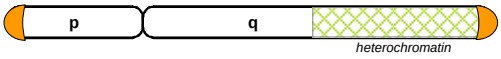


The Human Y-Chromosome:
Background and Use in Forensic DNA Typing

John M. Butler
National Institute of Standards and Technology

CODIS Conference – October 23, 2006
Arlington, VA

Summary of 2006 CODIS Survey Questions Regarding Y-STRs

171 labs

Questions #45a & #45b

- **Is your lab using or validating Y-STRs?**
  - **51 Yes** (30%)
    28 Yfiler, **15** PowerPlex Y, some both kits
  - **114** No
  - **6** no response

Summary of 2006 CODIS Survey Questions Regarding Y-STRs

171 labs

Question #50

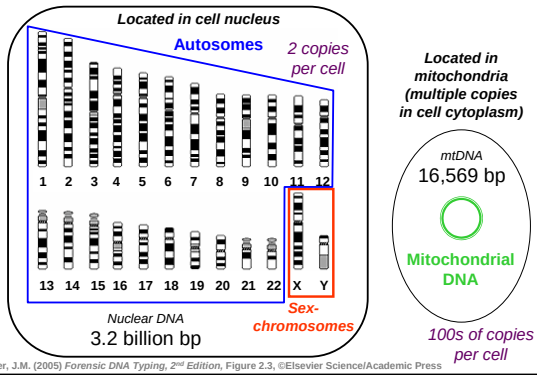
- Y-STR data can be entered in CODIS similar to entering the current STR loci in CODIS. **Do you think CODIS should include Y-STR loci in Popstats calculations?**
  - Yes – 116 (68%)
  - No – 18
  - No response – 37

Presentation Outline

- Background on human Y-chromosome
- Why Y is of interest in human identity testing
- Y-STR markers and kits available
- Different population databases and statistics for reporting matches

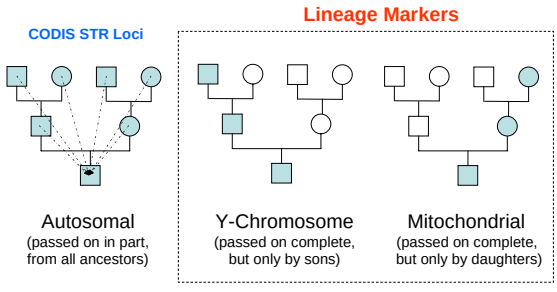
Human Genome

23 Pairs of Chromosomes + mtDNA



Butler, J.M. (2005) *Forensic DNA Typing, 2nd Edition,* Figure 2.3, ©Elsevier Science/Academic Press

Different Inheritance Patterns

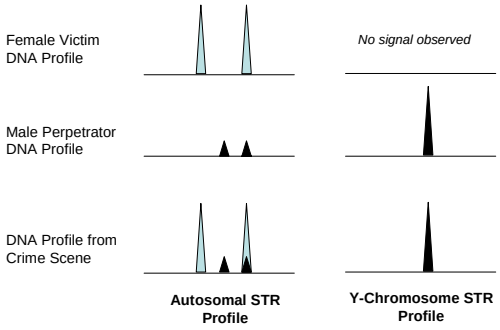


Butler, J.M. (2005) *Forensic DNA Typing, 2nd Edition,* Figure 9.1, ©Elsevier Science/Academic Press

Role of Y-STRs and mtDNA
Compared to Autosomal STRs

- **Autosomal STRs provide a higher power of discrimination and <u>are the preferred method</u> whenever possible**
- **Due to capabilities for male-specific amplification**, Y-chromosome STRs (**Y-STRs**) can be useful in extreme female-male mixtures (e.g., when differential extraction is not possible such as fingernail scrapings)
- **Due to high copy number**, mitochondrial DNA (**mtDNA**) may be the only source of surviving DNA in highly degraded specimens or low quantity samples such as hair shafts

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

Value of Y-Chromosome Markers

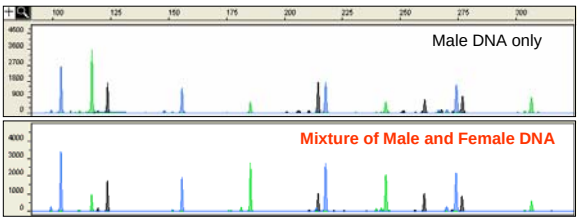
J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition; Table 9.1

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

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 Identify the Male Component even with Excess Female DNA



800X female DNA

Scenarios Where Y-STRs Can Aid Forensic Casework

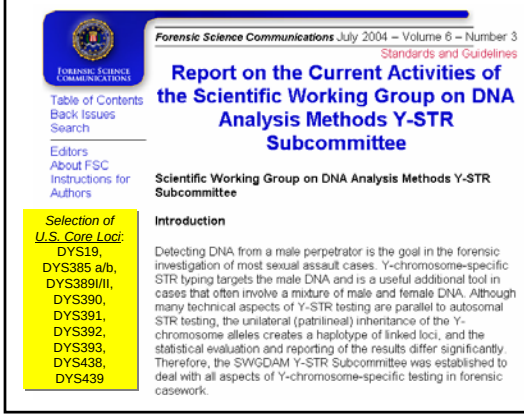
- Sexual assaults by vasectomized or azoospermic males (no sperm left behind for differential extraction)
- Extending length of time after assault for recovery of perpetrator's DNA profile (greater than 48 hours)
- Fingernail scrapings from sexual assault victims
- Male-male mixtures
- Other bodily fluid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

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)**

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



Forensic Science Communications July 2004 – Volume 6 – Number 3

Standards and Guidelines

Report on the Current Activities of the Scientific Working Group on DNA Analysis Methods Y-STR Subcommittee

Table of Contents
Back Issues
Search

Editors
About FSC
Instructions for Authors

Selection of U.S. Core Loci: DYS19, DYS385 a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS438, DYS439

Scientific Working Group on DNA Analysis Methods Y-STR Subcommittee

Introduction

Detecting DNA from a male perpetrator is the goal in the forensic investigation of most sexual assault cases. Y-chromosome-specific STR typing targets the male DNA and is a useful additional tool in cases that often involve a mixture of male and female DNA. Although many technical aspects of Y-STR testing are parallel to autosomal STR testing, the unilateral (patrilineal) inheritance of the Y-chromosome alleles creates a haplotype of linked loci, and the statistical evaluation and reporting of the results differ significantly. Therefore, the SWGDAM Y-STR Subcommittee was established to deal with all aspects of Y-chromosome-specific testing in forensic casework.

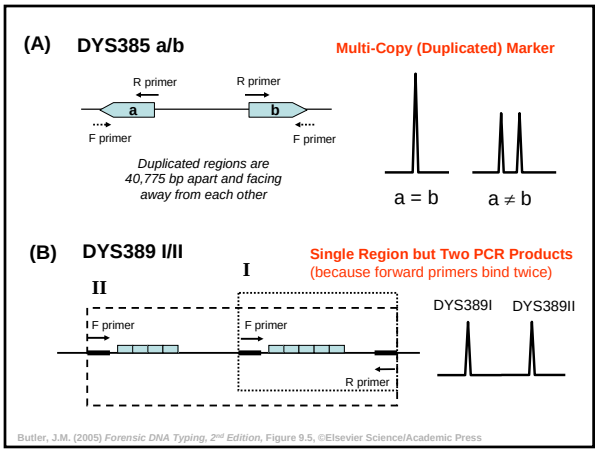
Core Y-STR Characteristics

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).

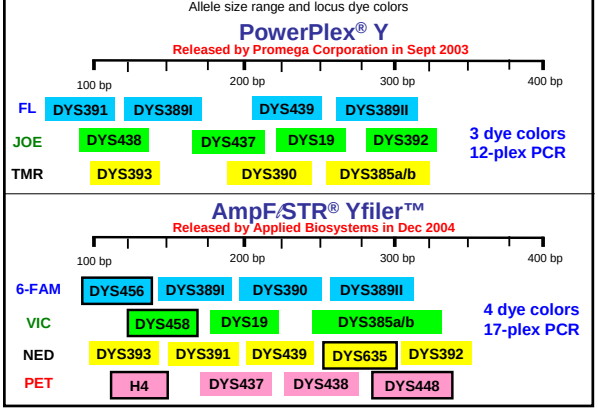
Butler, J.M. (2006) Genetics and genomics of core STR loci used in human identity testing. *J. Forensic Sci.*, 51(2): 253-265



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

Primary Commercial Y-STR Kits

Allele size range and locus dye colors



Haplotype Databases for Y-STR Kits

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

PowerPlex Y

1311 Caucasians
325 Asians
894 Hispanics
1108 African Americans
366 Native Americans

4,004 total
(as of March 2005)

Yfiler

1276 Caucasians
330 Asians
597 Hispanics
985 African Americans
106 Native Americans
105 Filipino
59 Sub-Saharan Africans
103 Vietnamese

3,561 total
(as of December 2004)

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

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

National U.S. Y-STR Population Database

- Efforts underway at the **University of Central Florida** (with NIJ funding) to consolidate all U.S. data on Y-STR loci for population
- Data from ReliaGene, Promega, Applied Biosystems being gathered plus any forensic lab population sample data available

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 profile 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.

Acknowledgments

Funding from interagency agreement 2003-IJ-R-029 between the National Institute of Justice and the NIST Office of Law Enforcement Standards

NIST Human Identity Project Team – Leading the Way in Forensic DNA…

John Butler | Margaret Kline | Pete Vallone | Jan Redman | Amy Decker | Becky Hill | Dave Duewer

Tom Reid (DNA Diagnostics Center) – supplying the father-son samples for mutation rate analysis