Y-Chromosome and Mitochondrial DNA Analysis

The Human Y-Chromosome:
Background and Y-SNPs

NEAFS 2006 Workshop
Rye Brook, NY
November 1, 2006

Northeastern Association of Forensic Scientists

Dr. John M. Butler
Dr. Michael D. Coble

john.butler@nist.gov
Michael.Coble@afip.osd.mil

Presentation Outline

- Characteristics of the Y-Chromosome and Y-SNPs
- Y-STR Markers, Core Loci, and Kits
- Populations, Mutations, and Statistics
- Work with Additional Loci to Separate Common Types
- Casework Examples and Resources

"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 few years…

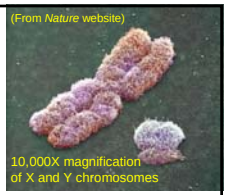
- "Full" Y-chromosome sequence became available in June 2003; over 350 Y-STR loci identified (only ~20 in 2000)
- **Selection of core Y-STR loci** (SWGDAM Jan 2003)
- **Commercial Y-STR kits released**
  - ~~Y PLEX 6,5,12 (2001-03),~~ **PowerPlex Y** (9/03), **Yfiler** (12/04)
- Many population studies performed and databases generated with thousands of Y-STR haplotypes
- Forensic casework demonstration of value of Y-STR testing along with court acceptance

Characteristics of the Y-Chromosome

THE HUMAN Y CHROMOSOME: AN EVOLUTIONARY MARKER COMES OF AGE

Mark A. Jobling & Chris Tyler-Smith

Nature Reviews Genetics (2003) 4, 598-612

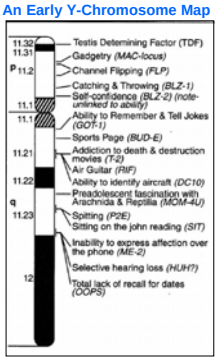


Abstract

- **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

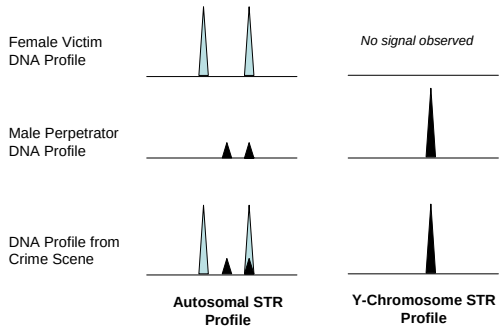
Disadvantages of the Y-Chromosome

- 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
  - **More like addition (10 + 10 + 10 = 30) than multiplication (10 x 10 x 10 = 1,000)**

Forensic Advantages of Y-STRs

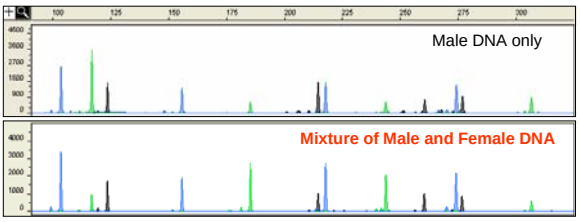
- **Male-specific amplification** extends range of cases accessible to obtaining probative DNA results (e.g., fingernail scrapings, sexual assault without sperm)
- **Technical simplicity due to single allele profile**; can potentially recover results with lower levels of male perpetrator DNA because 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

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



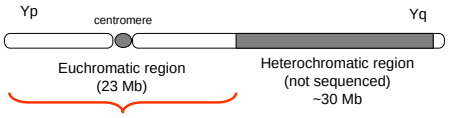
Butler, J.M. (2005) *Forensic DNA Typing, 2nd Edition*, Figure 9.2, ©Elsevier Science/Academic Press

Y-STRs Identify the Male Component even with Excess Female DNA



800X female DNA

"Full" Y-Chromosome Sequence



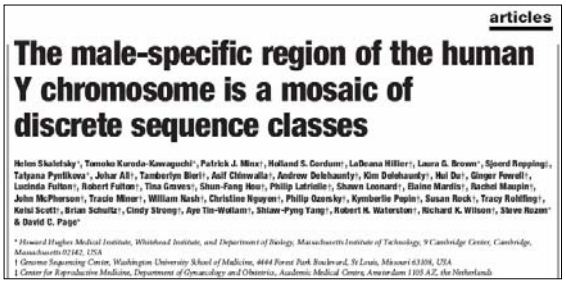
Sequence reported in June 2003 as part of Human Genome Project

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

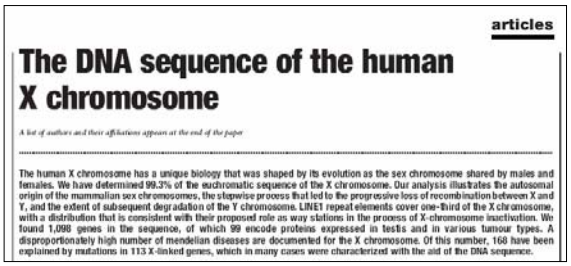
The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes

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)



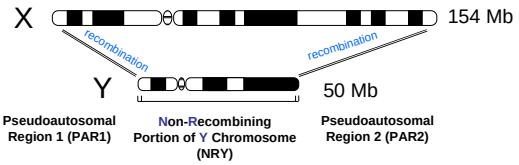
articles

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)

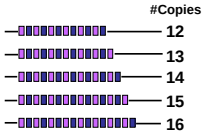
Focus in Forensics is on the NRY
PCR primers for Y-STRs need to be carefully designed to avoid X-chromosome homology

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

Various Types of Genetic Markers on the Human Y-Chromosome

Y-STRs
Short Tandem Repeats

–GATAGATAGATAGATA–



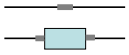
Multi-state characters
Quickly evolving (**2 x 10^{-3}**/gen)
High resolution haplotypes

Y-SNPs
Single Nucleotide Polymorphisms

–CG**A**TG–

–CG**G**TG–

Insertion/deletions (indels)

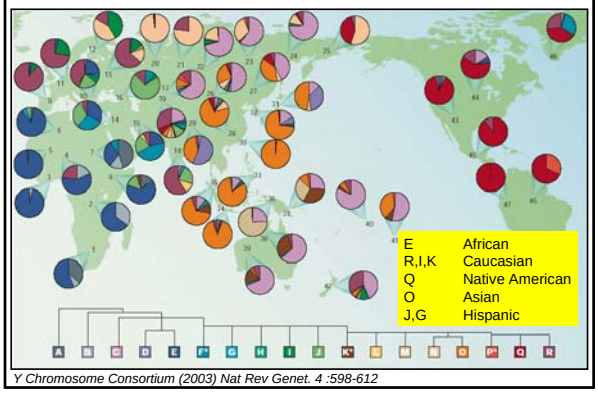


Binary characters
Slowly evolving (**~10^{-8}**/gen)
Low resolution haplogroups

Slide from Alan Redd (University of Arizona) presentation at Promega Oct 2002

Y-SNPs

Global Distribution of Y Haplogroups



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

Haplogroups Defined by Various Y-SNPs

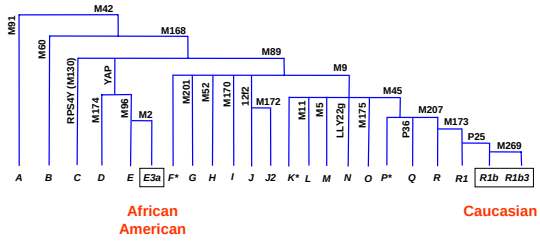


Figure 9.9, J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press

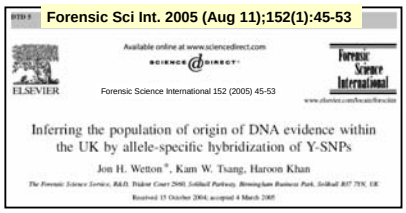
Potential Use for Y SNPs…

Good ethnic differentiation (African American)

Good ethnic differentiation (Caucasian)

SWGDAM Samples	M207 A/G	M45 G/A	M89 C/T	DYS391 C/G	M2 A/G	M170 A/C	M172 T/G	M201 G/T	M153 T/A	SRY10831 A/G	Hg	Frequency
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%

Recent Forensic Science Service Work with Y-SNPs



Forensic Sci Int. 2005 (Aug 11);152(1):45-53

Available online at www.sciencedirect.com
SCIENCE DIRECT

ELSEVIER

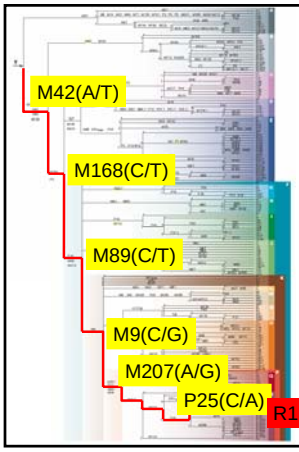
Forensic Science International 152 (2005) 45-53

Forensic Science International

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

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)

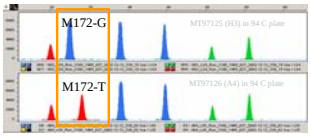
Approaches to Y SNP Typing

Multi-Color Capillary Electrophoresis
(ABI 310 or 3100)

Luminex 100 Flow Cytometer

Primer extension
(SNaPshot assays)

Allele-specific hybridization
(Marligen Signet Y SNP kit)



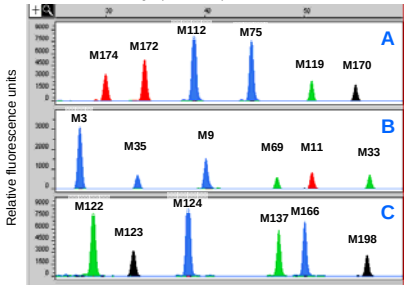
	M172-G	M172-T
	267	73.5
	114	238.5

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

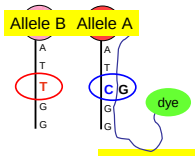


Equal multiplexing done at both PCR and SNP levels (6plexes)

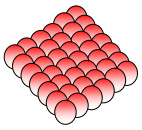
Vallone, P.M. and Butler, J.M. (2004) *J. Forensic Sci., 49(4): 723-732*

SNP Detection by Hybridization

Luminex Bead Array Assay



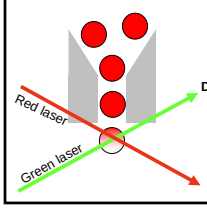
PCR product



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

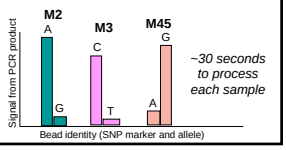
Luminex 100 Flow Cytometer

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

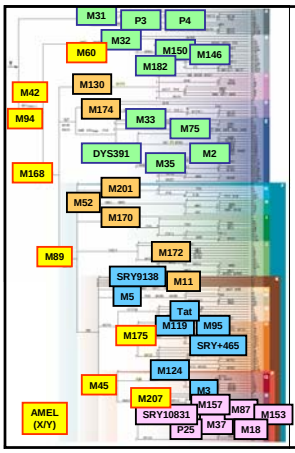


Detects labeled
PCR product

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)

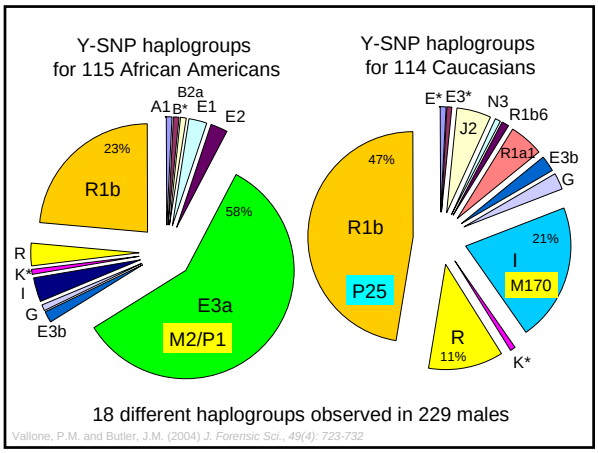
18 SNPs in 3 NIST-designed 6plexes (8 unique)
10 SNPs in 2 NIST-designed 5plexes (1 unique)

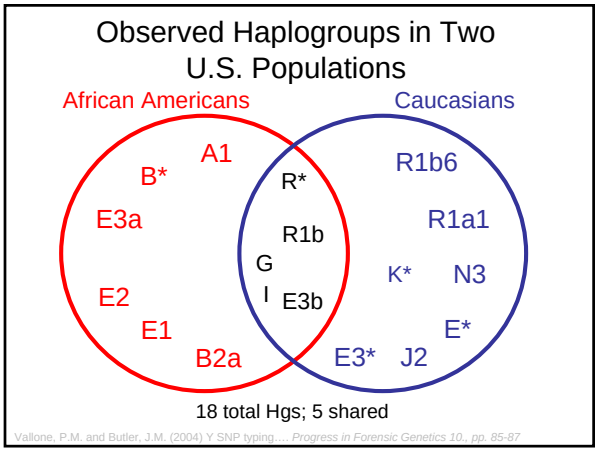
19 of the SNP sites overlapped…

Resulting in a total of 51 Y-SNPs

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

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





Publication on U.S. Groups with Y-SNPs

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*

Summary

- 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

Y-SNP Information Cataloged on STRBase Website

http://www.cstl.nist.gov/biotech/strbase/SNPs/YSNPs50.htm

Y position (Mb)	SNP Name	YCC Hg Defined	Multiplex	Polymorphism	African American (N=115)	Caucasian (N=114)
2,562,931	SRY+465	O2b	4	C->T	1.00/0.00	1.00/0.00
2,564,927	$SRY_{10831a,b}$	B-R, R1a	5	A->G, G->A	0.01/0.99, 1.00/0.00	0.00/1.00, 0.95/0.05
2,566,620	SRY_{9138}	K1	4	C->T	1.00/0.00	1.00/0.00
2,642,605	M130 (RPS4Y)	C	3	C->T	1.00/0.00	1.00/0.00
13,407,330	M2	E3a	2	A->G	0.42/0.58	1.00/0.00
13,413,670	DYS391	E3	2	C->G	0.40/0.60	0.96/0.04
14,124,138	M168	C-R	1	C->T	0.03/0.97	0.00/1.00
14,157,939	M170	I	3, A	A->C	0.97/0.03	0.79/0.21
14,179,223	M182	B2	2	C->T	0.99/0.01	1.00/0.00
14,232,730	Tat	N3	4	T->C	1.00/0.00	0.99/0.01
14,264,427	M174	D	3, A	T->C	1.00/0.00	1.00/0.00

Positions mapped against Human Genome reference sequence (July 2003) using BLAT

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

Recent Y-SNP Article from Mike Hammer's Group

ARTICLE IN PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT

ELSEVIER

Forensic Science International

Forensic Science International xxx (2005) xxx–xxx

www.elsevier.com/locate/forsciint

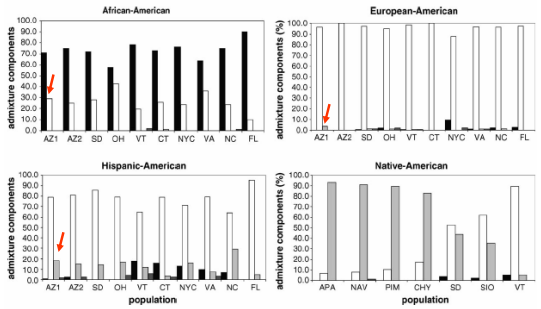
Population structure of Y chromosome SNP haplogroups in the United States and forensic implications for constructing Y chromosome STR databases

Michael F. Hammer [a,b,*], Veronica F. Chamberlain [a], Veronica F. Kearney [a], Daryn Stover [a], Gina Zhang [a], Tatiana Karafet [a], Bruce Walsh [b], Alan J. Redd [a,1]

[a] Division of Biotechnology, University of Arizona, Tucson, AZ 85721, USA
[b] Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721, USA

Received 17 August 2005; received in revised form 8 November 2005; accepted 8 November 2005

Some Admixture Exists Between U.S. Population Groups Based on Y-SNP Typing Results -- Hammer *et al.* (2006) *FSI in press*



Continental origin rather than current location in the U.S. determines major patterns of Y chromosome variation for most ethnic groups.