Y STRs

Bruce McCord
Professor Analytical/Forensic Chemistry
FIU Department of Chemistry
International Forensic Research Institute
Miami, FL 33199
mccordb@fiu.edu 305-348-7543
What if there is a mixture with a large amount of female DNA and a small amount of male DNA?

• Differential extractions may not work if
  – Large differences in DNA quantity between male and female exist in rape cases
  – Vasectimized males
  – Saliva samples
  – Such samples cry out for a male specific marker
Human Genome
23 Pairs of Chromosomes + mtDNA

Autosomes
2 copies per cell

Y chromosome
1 copy per cell

mtDNA
16,569 bp

Different Inheritance Patterns

CODIS STR Loci

Autosomal
(passed on in part, from all ancestors)

Y-Chromosome
(passed on complete, but only by sons)

Mitochondrial
(passed on complete, but only by daughters)

Lineage Markers

Sequential detection of mtDNA and multicopy nuclear DNA (Alu)
Role of Y-STRs (and mtDNA) Compared to Autosomal STRs

• Autosomal STRs provide a higher power of discrimination and are the preferred method whenever possible

• Due to capabilities for male-specific amplification, Y-chromosome STRs (Y-STRs) can be useful in extreme female-male mixtures (bite marks, touch samples, etc.)

• Y STR typing can be very sensitive as the kit is designed to detect mixtures of DNA
Y Chromosome DNA Typing:
Effect of concentration on the ability to recover male DNA in Powerplex
The recovery of a male profile (upper egram) when no mixture is evident with identifiler. (lower egram) Duo showed a 1/1000 male/female mixture.

Figure 5. Yfiler® kit profile generated from 1:1000 male:female DNA mixture.

Figure 6. Identifiler® kit profile generated from 1:20 male:female DNA mixture. Only a single source female profile was obtained.

Y Chromosome STRs

Advantages:
1. Specific to males (inherited solely through the male-line descendants. Non recombining.
2. Differential extraction in rape kits is unnecessary.
3. Men commit the vast majority of violent crimes.
4. Because they don’t recombine, Y chromosomes retain a unique genological record of mutations in their junk DNA

1. Like mtDNA, statistics are based on database size.
2. Mixed profiles can be difficult to interpret statistically
3. There are odd issues with the Y chromosome: duplications, deletions are possible – ie
4. When sample limited choices must be made- autosomal, Y, mtDNA?
When will Y testing be used

1. When regular STR testing fails indicates little male DNA is present
2. When male saliva is collected from a female body
3. In a mixed blood stain – male cuts himself in a stabbing of a female
4. Fingernail clippings, touch samples
5. Ligatures around strangulation victim.

Sample collection will be important in such cases and QPCR will assist greatly in determining value.
Important Questions Can Be Answered with Y Chromosome Tests...

I need a good test to be sure if Darth Vader is really my father...

Luke,
We have the same Y-STR haplotype...

But how can I be sure if I am using the right quantity of DNA in a mixture?

Apologies to George Lukas, John Butler and other B movie heroes
Y- Chromosomes and Thomas Jefferson

(1) Y Chromosomes STRs were used to trace the lineage of Thomas Jefferson - this question arose at the time Clinton has having his troubles

(a) The problem basically was this: Did Thomas Jefferson father an illegitimate son?

(i) The potential mother was Sally Hemmings, whom Jefferson used as a personal secretary and who was a slave in his household.

1. At the time there were rumors floating around that Jefferson had fathered illegitimate children
2. One of them was thought to be Easton Hemmings, who was a son of Sally Hemmings.
locade the male line descendents of both Jefferson and Hemmings and track polymorphic regions in their Y chromosomes.

1. Apart from mutations there would be no possibility of change in these markers
2. This is unlike all X chromosomes which can result from either sex

(i) The only problem was that Jefferson had no male descendant- but he did have a paternal uncle - Field Jefferson
1. Male lines were located for two of Hemmings sons- Easton, and Thomas Woodson and Samuel and Peter Carr, Jefferson's nephew

Could DNA typing solve this problem?

Perhaps
Y Chromosome analysis of Field Jefferson’s descendants (Thomas Jefferson had no surviving sons) *(Nature, Nov. 5, 1998)*

<table>
<thead>
<tr>
<th>Family</th>
<th>Pedigree</th>
<th>Biliallic Markers</th>
<th>Microsatellite STR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Field</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J41</td>
<td></td>
<td>15,12,4,11,3,9,11,10,15,13,7</td>
<td></td>
</tr>
<tr>
<td>J42</td>
<td></td>
<td>15,12,4,11,3,9,11,10115,13,7</td>
<td></td>
</tr>
<tr>
<td>J47</td>
<td></td>
<td>15,12,4,11,3,9,11,10,15,13,7</td>
<td></td>
</tr>
<tr>
<td>J49</td>
<td></td>
<td>15,12,4,11,3,9,11,10,15,13,7</td>
<td></td>
</tr>
<tr>
<td>J50</td>
<td></td>
<td>15,12,4,11,3,9,11,10,16,13,7</td>
<td></td>
</tr>
<tr>
<td><strong>Jefferson</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J47</td>
<td></td>
<td>15,12,4,11,3,9,11,10,15,13,7</td>
<td></td>
</tr>
<tr>
<td>J49</td>
<td></td>
<td>15,12,4,11,3,9,11,10,15,13,7</td>
<td></td>
</tr>
<tr>
<td>J50</td>
<td></td>
<td>15,12,4,11,3,9,11,10,16,13,7</td>
<td></td>
</tr>
<tr>
<td><strong>Easton</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H21</td>
<td></td>
<td>15,12,4,11,3,9,11,10,15,13,7</td>
<td></td>
</tr>
<tr>
<td><strong>Sally’s son</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C27</td>
<td></td>
<td>14,12,5,12,3,10,11,13,13,13,7</td>
<td></td>
</tr>
<tr>
<td>C29</td>
<td></td>
<td>14,12,5,11,3,10,11,13,13,13,7</td>
<td></td>
</tr>
<tr>
<td>C31</td>
<td></td>
<td>14,12,5,12,3,10,11,13,13,13,7</td>
<td></td>
</tr>
<tr>
<td><strong>Jefferson’s nephews</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C27</td>
<td></td>
<td>14,12,5,12,3,10,11,13,13,13,7</td>
<td></td>
</tr>
<tr>
<td>C29</td>
<td></td>
<td>14,12,5,11,3,10,11,13,13,13,7</td>
<td></td>
</tr>
<tr>
<td><strong>Peter Carr</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C27</td>
<td></td>
<td>14,12,5,12,3,10,11,13,13,13,7</td>
<td></td>
</tr>
<tr>
<td>C29</td>
<td></td>
<td>14,12,5,11,3,10,11,13,13,13,7</td>
<td></td>
</tr>
<tr>
<td><strong>Thomas Hemmings</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W55</td>
<td></td>
<td>14,12,5,11,3,10,11,13,13,13,7</td>
<td></td>
</tr>
<tr>
<td>W56</td>
<td></td>
<td>14,12,5,11,3,10,11,13,13,13,7</td>
<td></td>
</tr>
<tr>
<td><strong>Sally’s Son</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W69</td>
<td></td>
<td>14,12,5,11,3,10,11,13,13,13,7</td>
<td></td>
</tr>
<tr>
<td>W70</td>
<td></td>
<td>17,12,6,11,3,11,8,10,11,14,6</td>
<td></td>
</tr>
</tbody>
</table>
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... 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

The more modern version


Sex determining region Y

Amelogenin
Issues surrounding the Y chromosome

• ca 50 Mbases long (third smallest after 21 and 22
• Contains 90 -300 genes
• 23 Mbases have been sequenced, the rest consists of highly repetitive DNA, difficult to sequence (Heterochromic region)
• 95% does not combine with the X chromosome, 2 regions at the tips however do, PAR1 and PAR2 (Pseudo autosomal region)
• The middle portion is known as the male specific region
  – It used to be called the non recombining region, but in fact it recombines with itself, and much sequence homology exists via back and forth gene conversion
• A variety of genetic diseases exist in which an X chromosome is damaged. This creates few problems to the mother but can be deadly to a male son. (hemophilia is an example)
Another issue with the Y chromosome is that there is recombination with the X chromosome at its tips (Pseudoautosomal regions 1 & 2).

Pseudoautosomal regions consist of 5% of the Y and code for 9 found genes. 80 found genes exist on the 95% of the rest of the DNA.
Classes of sequences in the Y chromosome

MSY Region - The euchromic region  23MB sequenced of 50 MB

• X transpose
  - 99% identical to Xq21, 2 coding genes on 2 portions of the short arm (3.4Mb)

• X degenerate
  - 96% similarity to X

• The above sequences reflect the ancient common origin of the two chromosomes and provide evidence of a stepwise decay over time

• Ampliconic
  - Sets of very similar sequences, some of which are palindromes. Many relate to male sex genes
  - The similarity of these sequences caused difficulties in the human genome project

• Heterochromic Region – region of tightly wound DNA, not expressed or sequenced

http://www.nature.com/nature/journal/v423/n6942/full/423810a.html
**A)** DYS385 a/b

Duplicated regions are 40,775 bp apart and facing away from each other.

**B)** DYS389 I/II

Single Region but Two PCR Products (because forward primers bind twice)
Y-Chromosome Haplotype Reference Database (YHRD)

US haplotype requires 2 additional loci:
- DYS438
- DYS439

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

Commercial Y-STR kits exist to amplify all of the core loci in a single reaction (plus a few additional markers).
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
- PowerPlex Y (12 loci)
  - 0 matches in 4,004
- Yfiler (17 loci)
  - 0 matches in 3,561
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:

\[ p = \frac{X}{N} \]

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

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

7 matches in 27,773

\[ p = \frac{7}{27,773} = 0.000252 = 0.025\% \]

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

\[ = 0.000252 + 0.000187 = 0.000439 \]

\[ = 0.044\% \text{ (~1 in 2270)} \]
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

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

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

$$1 - \alpha^{1/N} = 1 - (0.05)^{[1/4,004]} = 0.000748$$

$$= 0.075\% \sim (1 \text{ in } 1340)$$

If using database of 2,443, then the best you can do is 1 in 816.
### Core Y-STR Characteristics

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

Promega PowerPlex® Y Allelic Ladders

U.S. Core Loci + DYS437
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?
Problems with Most Common Type (MCT)

Most common type in Europe occurs about 3% of the time. 530 out of 19.4K samples. Other types can be rare.

www.ystr.org
Researchers have found one Y chromosome fingerprint that was identical in eight percent of the male population of Central Asia. "This was highly unusual and suggested that they may all have descended from one man living in the fairly recent past.

By seeing what small changes had occurred, it was possible to estimate the time at which this common ancestor lived, and it was consistent with an origin in the 12th or 13th century," Sykes said. Matching that evidence with the overlap between where the chromosome was abundant and the geographical extent of the Mongol empire established by Genghis Khan in the 12th century, the researchers concluded it was Genghis' chromosome.
One Y chromosome type is present in 8% of the population of Central Asia. Slight mutation events lead back to the 12th Century legacy of Genghis Khan?
Why the localization of this haplotype?

• The Mongol emperor's habit of killing the men and inseminating the women when his army conquered a new territory, coupled with handing the Empire and other wealth to his sons, and their sons, would explain how the chromosome came to such prevalence today, said Sykes.

• The final piece of evidence came from the Hazara, a hill tribe in Pakistan who had a strong oral history of being descended from Genghis Khan.

• "The Y chromosome was present in the Hazara, but not in the surrounding tribes, who did not have this oral history. Though the evidence is circumstantial, it is, I believe, very strong," Sykes said.

• Finding Genghis Khan's tomb, one of the great secrets of all time, could provide the definitive evidence, leading to a direct comparison of Genghis' Y chromosome with those of modern men.
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.
Y STR typing can do amazing things in the presence of overwhelming excesses of female DNA.
1μL of male blood combined with 100-1000 μL female blood

Y STR concentration is important however and it is important to understand that kits can be highly sensitive and occasionally suppressed by the female signal, depending on concentration.
Effect of Concentration: Single source Y STR analysis (ABI Yfiler)

Y Filer is a very sensitive single source assay

Figure 1: comparison of Y-Filer alleles with decreasing amounts of DNA (500pg, 250 pg, 100pg, 50pg and 10pg). Only 4 of the total 17 alleles are detected with 10 pg of DNA.
Issues with mixtures and inhibition

• The important issue with Y STRs and QPCR is the ability to parse out mixtures

• Goal should be to use the autosomal/Y ratio to determine if a sample can be analyzed by conventional STRs or if Y is necessary.

• But: Weird things can happen at low quantities of sample.
Effect of Concentration: Mixed source Y STR analysis (ABI Yfiler)
Relative sensitivity drops in a low level mixture with female DNA

Comparison between a male control sample at 250pg and a mixture of 250 pg male DNA with 1ng of female DNA. In the mixture, the loss of 6 alleles is observed.
<table>
<thead>
<tr>
<th>Locus</th>
<th>1:0</th>
<th>1:100</th>
<th>1:200</th>
<th>1:400</th>
<th>1:600</th>
<th>1:800</th>
</tr>
</thead>
<tbody>
<tr>
<td>DYS392</td>
<td>3337</td>
<td>1406</td>
<td>642</td>
<td>493</td>
<td>467</td>
<td>347</td>
</tr>
<tr>
<td>DYS390</td>
<td>2457</td>
<td>1892</td>
<td>1589</td>
<td>1257</td>
<td>1467</td>
<td>1963</td>
</tr>
<tr>
<td>DYS385 a/b</td>
<td>2839</td>
<td>1547</td>
<td>1212</td>
<td>843</td>
<td>1105</td>
<td>811</td>
</tr>
<tr>
<td>DYS389I</td>
<td>3212</td>
<td>1795</td>
<td>1136</td>
<td>685</td>
<td>841</td>
<td>655</td>
</tr>
<tr>
<td>DYS393</td>
<td>3532</td>
<td>3080</td>
<td>3201</td>
<td>4277</td>
<td>3472</td>
<td>4032</td>
</tr>
<tr>
<td>DYS389II</td>
<td>1152</td>
<td>944</td>
<td>959</td>
<td>786</td>
<td>911</td>
<td>1033</td>
</tr>
<tr>
<td>DYS391</td>
<td>1733</td>
<td>1613</td>
<td>1187</td>
<td>803</td>
<td>1578</td>
<td>1208</td>
</tr>
<tr>
<td>DYS389II</td>
<td>1415</td>
<td>625</td>
<td>422</td>
<td>314</td>
<td>213</td>
<td>229</td>
</tr>
<tr>
<td>X</td>
<td>1427</td>
<td>&gt;5000</td>
<td>&gt;5000</td>
<td>&gt;5000</td>
<td>&gt;5000</td>
<td>&gt;5000</td>
</tr>
<tr>
<td>Amelogenin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td>1465</td>
<td>176</td>
<td>85</td>
<td>75</td>
<td>&lt;75</td>
<td>&lt;75</td>
</tr>
<tr>
<td>DYS19</td>
<td>2221</td>
<td>1257</td>
<td>631</td>
<td>351</td>
<td>505</td>
<td>399</td>
</tr>
<tr>
<td>DYS439</td>
<td>1419</td>
<td>1639</td>
<td>1222</td>
<td>691</td>
<td>960</td>
<td>1075</td>
</tr>
<tr>
<td>DYS438</td>
<td>1936</td>
<td>548</td>
<td>386</td>
<td>195</td>
<td>231</td>
<td>161</td>
</tr>
</tbody>
</table>

*The mixture samples were prepared by taking 0.5 ng of male DNA and increasing the quantity of female DNA to obtain the indicated ratio.
Effect of Inhibitors on YSTRs

Fig. 2. Humic acid inhibition of STR amplifications. PowerPlex® 16 (Panel A) or PowerPlex® Y (Panel B) amplifications of 1 ng of DNA were performed in the presence of the indicated amounts of humic acid.
Conclusions on Y STR applications

• Y STRs are very valuable in mixtures
• Databases are expanding but some issues remain – significance of a match, common haplotypes,
• Low level Y STRs are sensitive to inhibition especially in mixtures
So the Big Question: When to Use Y? And how will QPCR help?

Modern QPCR kits – Quantifiler Duo and Plexor HY will provide a ratio of Autosomal to Male DNA

Laboratories will need to evaluate these kits in combination with their sensitivity thresholds

Issues will be:

1. **What’s the Question?** Is male DNA present? Or Is a mixture present at some ratio?
2. **Type of sample:** differential extraction, bitemark, fingernail, touch /digital penetration sample
3. **System sensitivity for autosomal mixtures** ie: at what M/F ratio does it become impossible to recover/interpret a mixed profile? (note this ratio will change with total input DNA and A/Y ratio.
4. **Availability of validated Y STR typing:** In - house or outsourced? Would you ever not perform an autosomal STR profile?
5. **Precision and sensitivity of estimate.** Single copy assay may produce more precise ratios while multicopy assays will be more sensitive
How to use Y quantification in casework

• Use it as a presumptive test for the presence of male DNA. - find and amplify the most probabtive samples

• Use it for an estimate of amplification success. If Y ratio is above 10% then Autosomal STR analysis may provide a useful result.

• Use it to estimate the amount of DNA template to get a useful Y STR result.
Conclusions

- The Y Chromosome is a complex and interesting piece of DNA
- Y STR typing while not as valuable as autosomal can provide results and statistics based on the counting method
- Y DNA quantification can be used as a presumptive test for evidence screening,
- to determine mixtures and if autosomal DNA will work
- to determine the quantity of male DNA for Y STR analysis
Acknowledgements

• Heather LaSalle
• Silvia Zoppis
• George Duncan
• Eric Buel