3**	
G	DYS540	NED	265-277	4	
J	DYS522	VIC	359-370	4	
A	DYS495	FAM	133-146	5	
G	DYS594	VIC	264-289	5	
C	DYS505	FAM	168-187	**6**	
C	DYS508	VIC	180-200	**6**	
C	DYS520	NED	181-201	6	
E	DYS533	VIC	209-230	6	
J	DYS446	FAM	298-325	**6**	
H	DYS444	NED	299-325	6	
B	DYS643	FAM	135-167	7	
D	DYS534	NED	202-227	**7**	
D	DYS557	VIC	196-219	7	
H	DYS525	VIC	303-328	7	
J	DYS532	NED	466-490	7	
D	DYS576	FAM	178-205	8	
A	DYS485	NED	138-164	8	
H	DYS504	FAM	270-303	9	
F	DYS463	FAM	227-277	9	←
F	DYS570	NED	251-288	10	←
A	DYS490	VIC	126-158	10	
B	DYS449	NED	341-381	12	←

Loci Not Pursued Further…

Low Number of Alleles
DYS632
DYS575

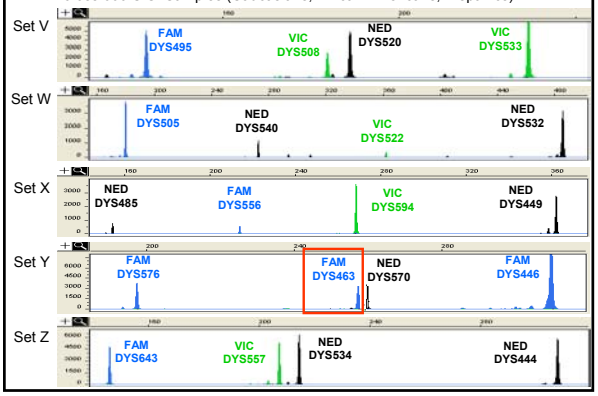
Primers Gave Artifacts in Female
DYS490 – duplicated and on chr X
DYS504
DYS525
DYS557

14,535 types generated across 27 loci

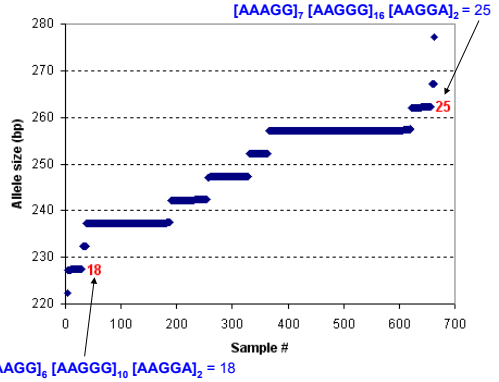
Combo 1 plate
31 Caucasians
32 African Americans
32 Hispanics

5 Panels Used to Type 20 New Y-STRs

Across 660 U.S. Samples (Caucasians, African Americans, Hispanics)



DYS463 Alleles Observed



Conversion to Allele Frequency Information

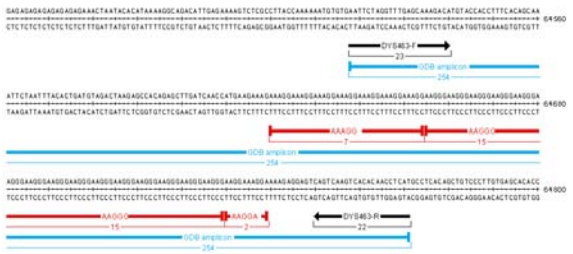
Locus	Allele	Size Range (bp)	Count	Combined Freq (N = 661)
DYS463	17	222.45	1	0.0015
	18	227.34-227.44	27	0.0408
	19	232.30-232.39	7	0.0106
	20	237.24-237.44	151	0.2284
	21	242.21-242.41	67	0.1014
	22	247.12-247.40	74	0.1120
	23	252.13-252.33	35	0.0530
	24	257.05-257.49	256	**0.3873**
	25	262.01-262.26	37	0.0560
	26	267.05-267.21	5	0.0076
	28	277.22	1	0.0015
		failure	2	
				STR diversity
		TOTAL	**661**	**0.7684**

$D = (n/n-1)(1 - \sum x^2)$

DYS463

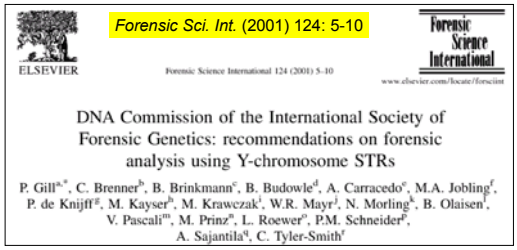
GenBank AC007275 reference allele = 24 (7+15+2) repeats

Nomenclature from Redd *et al.* (2002) *Forensic Sci. Int.* 130: 97-111



CAACCATGAAGAAAG **[AAAGG]7 [AAGGG]15 [AAGGA]2** AAAGAGGAGT

(H09_GT37888) = 18 = **[AAAGG]6 [AAGGG]10 [AAGGA]2**
(G10_TT51399) = 25 = **[AAAGG]7 [AAGGG]16 [AAGGA]2**



ELSEVIER

Forensic Sci. Int. (2001) 124: 5-10

Forensic Science International

Forensic Science International 124 (2001) 5–10

www.elsevier.com/locate/forsciint

DNA Commission of the International Society of Forensic Genetics: recommendations on forensic analysis using Y-chromosome STRs

P. Gill[a,*], C. Brenner[b], B. Brinkmann[c], B. Budowle[d], A. Carracedo[e], M.A. Jobling[f], P. de Knijff[g], M. Kayser[h], M. Krawczak[i], W.R. Mayr[j], N. Morling[k], B. Olaisen[l], V. Pascali[m], M. Prinz[n], L. Roewer[o], P.M. Schneider[p], A. Sajantila[q], C. Tyler-Smith[r]

ISFG Guidelines for Y STRs

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

ISFG Updated Y-STR Recommendations

DTD 5
ARTICLE IN PRESS
Available online at www.sciencedirect.com
SCIENCE DIRECT
Forensic Science International xxx (2005) xxx–xxx
www.elsevier.com/locate/forsciint
ELSEVIER
Forensic Science International
Short communication

DNA Commission of the International Society of Forensic Genetics (ISFG): An update of the recommendations on the use of Y-STRs in forensic analysis☆

L. Gusmão[a], J.M. Butler[b], A. Carracedo[c], P. Gill[d], M. Kayser[e], W.R. Mayr[f], N. Morling[g], M. Prinz[h], L. Roewer[i], C. Tyler-Smith[j], P.M. Schneider[k,*]

[a] IPATIMUP, Porto, Portugal
[b] National Institute of Standards and Technology, Gaithersburg, MD, USA

Probability of Finding No Mutation or at Least One Mutation Between Two Y-STR Haplotypes in a Single Generation

Using average mutation rate of 0.28% (Kayser et al. AJHG 2000, 66:1580-1588)

# STRs	Prob. no mutation	Prob. at least one mutation	
1	0.99720000	0.00280000	
2	0.99440784	0.00559216	
3	0.99162350	0.00837650	
4	0.98884695	0.01115305	
5	0.98607818	0.01392182	
6	0.98331716	0.01668284	
7	0.98056387	0.01943613	
8	0.97781829	0.02218171	
9	0.97508040	0.02491960	
10	0.97235018	0.02764982	
11	0.96962760	0.03037240	
12	0.96691264	0.03308736	**3.3% with 12 Y-STRs**
…			
40	0.89390382	**0.10609618**	

Gusmao *et al.* (2005) *Forensic Sci. Int., in press*

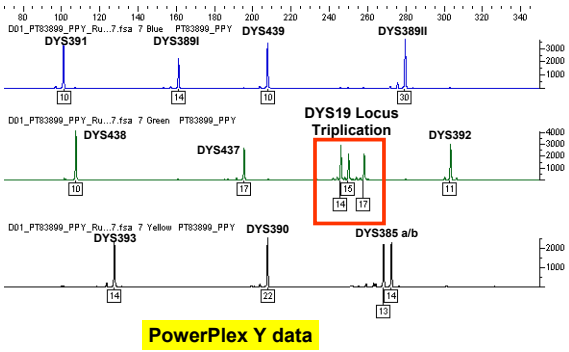
Separating Brothers with 47 Y-STRs

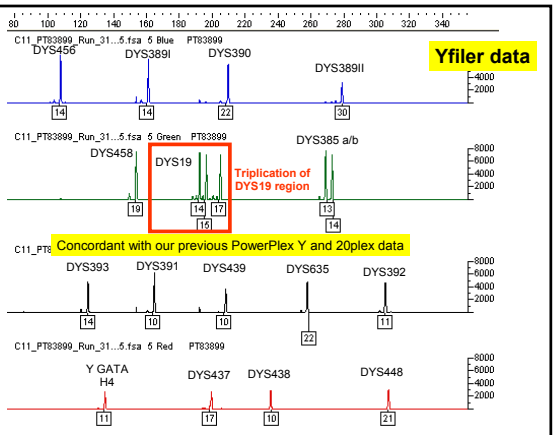
- Two suspected brothers (ZT79338 and ZT79339) are part of our ~660 U.S. sample dataset at NIST.
- Thus far, we have generated 47 Y-STR allele calls on these samples.
- **A mutation at DYS391 separates these individuals** (one contains allele 11 and the other allele 10).
- These samples share autosomal STR alleles and contain identical mtDNA sequences.

Locus Duplication and Deletion

Events that impact Y-STR interpretation

Locus Triplication at DYS19





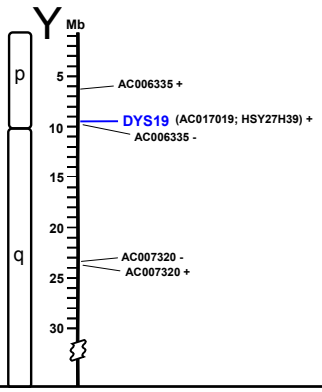
DYS19 Flanking Region Duplication

Int. J. Legal Med. (2004) 118: 178-183

TECHNICAL NOTE

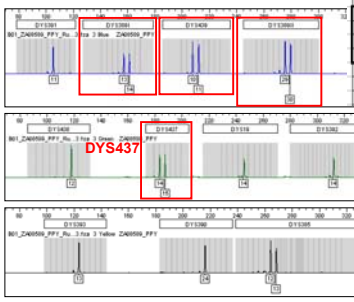
John M. Butler · Richard Schoske

Duplication of DYS19 flanking regions in other parts of the Y chromosome



Different primers around DYS19 repeat result in selection of different regions of the Y-chromosome

Duplication at Multiple Loci with Single-Source Sample



Y-chromosome mapping

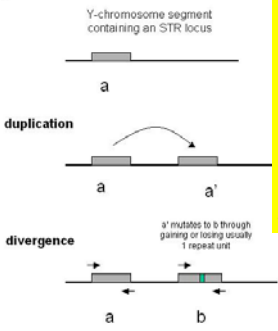
Y STR Marker	Position (Mb)
DYS391	13.413
DYS635 (C4)	13.690
DYS434	13.777
DYS437	13.778
DYS435	13.807
DYS439	13.826
DYS389 I/II	13.923
DYS388	14.057
DYS442	14.071
DYS438	14.248

Entire region of Y-chromosome has likely been duplicated and then diverged

Most duplications have a single repeat spread in allele patterns

Butler et al. (2005), *J. Forensic Sci., in press (July issue)*

Duplication and Divergence Model



Locus	# dup*	>1 repeat
DYS19	23	2
DYS389I	5	0
DYS389II	9	2
DYS390	1	0
DYS391	3	1
DYS392	0	0
DYS393	3	0
DYS385a/b	17	0

*from www.yhrd.org, literature, and our work

92% have single repeat difference

Since single-step mutations are most common, then single repeat spacing in duplicated alleles is expected

Butler *et al.* (2005), *JFS, in press*

Deciphering between a Mixture of Multiple Males and Locus Duplication

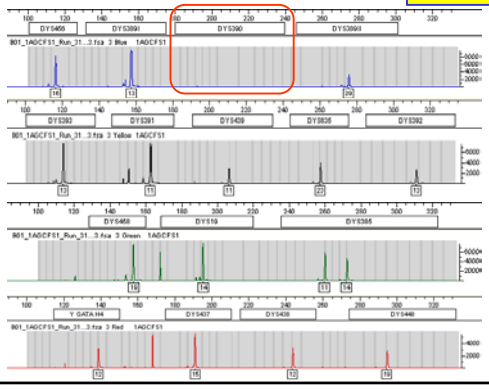
- ***Note the number of loci containing >1 allele (other than multi-copy DYS385)***
- ***Consider relative position on the Y-chromosome if multiple loci have two alleles***
- ***See if repeat spread is >1 repeat unit***
- ***Examine DYS385 for presence of >2 alleles***

Locus duplication along the Y-chromosome is in many ways analogous to heteroplasmy in mitochondrial DNA, which depending on the circumstances can provide greater strength to a match between two DNA samples.

Butler et al. (2005), *J. Forensic Sci., in press (July issue)*

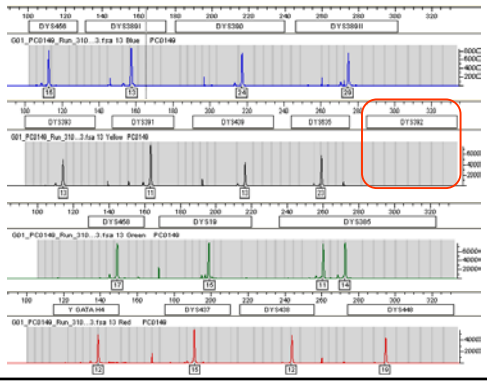
CFS-1 with Y filer

DYS390 is deleted



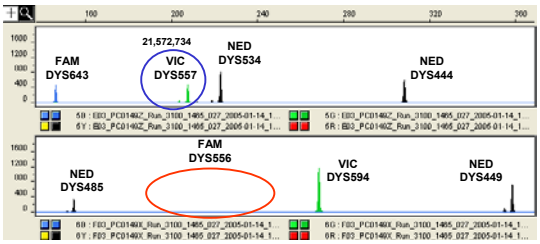
Sample PC0149 with Yfiler

DYS392 is deleted



PC0149 with Additional Y-STRs

One of the closest available loci fails



DYS556 ~32,000 bp away from DYS392 is **missing**

DYS557 ~600,000 bp away from DYS392 is present

Deletions of some Y-STRs can be an inadvertent diagnosis of male infertility

King *et al.* (2005) Inadvertent diagnosis of male infertility through genealogical DNA testing. *J. Med. Genet.* 42:366-368

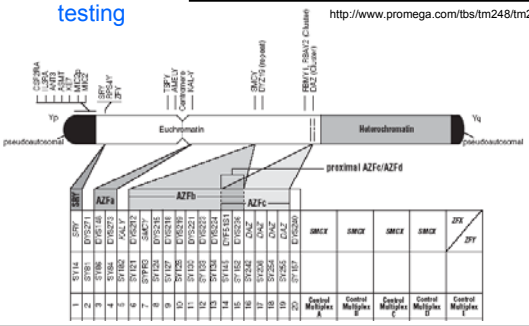
- ***AZFa* deletion** (<1 in 100,000 men): expected to lack **DYS389I/II, DYS437, DYS438, DYS439**
- ***AZFb* deletion** (very rare): expected to lack **DYS385 and DYS392**
- ***AZFc* deletion** (1 in 4,000 men): expected to lack DYS464
- Possible that "incomplete" haplotypes are not being submitted to the Y-STR haplotype databases
- **Thus, Y-STRs are not neutral with respect to fertility information**

Promega sells a Y-deletion test for infertility testing

Y Chromosome Deletion Detection System, Version 2.0

Promega

http://www.promega.com/tbs/tm248/tm248.pdf



Y-SNPs

Recent Forensic Science Service Work with Y-SNPs

ARTICLE IN PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT

Forensic Science International

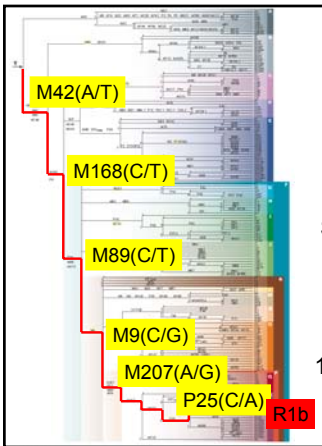
ELSEVIER

Inferring the population of origin of DNA evidence within the UK by allele-specific hybridization of Y-SNPs

Jon H. Wetton*, Kam W. Tsang, Haroon Khan

Abstract

Marked differences in Y-SNP allele frequencies between continental populations can be used to predict the biogeographic origin of a man's ancestral paternal lineage. Using 627 samples collected from individuals within the UK with pale-skinned Caucasian, dark-skinned Caucasian, African/Caribbean, South Asian, East Asian or Middle Eastern appearance **we demonstrate that an individual's Y-SNP haplogroup is also strongly correlated with their physical appearance.** Furthermore, experimental evaluation of the Marligen SignetTM Y-SNP kit in conjunction with the Luminex 100 detection instrument indicates that reliable and reproducible haplogrouping results can be obtained from 1 ng or more of target template derived from a variety of forensic evidence types including, blood, saliva and post-coital vaginal swabs. The test proved highly male-specific with reliable results being generated in the presence of a 1000-fold excess of female DNA, and no anomalous results were observed during degradation studies despite a gradual loss of typable loci. Hence, **Y-SNP haplogrouping has considerable potential forensic utility in predicting likely ethnic appearance.**



The Y Chromosome Consortium Map (2003) Nat Rev Genet. 4 :598-612

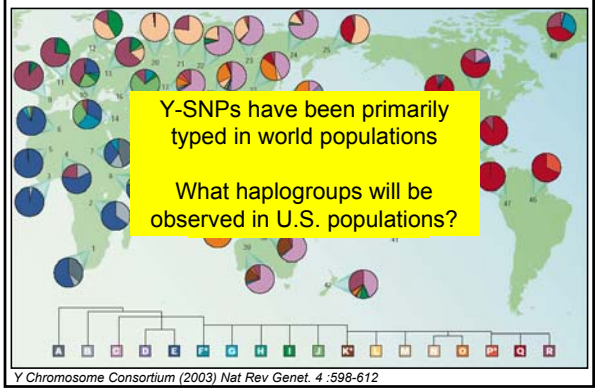
Tree contains over 250 Y-SNPs

Samples were typed for 48 world populations

18 main groups A-R

159 haplogroups defined

Global Distribution of Y Haplogroups



Y Chromosome Consortium (2003) Nat Rev Genet. 4 :598-612

Y-SNPs in U.S. populations

What haplogroups will be observed?

How specific will certain Y-SNPs be for a U.S. population group?

Forensic utility in comparison/addition to Y-STRs

Commercial kit (Marligen) 42 Y-SNPs

Medium sized multiplexes developed in-house (CE or MS)

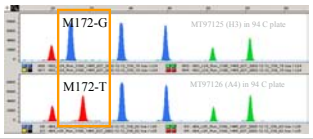
Approaches to Y SNP Typing

Multi-Color Capillary Electrophoresis (ABI 310 or 3100)

Luminex 100 Flow Cytometer

Primer extension (SNaPshot assays)

Allele-specific hybridization (Marligen Signet Y SNP kit)



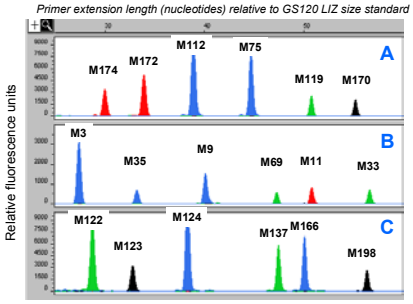
M172-G	M172-T
267	73.5
114	238.5

Allele fluorescent counts on Luminex system

Y SNP Assays Using Primer Extension (SNaPshot)

18 loci in 3 multiplex assays

Primer extension length (nucleotides) relative to GS120 LIZ size standard

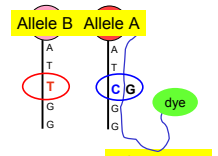


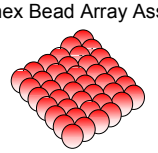
Equal multiplexing done at both PCR and SNP levels (6plexes)

Vallone, P.M. and Butler, J.M. (2004) *J. Forensic Sci., 49(4): 723-732*

SNP Detection by Hybridization

Luminex Bead Array Assay

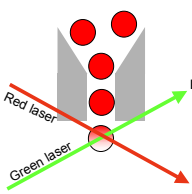




100 different colored beads are possible (potential for multiplexing 50 SNP markers)

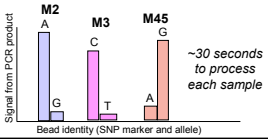
Luminex 100 Flow Cytometer

MARLIGEN BIOSCIENCES **Signet™ Y-SNP Typing System** (42 Y-SNPs + AMEL)

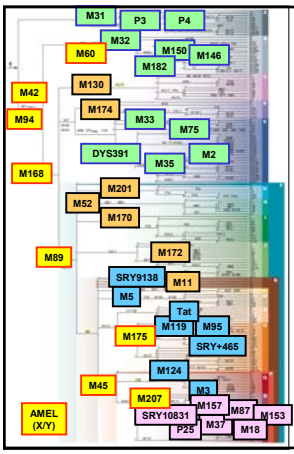


Detects labeled PCR product

Identity of bead (probe)



~30 seconds to process each sample



The Y Chromosome Consortium Map (2003) Nat Rev Genet. 4 :598-612

8+1 **Position of Marligen *Multiplex 1* Y SNPs**

12 **Position of Marligen *Multiplex 2* Y SNPs**

7 **Position of Marligen *Multiplex 3* Y SNPs**

8 **Position of Marligen *Multiplex 4* Y SNPs**

7 **Position of Marligen *Multiplex 5* Y SNPs**

46(159) haplogroups defined

Y-SNPs Typed at NIST

42 SNPs + Amelogenin present in 5 multiplexes
(commercially available kit from Marligen)

18 SNPs in 3 NIST-designed 6plexes (8 unique)
10 SNPs in 2 NIST-designed 5plexes (1 unique)

19 of the SNP sites overlapped…

Resulting in a total of 51 Y-SNPs

115 African Americans
114 Caucasians
95 Hispanics (presently typed for 10 Y-SNPs)

Potential Use for Y SNPs…

Good ethnic separation (African American)

Good ethnic separation (Caucasian)

SWGDAM	M207	M45	M89	DYS391	M2	M170	M172	M201	M153	SRY10831	Hg	Frequency
Samples	A/G	G/A	C/T	C/G	A/G	A/C	T/G	G/T	T/A	A/G		
AA1	A	G	C	G	G	A	T	G	T	G	E3a	40%
AA2	A	G	C	G	G	A	T	G	T	G		
AA3	A	G	C	G	G	A	T	G	T	G		
AA4	A	G	C	G	G	A	T	G	T	G		
AA6	A	G	C	G	G	A	T	G	T	G		
AA7	A	G	C	G	G	A	T	G	T	G		
AA8	A	G	C	G	G	A	T	G	T	G		
AA10	A	G	C	G	G	A	T	G	T	G		
AA11	A	G	C	G	G	A	T	G	T	G		
AA12	A	G	C	G	G	A	T	G	T	G		
AA15	A	G	C	G	G	A	T	G	T	G		
AA16	A	G	C	G	G	A	T	G	T	G		
AA18	A	G	C	G	G	A	T	G	T	G		
AA19	A	G	C	G	G	A	T	G	T	G		
AA20	A	G	C	G	G	A	T	G	T	G		
AA5	A	G	C	G	G	A	T	G	T	G		
C9	A	G	C	G	A	A	T	G	T	G	E3*	3%
C6	A	G	T	C	A	A	G	G	T	G	J2	3%
C7	A	G	T	C	A	A	T	T	T	G	G	3%
AA9	A	G	T	C	A	C	T	G	T	G	I	10%
AA14	A	G	T	C	A	C	T	G	T	G		
C3	A	G	T	C	A	C	T	G	T	G		
C18	A	G	T	C	A	C	T	G	T	G		
AA13	G	A	T	C	A	A	T	G	T	G	R	38%
AA17	G	A	T	C	A	A	T	G	T	G		
C1	G	A	T	C	A	A	T	G	T	G		
C2	G	A	T	C	A	A	T	G	T	G		
C4	G	A	T	C	A	A	T	G	T	G		
C5	G	A	T	C	A	A	T	G	T	G		
C8	G	A	T	C	A	A	T	G	T	G		
C10	G	A	T	C	A	A	T	G	T	G		
C11	G	A	T	C	A	A	T	G	T	G		
C13	G	A	T	C	A	A	T	G	T	G		
C14	G	A	T	C	A	A	T	G	T	G		
C16	G	A	T	C	A	A	T	G	T	G		
C17	G	A	T	C	A	A	T	G	T	G		
C19	G	A	T	C	A	A	T	G	T	G		
C20	G	A	T	C	A	A	T	G	T	G		
C12	G	A	T	C	A	A	T	G	T	A	R1a	3%
C15	G	A	T	C	A	A	T	G	A	G	R1b6	3%

Publication on U.S. Groups with Y-SNPs

J. Forensic Sci. 2004; 49(4): 723-732

J Forensic Sci, July 2004, Vol. 49, No. 4
Paper ID JFS2003303
Available online at: www.astm.org

Peter M. Vallone,[1] Ph.D. and John M. Butler,[1] Ph.D.

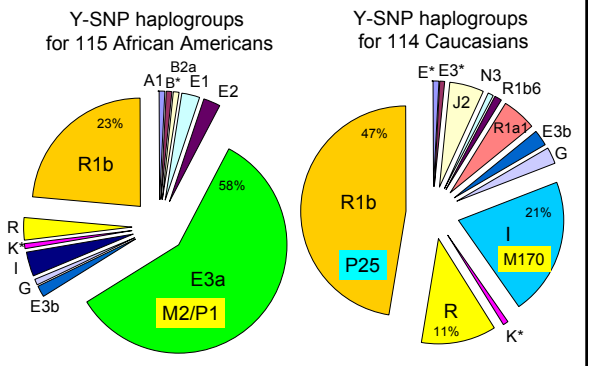
Y-SNP Typing of U.S. African American and Caucasian Samples Using Allele-Specific Hybridization and Primer Extension*

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

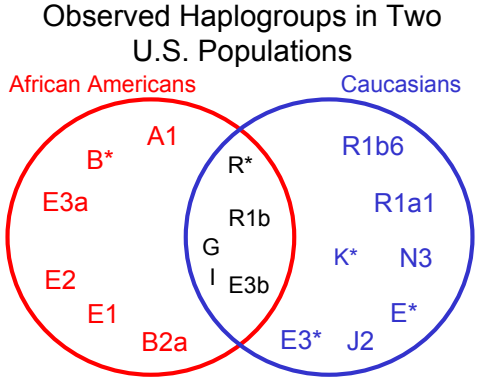
Current Y-SNPs appear to have limited value for ethnic differentiation in U.S. populations (with the exception of M2 that is only found in African Americans and not in Caucasians)

24 plates x 96 samples x 6plexes = >13,000 Y SNP allele calls



18 different haplogroups observed in 229 males

Vallone, P.M. and Butler, J.M. (2004) *J. Forensic Sci., 49(4): 723-732*



18 total Hgs; 5 shared

Vallone, P.M. and Butler, J.M. (2004) Y SNP typing…. *Progress in Forensic Genetics 10., pp. 85-87*

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., 49(4): 723-732*

Summary of NIST Y-Chromosome Work

- Standardize information resources on Y-STRs and nomenclature for alleles
- Understand variation in U.S. populations so the best loci can be selected for commercial kits
- Construct multiplex assays to quickly evaluate loci
- Provide reference material for laboratory calibration (SRM 2395)

Y-Chromosome Publications from NIST (1)

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

- Butler, J.M., Schoske, R., Vallone, P.M., Kline, M.C., Redd, A.J., Hammer, M.F. (2002) A novel multiplex for simultaneous amplification of 20 Y chromosome STR markers. *Forensic Sci. Int.* 129: 10-24.
- Redd, A.J., Agellon, A.B., Kearney, V.A., Karafet, T., de Knijff, P., Park, H., Butler, J.M., Hammer, M.F. (2002) Forensic value of fourteen novel STRs on the human Y chromosome. *Forensic Sci. Int.* 130: 97-111.
- Butler, J.M. (2003) Recent developments in Y-short tandem repeat and Y-single nucleotide polymorphism analysis. *Forensic Sci. Rev.* 15:91-111.
- Schoske, R., Vallone, P.M., Ruitberg, C.M., Butler, J.M. (2003) Multiplex PCR design strategy used for the simultaneous amplification of 10 Y chromosome short tandem repeat (STR) loci. *Anal. Bioanal. Chem.*, 375: 333-343.
- Butler, J.M., Schoske, R., Vallone, P.M. Highly multiplexed assays for measuring polymorphisms on the Y-chromosome. (2003) *Progress in Forensic Genetics 9* (Brinkmann, B. and Carracedo, A., eds.), Elsevier Science: Amsterdam, The Netherlands, International Congress Series 1239, pp. 301-305.

Y-Chromosome Publications from NIST (2)

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

- Schoske, R. (2003) The design, optimization and testing of Y chromosome short tandem repeat megaplexes. PhD dissertation, American University, 270 pp.
- Schoske, R., Vallone, P.M., Kline, M.C., Redman, J.W., Butler, J.M. (2004) High-throughput Y-STR typing of U.S. populations with 27 regions of the Y chromosome using two multiplex PCR assays, *Forensic Sci. Int.* 139: 107-121.
- Vallone, P.M. and Butler, J.M. (2004) Multiplexed assays for evaluation of Y-SNP markers in U.S. populations. *Progress in Forensic Genetics 10*, Elsevier Science: Amsterdam, The Netherlands, International Congress Series 1261, 85-87.
- Butler, J.M. and Schoske, R. (2004) Forensic value of the multi-copy Y-STR marker DYS464. *Progress in Forensic Genetics 10*, Elsevier Science: Amsterdam, The Netherlands, International Congress Series 1261, 278-280.
- Butler, J.M. and Schoske, R. (2004) Duplication of DYS19 flanking regions in other parts of the Y chromosome. *Int. J. Legal Med.,* 118: 178-183.

Y-Chromosome Publications from NIST (3)

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

- Vallone, P.M. and Butler, J.M. (2004) Y-SNP typing of U.S. African American and Caucasian samples using allele-specific hybridization and primer extension. *J. Forensic Sci.* 49(4): 723-732.
- Butler, J.M. (2005) Constructing STR multiplex assays. *Methods in Molecular Biology: Forensic DNA Typing Protocols* (Carracedo, A., ed.), Humana Press: Totowa, New Jersey, 297: 53-66.
- Vallone, P.M., Fahr, K., Kostrzewa, M. (2005) Genotyping SNPs using a UV-photocleavable oligonucleotide in MALDI-TOF MS. *Methods in Molecular Biology: Forensic DNA Typing Protocols* (Carracedo, A., ed.), Humana Press: Totowa, New Jersey, 297: 169-178.
- Butler, J.M., Decker, A.E., Kline, M.C., Vallone, P.M. (2005) Chromosomal duplications along the Y-chromosome and their potential impact on Y-STR interpretation, *J. Forensic Sci., in press.*

Y-Chromosome Publications from NIST (4)

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

- Butler, J.M. and Schoske, R. (2005) U.S. population data for the multi-copy Y-STR locus DYS464. *J. Forensic Sci., in press.*
- Butler, J.M., Appleby, J.E., Duewer, D.L. (2005) Locus-specific brackets for reliable typing of Y-chromosome short tandem repeat markers, *Electrophoresis, in press.*
- Butler, J.M., Decker, A.E., Vallone, P.M., Kline, M.C. (2005) Allele frequencies for 27 Y-STR Loci with U.S. Caucasian, African American, and Hispanic Samples, *Forensic Sci. Int., in press.*
- Gusmão L, Butler JM, Carracedo A, Gill P, Kayser M, Mayr WR, Morling N, Prinz M, Roewer L, Schneider PM, Tyler-Smith C (2005) DNA Commission of the International Society of Forensic Genetics (ISFG): An update of the recommendations on the use of Y-STRs in forensic analysis. *Forensic Sci. Int., in press*

International Forensic Y-User Workshops

- Next meeting (5th): Sept 26-30, 2006 (Innsbruck, Austria) – will also cover mtDNA
- 1st – Berlin, Germany June 1996
- 2nd – Berlin, Germany June 2000
- 3rd – Porto, Portugal Nov 2002
- 4th – Berlin, Germany Nov 2004

For more information, see: http://www.yhrd.org/index.html