New Autosomal and Y-STR Loci and Kits: Making Data Driven Decisions

Organized by John M. Butler
National Institute of Standards and Technology

Schedule

1:00 – 1:20 pm  Welcome and Introductory Remarks  John Butler
1:20 – 1:50 pm  NIST Studies: Kit Concordance and U.S. Population Data  Mike Coble (for Becky Hill)
1:50 – 2:30 pm  Experience with PowerPlex Fusion  Jeff Nye
2:30 – 2:45 pm  BREAK  Hope Olson
2:45 – 3:15 pm  Experience with GlobalFiler  Jason Kokoszka
3:15 – 3:45 pm  NIST Studies with New Y-STR Loci & Kits  Mike Coble
3:45 – 4:00 pm  STRBase Resources and Wrap-up  John Butler

This workshop describes characteristics of additional STR loci that are part of new STR kits created to meet the expanded European and U.S. core loci requirements. Experience will be shared with U.S. population data, STR kit concordance studies, issues with validating and implementing new STR kits, benefits of additional Y-STR loci in recent Y-STR kits, and resources available on the NIST STRBase website.

Points of view are those of the presenters and do not necessarily represent the official position or policies of the US Department of Justice or the National Institute of Standards and Technology. Certain commercial equipment, instruments and materials are identified in order to specify experimental procedures as completely as possible. In no case does such identification imply a recommendation or endorsement by the National Institute of Standards and Technology nor does it imply that any of the materials, instruments or equipment identified are necessarily the best available for the purpose.
# Autosomal STR Loci

<table>
<thead>
<tr>
<th>Chr</th>
<th>STR Locus</th>
<th>Repeat</th>
<th>Allele Range</th>
<th>PowerPlex 16 (HS)</th>
<th>PowerPlex 18D</th>
<th>PowerPlex ES/ESX 16</th>
<th>PowerPlex ES/ESX 17</th>
<th>PowerPlex CS7</th>
<th>PowerPlex Fusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1q31</td>
<td>F13B</td>
<td>AAAT</td>
<td>6 to 11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1q42</td>
<td>D1S1656</td>
<td>TAGA</td>
<td>10 to 19.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2p25.3</td>
<td>TPOX</td>
<td>AATG</td>
<td>5 to 13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2p14</td>
<td>D2S441</td>
<td>TCWA</td>
<td>8 to 17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2q35</td>
<td>D2S1338</td>
<td>TKCC</td>
<td>15 to 27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3p21.31</td>
<td>D3S1358</td>
<td>TCTR</td>
<td>11 to 20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4q31.3</td>
<td>FGA</td>
<td>YTTY</td>
<td>16.2 to 43.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5q23.2</td>
<td>D5S818</td>
<td>AGAT</td>
<td>7 to 15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5q33.1</td>
<td>CSF1PO</td>
<td>AGAT</td>
<td>7 to 15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6p24</td>
<td>F13A01</td>
<td>AAAG</td>
<td>3.2 to 17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6q14</td>
<td>SE33</td>
<td>AAAG</td>
<td>6.3 to 36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6q15</td>
<td>D6S1043</td>
<td>AGAY</td>
<td>8 to 26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7q21.11</td>
<td>D7S820</td>
<td>GATA</td>
<td>6 to 14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8p22</td>
<td>LPL</td>
<td>AAAT</td>
<td>7 to 15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8q24.13</td>
<td>D8S1179</td>
<td>TCTR</td>
<td>8 to 18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9p13</td>
<td>Penta C</td>
<td>AAAAC</td>
<td>5 to 16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10q26.3</td>
<td>D10S1248</td>
<td>GGAA</td>
<td>8 to 19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11p15.5</td>
<td>TH01</td>
<td>TCAT</td>
<td>5 to 11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12p13.31</td>
<td>vWA</td>
<td>TCTR</td>
<td>11 to 21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12p13.2</td>
<td>D12S391</td>
<td>AGAY</td>
<td>14 to 27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13q31.1</td>
<td>D13S317</td>
<td>TATC</td>
<td>8 to 15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15q25</td>
<td>FESFPS</td>
<td>ATTT</td>
<td>5 to 14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15q26.2</td>
<td>Penta E</td>
<td>AAAGA</td>
<td>5 to 25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16q24.1</td>
<td>D16S539</td>
<td>GATA</td>
<td>5 to 15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18q21.33</td>
<td>D18S51</td>
<td>AGAA</td>
<td>9 to 28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19q12</td>
<td>D19S433</td>
<td>WAGG</td>
<td>9 to 18.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21q21.1</td>
<td>D21S11</td>
<td>TCTR</td>
<td>24.2 to 39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21q22.3</td>
<td>Penta D</td>
<td>AAAGA</td>
<td>2.2 to 17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22q12.3</td>
<td>D22S1045</td>
<td>ATT</td>
<td>8 to 19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xp, Yp</td>
<td>Amelogenin</td>
<td>--</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yq11.21</td>
<td>DYS391</td>
<td>TCTA</td>
<td>7 to 13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## Promega STR kits

## Life Technologies (ABI) STR kits

## Qiagen STR kits

**Codis 13 (US 1997-present):**
CODIS 20 (US future)
ESS 12 (EU 2009-present)
PowerPlex 16 (HS)
PowerPlex 18D
PowerPlex ES/ESX 16
PowerPlex CS7
PowerPlex Fusion

**Autosomal STR Kits**

- COfiler
- SGM Plus
- SEfiler Plus
- SinoFiler
- MiniFiler
- Identifiler (Plus)
- VeriFiler
- NGM
- GlobalFiler
- ESSplex
- ESSplex SE
- Hexaplex ESS
- Nonaplex ESS
- Decaplex SE
- IDplex
- 24plex

Butler et al. 2012 required.

Promega STR kits

Life Technologies (ABI) STR kits

Qiagen STR kits
Promega STR Kits (Internal Size Standard CXR ILS600 – 4-dye; CXR ILS 550 – 5-dye)

**PowerPlex 16**
- D3S1358
- TH01
- D21S11
- D18S51
- Penta E
- D5S818
- D13S317
- D7S820
- D16S539
- CSF1PO
- Penta D
- AM
- vWA
- D8S1179
- TPOX
- FGA

**PowerPlex ESI 17 Pro**
- AM
- D3S1358
- D19S433
- D2S1338
- D22S1045
- TH01
- vWA
- D21S11
- D12S391
- D8S1179
- FGA
- SE33

**PowerPlex Fusion**
- AM
- D3S1358
- D1S1656
- D2S441
- D10S1248
- D13S317
- Penta E
- D16S539
- D18S51
- D2S1338
- CSF1PO
- Penta D
- TH01
- vWA
- D21S11
- D7S820
- D5S818
- TPOX
- D12S391
- D8S1179
- D19S433
- FGA
- D22S1045

- 16plex (4-dye)
- 17plex (5-dye)
- 24plex (5-dye)
Life Technologies/ABI STR Kits (Internal Size Standard LIZ GS500 – 5-dye; LIZ GS600 – 6-dye)

16plex (5-dye)
- D8S1179
- D21S11
- D7S820
- CSF1PO
- D3S1358
- TH01
- D13S317
- D16S539
- D2S1338
- D19S433
- vWA
- TPOX
- D18S51
- AM
- D5S818
- FGA

17plex (5-dye)
- D10S1248
- vWA
- D16S539
- D2S1338
- AM
- D8S1179
- D21S11
- D18S51
- D22S1045
- D19S433
- TH01
- FGA
- D2S441
- D3S1358
- D1S1656
- D12S391
- SE33

24plex (6-dye)
- D3S1358
- vWA
- D16S539
- CSF1PO
- TPOX
- Y±
- AM
- D8S1179
- D21S11
- D18S51
- D2S441
- D19S433
- TH01
- FGA
- D22S1045
- D5S818
- D13S317
- D7S820
- D2S1338
- D10S1248
- D1S1656
- D12S391
- SE33
Qiagen Investigator STR Kits (Internal Size Standard 550 BTO – 5th or 6th dye)

16plex (5-dye)
- AM
- TH01
- D3S1358
- vWA
- D21S11
- D7S820
- D19S433
- D5S818
- D2S1338
- D18S51
- CSF1PO
- D13S317
- FGA
- D8S1179

17plex (5-dye)
- AM
- TH01
- D3S1358
- vWA
- D21S11
- D7S820
- D19S433
- D5S818
- D2S1338
- D18S51
- CSF1PO
- D13S317
- FGA
- D8S1179

24plex (6-dye)
- AM
- TH01
- D3S1358
- vWA
- D21S11
- TPOX
- DYS391
- D1S1656
- D12S391
- SE33
- D8S1179
- D2S1338
- D10S1248
- D22S1045
- D19S433
- D2S441
- D18S51
- D13S317
- FGA
- D8S1179
- D7S820
- QS1
- D16S539
- CSF1PO
- D5S818
- QS2
- D7S820
Relative Sizes of STR Loci in 24plex Kits

PowerPlex Fusion
- D16S539
- D18S51
- D2S441
- D2S1338
- CSF1PO
- TH01
- vWA
- D21S11
- D7S820
- D5S818
- TPOX
- DYS391
- D22S1045

GlobalFiler
- Y±
- AM
- D10S1248
- D18S51
- D2S441
- D19S433
- TH01
- vWA
- D21S11
- D16S539
- CSF1PO
- D16S531
- D1S1656
- D12S391
- D19S433
- D7S820
- SE33
- D2S1338

Investigator 24plex
- AM
- TH01
- D3S1358
- vWA
- D16S539
- CSF1PO
- D13S317
- D21S11
- D22S1045
- D10S1248
- D18S51
- D1S1656
- D2S441
- D19S433
- D8S1179
- D2S1338
- D5S818
- SE33
- D7S820
- FGA
- QS1
- QS2
Characteristics of Y-STR Loci and Y-Chromosome Sex-Typing Markers in Commercial Kits

Markers in bold font are the 11 recommended by SWGDAM and present in all kits. DYS391 is present in PowerPlex Fusion and GlobalFiler to aid sex-typing. The Y-chromosome positions were determined using the February 2009 human reference sequence and BLAT (http://genome.ucsc.edu/cgi-bin/hgBlat). *Allele range listed is for PowerPlex Y23 and Yfiler Plus allelic ladder alleles. Information from Butler, J.M. (2015) Advanced Topics in Forensic DNA Typing: Interpretation, Table 1.3.

<table>
<thead>
<tr>
<th>ChrY Position (Mb)</th>
<th>Y-STR Marker</th>
<th>Repeat Motif</th>
<th>Allele Range*</th>
<th>Present in Y-STR Kit</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.13</td>
<td>DYS393</td>
<td>AGAT</td>
<td>7 to 18</td>
<td>PPY, Yfiler, PPY23, Yfiler Plus</td>
</tr>
<tr>
<td>4.27</td>
<td>DYS456</td>
<td>AGAT</td>
<td>11 to 23</td>
<td>Yfiler, PPY23, Yfiler Plus</td>
</tr>
<tr>
<td>6.74</td>
<td>AMEL Y</td>
<td>+AAAGTG</td>
<td></td>
<td>Fusion, GlobalFiler, etc.</td>
</tr>
<tr>
<td>6.86</td>
<td>DYS570</td>
<td>TTTC</td>
<td>10 to 25</td>
<td>PPY23, Yfiler Plus</td>
</tr>
<tr>
<td>7.05</td>
<td>DYS576</td>
<td>AAAG</td>
<td>11 to 23</td>
<td>PPY23, Yfiler Plus</td>
</tr>
<tr>
<td>7.87</td>
<td>DYS458</td>
<td>GAAA</td>
<td>10 to 24</td>
<td>Yfiler, PPY23, Yfiler Plus</td>
</tr>
<tr>
<td>8.22</td>
<td>DYS449</td>
<td>TTTC</td>
<td>22 to 40</td>
<td>Yfiler Plus</td>
</tr>
<tr>
<td>8.43</td>
<td>DYS481</td>
<td>CTT</td>
<td>17 to 32</td>
<td>PPY23, Yfiler Plus</td>
</tr>
<tr>
<td>8.65</td>
<td>DYS627</td>
<td>AAAG</td>
<td>11 to 27</td>
<td>Yfiler Plus</td>
</tr>
<tr>
<td>9.52</td>
<td>DYS19</td>
<td>TAGA</td>
<td>9 to 19</td>
<td>PPY, Yfiler, PPY23, Yfiler Plus</td>
</tr>
<tr>
<td>14.10</td>
<td>DYS391</td>
<td>TCTA</td>
<td>5 to 16</td>
<td>PPY, Yfiler, PPY23, Yfiler Plus, Fusion, GlobalFiler</td>
</tr>
<tr>
<td>14.38</td>
<td>DYS635</td>
<td>TSTA</td>
<td>15 to 28</td>
<td>Yfiler, PPY23, Yfiler Plus</td>
</tr>
<tr>
<td>14.47</td>
<td>DYS437</td>
<td>TCTR</td>
<td>11 to 18</td>
<td>PPY, Yfiler, PPY23, Yfiler Plus</td>
</tr>
<tr>
<td>14.51</td>
<td>DYS439</td>
<td>AGAT</td>
<td>6 to 17</td>
<td>PPY, Yfiler, PPY23, Yfiler Plus</td>
</tr>
<tr>
<td>14.61</td>
<td>DYS389 I/II</td>
<td>TCTR</td>
<td>9 to 17 / 24 to 35</td>
<td>PPY, Yfiler, PPY23, Yfiler Plus</td>
</tr>
<tr>
<td>14.94</td>
<td>DYS438</td>
<td>TTTTC</td>
<td>6 to 16</td>
<td>PPY, Yfiler, PPY23, Yfiler Plus</td>
</tr>
<tr>
<td>15.51</td>
<td>M175</td>
<td>[TTCTC/-]</td>
<td>“1” or “2”</td>
<td>GlobalFiler Y-InDel (Y±)</td>
</tr>
<tr>
<td>17.27</td>
<td>DYS390</td>
<td>TCTR</td>
<td>17 to 29</td>
<td>PPY, Yfiler, PPY23, Yfiler Plus</td>
</tr>
<tr>
<td>17.32</td>
<td>DYS518</td>
<td>AAAG</td>
<td>32 to 49</td>
<td>Yfiler Plus</td>
</tr>
<tr>
<td>17.43</td>
<td>DYS643</td>
<td>CTTTT</td>
<td>6 to 17</td>
<td>PPY23</td>
</tr>
<tr>
<td>18.39</td>
<td>DYS533</td>
<td>ATCT</td>
<td>7 to 17</td>
<td>PPY23, Yfiler Plus</td>
</tr>
<tr>
<td>18.74</td>
<td>GATA-H4</td>
<td>TAGA</td>
<td>8 to 18</td>
<td>Yfiler, PPY23, Yfiler Plus</td>
</tr>
<tr>
<td>20.80, 20.84</td>
<td>DYS385 a/b</td>
<td>GAAA</td>
<td>7 to 28</td>
<td>PPY, Yfiler, PPY23, Yfiler Plus</td>
</tr>
<tr>
<td>21.05</td>
<td>DYS460</td>
<td>ATAG</td>
<td>7 to 14</td>
<td>Yfiler Plus</td>
</tr>
<tr>
<td>21.52</td>
<td>DYS549</td>
<td>GATA</td>
<td>7 to 17</td>
<td>PPY23</td>
</tr>
<tr>
<td>22.63</td>
<td>DYS392</td>
<td>TAT</td>
<td>4 to 20</td>
<td>PPY, Yfiler, PPY23, Yfiler Plus</td>
</tr>
<tr>
<td>24.36</td>
<td>DYS448</td>
<td>AGAGAT</td>
<td>14 to 24</td>
<td>Yfiler, PPY23, Yfiler Plus</td>
</tr>
<tr>
<td>25.93, 28.03</td>
<td>DYF387S1 a/b</td>
<td>RAAG</td>
<td>30 to 44</td>
<td>Yfiler Plus</td>
</tr>
</tbody>
</table>
We will not discuss FBI Project data

- This workshop will NOT discuss Consortium Validation Project data being used by the FBI CODIS Unit in the U.S. core loci expansion

- We will discuss STR loci and what we know about the latest autosomal and Y-STR kits

Product Disclaimer

- We will mention commercial STR kit names and information, but we are in no way attempting to endorse any specific products.

- NIST Disclaimer: Certain commercial equipment, instruments and materials are identified in order to specify experimental procedures as completely as possible. In no case does such identification imply a recommendation or it imply that any of the materials, instruments or equipment identified are necessarily the best available for the purpose.

- Points of view are the speakers and do not necessarily represent the official position of the National Institute of Standards and Technology or the U.S. Department of Justice. The NIST Applied Genetics Group receives or has received funding from the FBI Laboratory and the National Institute of Justice.

http://www.cstl.nist.gov/strbase/training.htm
Three major reasons for expanding the CODIS core loci in the United States

- To reduce the likelihood of adventitious matches as the number of profiles stored at NDIS continues to increase each year
- To increase international compatibility to assist law enforcement data sharing efforts
- To increase discrimination power to aid missing persons cases

International Comparability

There are currently

- 29 autosomal STR markers present in commercial kits
- 13 CODIS loci
- 7 ESS loci

<table>
<thead>
<tr>
<th>China</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 loci adopted in 2009 to expand to 12 ESS loci</td>
</tr>
<tr>
<td>Europe</td>
</tr>
<tr>
<td>13 CODIS loci + 5 additional loci in Powerplex CS7</td>
</tr>
<tr>
<td>3 miniSTR loci developed at NIST</td>
</tr>
</tbody>
</table>

Amelogenin for Sex-Typing

- Deletions and primer site polymorphisms can lead to incorrect sex-typing results
- Amelogenin is located at 6.74 Mb on ChrY (short arm) and 11.31 Mb on ChrX
- Using another marker on the Y-chromosome can help verify male DNA samples (e.g., DYS391)

Why Consider DYS391?

- DYS391 is located on the long arm of the Y-chromosome over 7 Mb away from amelogenin. Thus, it is likely to be detected in the event of an amelogenin Y deletion that could make a male sample falsely appear as a female (X-).
- DYS391 is not very polymorphic. From a data set of 97,575 haplotypes available on the Y-chromosome Haplotype Reference Database, over half of them possess allele 10. However, only two null alleles have been reported and 0.01% duplication events (11 total) have been seen in over 700 different population groups from around the world. Thus, it is a stable locus with a relatively narrow allele range.
- DYS391 has a mutation rate of 0.26%, which is comparable to most autosomal STRs commonly in use. There have been 38 mutations observed so far in the 14,621 meioses reported in the literature and compiled on YHRD.

Novel Y-indel in GlobalFiler Kit

- Can be either “1” (deletion) or “2” (insertion)
- Small size (81 or 86 nt) enabling successful results with degraded DNA samples
- Likely an insertion/deletion (InDel) known as M175 (175th marker discovered by Peter Underhill from Stanford University using denaturing HPLC)
- Exhibits deletion of “TTCTC” with Y-SNP Haplogroup O individuals (East or SE Asians)
- See van Oven et al. (2012) J Human Genet 57: 65-69
- Most samples will be “2” (the ancestral “insertion” form) unless they are Asian in origin

Relative positions of amelogenin (AMEL Y), DYS391, and Y-indel in GlobalFiler

Deletions of the Y-chromosome can encompass >1 Mb around the AMEL Y region (DYS458 is often lost in these situations)

http://www.cstl.nist.gov/strbase/training.htm
Reference List Compiled for Workshop

268 Articles and Websites

<table>
<thead>
<tr>
<th>Autosomal STR Topics</th>
<th>#</th>
<th>Y-STR Topics</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>European &amp; US Core Loci</td>
<td>10</td>
<td>SWGDAM Y-STR Guidelines</td>
<td>1</td>
</tr>
<tr>
<td>STR kits &amp; new assays</td>
<td>38</td>
<td>Y-STR haplotype databases</td>
<td>9</td>
</tr>
<tr>
<td>NIST U.S. population data</td>
<td>6</td>
<td>Y-STR kits</td>
<td>5</td>
</tr>
<tr>
<td>On-line population databases</td>
<td>6</td>
<td>PowerPlex Y23 population data</td>
<td>5</td>
</tr>
<tr>
<td>Population data on new STR loci</td>
<td>25</td>
<td>Rapidly mutating (RM) Y-STRs</td>
<td>3</td>
</tr>
<tr>
<td>Information on STR loci</td>
<td>28</td>
<td>Early Y-STR work at NIST</td>
<td>11</td>
</tr>
<tr>
<td>Concordance studies</td>
<td>15</td>
<td>Impact of additional Y-STR loci</td>
<td>14</td>
</tr>
<tr>
<td>Next-Gen Sequencing of STRs</td>
<td>13</td>
<td>Y-STR mutations</td>
<td>26</td>
</tr>
<tr>
<td>Amelogenin &amp; anomalies</td>
<td>28</td>
<td>Y-STR profile anomalies</td>
<td>8</td>
</tr>
<tr>
<td>D12S391 &amp; VWA studies for LD</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Potential Disease Linkage?</td>
<td>12</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

186 82

Relative Sizes of STR Loci in 24plex Kits

<table>
<thead>
<tr>
<th>Loci</th>
<th>600 bp</th>
<th>400 bp</th>
<th>200 bp</th>
</tr>
</thead>
<tbody>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
<tr>
<td>AM</td>
<td>DYS516</td>
<td>DYS509</td>
<td>DYS506</td>
</tr>
</tbody>
</table>
Different Internal Size Standards

- **Life Technologies/ABI**
  - GS500-ROX and GS500-LIZ
  - GS600-LIZ

- **Promega**
  - ILS-500

- **Qiagen**
  - 550 BTO

**Local Southern sizing requires two size standard peaks on either size of measured allele**
- **20 bp spacing**
- **25 bp spacing**

Labeled with 5-dye (orange)
- Contains 21 fragments with even spacing
- Low end: 60 & 65 bp
- High end: 475 & 500 bp

**Local Southern sizing possible from 66 bp to 474 bp**

**Questions for Workshop Participants**

- **STR kit(s) in your lab?**
  - Currently in use: Identifiler, PP16, Pro/CO
  - Considering: Fusion, GlobalFiler, other

- **Y-STR kit(s):**
  - PPY, PPY23, Yfiler, Yfiler Plus

- **CE instrument(s)?**
  - Currently: ABI 310, ABI 3130xl, ABI 3500
  - Considering: 3500, 3130xl (6-dye conversion)

- **Analysis software?**
  - GeneMapperID, GMID-X, GeneMarkerHID, OSIRIS

**Acknowledgments**

$ NIST Office of Special Programs
(and previous funding from National Institute of Justice)

Becky Hill and Mike Coble
(NIST Applied Genetics Group)

**Contact info:**

john.butler@nist.gov

+1-301-975-4049

Final version of this presentation available at: http://www.cstl.nist.gov/strbase/NISTpub.htm
Presentation Outline

- STR kits (including Fusion and GlobalFiler)
- NIST U.S. population samples
- Kit concordance study design
- Concordance study results
- U.S. allele frequencies – FSI Genetics article

GlobalFiler STR Kit

Launched Friday, September 14, 2012

- 24 STR loci in 6 dyes (3500 use or 3130 upgrade required)
  - Includes SE33 and a Y-indel
- GlobalFiler Express: direct amplification capabilities
  - Single source samples: 40 min amplification
  - GlobalFiler Casework
    - Casework samples: 80 min amplification
    - GlobalFiler gives ~12 orders of magnitude improvement using the NIST 1036 data set

http://www.cstl.nist.gov/strbase/training.htm
SRM 2391c is fully concordant at all loci for GlobalFiler kit – Component A Profile

1 ng, 29 cycles, 3500xl

SRM 2391c Mixture Component D

1 ng DNA, 29 cycles, 3500xl

PowerPlex Fusion

Launched Friday, September 14, 2012


• 24 STR loci in 5 dyes (3130 and 3500 instrument use)
  – Includes Penta D and E
• Direct amplification and casework capabilities: 85 min amp for both (one kit)
• PowerPlex Fusion gives ~13 orders of magnitude improvement using the NIST 1036 data set


SRM 2391c is fully concordant at all loci for PP Fusion kit – Component A Profile

1 ng DNA, 30 cycles, 3130xl

SRM 2391c Mixture Component D

1 ng DNA, 30 cycles, 3130xl
C.R. Hill - ISHI Workshop on New Loci & Kits
NIST Studies: Kit Concordance & U.S. Population Data

October 2, 2014

Qiagen Investigator 24plex

Available worldwide from Oct 2014 and not before Oct 2015 in the U.S.

There’s been a sample data profile — Global STR Analysis Including Quality Control


Qiagen Investigator 24plex

Available worldwide from Oct 2014 and not before Oct 2015 in the U.S.

<table>
<thead>
<tr>
<th>AM</th>
<th>TH01</th>
<th>D3S1358</th>
<th>D16S539</th>
<th>D5S818</th>
<th>D2S1338</th>
</tr>
</thead>
<tbody>
<tr>
<td>D18S51</td>
<td>D21S11</td>
<td>D2S441</td>
<td>D19S433</td>
<td>D5S380</td>
<td>D8S1179</td>
</tr>
<tr>
<td>D22S1045</td>
<td>D19S433</td>
<td>D16S539</td>
<td>D21S11</td>
<td>D5S380</td>
<td>D8S1179</td>
</tr>
<tr>
<td>D8S1179</td>
<td>D21S11</td>
<td>D3S1358</td>
<td>D5S380</td>
<td>D8S1179</td>
<td>D21S11</td>
</tr>
<tr>
<td>D8S1179</td>
<td>D21S11</td>
<td>D3S1358</td>
<td>D5S380</td>
<td>D8S1179</td>
<td>D21S11</td>
</tr>
<tr>
<td>D8S1179</td>
<td>D21S11</td>
<td>D3S1358</td>
<td>D5S380</td>
<td>D8S1179</td>
<td>D21S11</td>
</tr>
</tbody>
</table>

- 23 STR loci in 6 dyes (3500 use or 3130 upgrade required)
  - Includes SE33
- Includes Quality Sensors for detecting degraded and/or inhibited DNA
- 24plex GO!: direct amplification capabilities
  - Single source samples: 45 min amplification
  - 24plex: Casework samples: 60 min amplification
- Qiagen 24plex gives ~12 orders of magnitude improvement using the NIST 1036 data set

Two separate kits

SRM 2391c is **fully concordant** at all loci for 24plex kit — **Component A Profile**

1 ng DNA, 30 cycles, 3500xI

SRM 2391c Mixture Component D

1 ng DNA, 30 cycles, 3500xI

NIST U.S. Samples (>1450)

- **NIST U.S. population samples**
  - 260 African American, 260 Caucasian, 140 Hispanic, 3 Asian
- **U.S. father/son paired samples**
  - ~100 fathers/100 sons for each group: 200 African American, 200 Caucasian, 200 Hispanic, 200 Asian
- **NIST SRM 2391b**, PCR-based DNA Profiling Standard (highly characterized)
  - 10 genomic DNA samples, 2 cell line samples
  - Includes 9947A and 9948
- **NIST SRM 2391c**, PCR-based DNA Profiling Standard
  - 4 genomic DNA (one mixture)
  - 2 cell lines (903 and FTA paper)


http://www.cstl.nist.gov/strbase/training.htm
NIST Studies: Kit Concordance & U.S. Population Data

17. Lue et al. (2010) – Human Mutation – 24 ancestry SNPs, Y-SNPs, mtDNA
27. Baloghy et al. (2014) – Human Mutation – 12 rapidly mutating (RM) Y-STRs with global population

Benefits of NIST 1036 Data Set

- Elimination of potential null alleles due to primer binding site mutations through extensive concordance testing performed with different PCR primer sets from all available commercial STR kits
- Ancestry testing performed on DNA samples with autosomal SNPs, Y-SNPs, and mtDNA sequencing to verify self-declared ancestry categorization
- Related individuals removed based on Y-STR and mtDNA results

Concordance Testing at NIST

STRA Kit Concordance Testing

- Many of these STR kits have different primer sequences for amplifying the same STR locus
- Need to analyze the same DNA samples with different STR typing kits looking for differences
- In some rare cases, allele dropout may occur due to mutations in primer binding regions

Purpose of Concordance Studies

When different primer sets are utilized, there is a concern that allele dropout may occur due to primer binding site mutations that impact one set of primers but not another.

- Use of non-overlapping primers
- Permits detection of allele dropout

Data available on STRBase: http://www.cstl.nist.gov/strbase/NISTpop.htm

NIST 1036 U.S. Population Samples

- 1032 males + 4 females
- 97 Asians (1 female)
- 361 Caucasians (3 females)
- 342 African Americans (1 female)
- 236 Hispanics

Unrelated samples

All known or potential related individuals (based on autosomal & lineage marker testing) have been removed from the 1036 data set (e.g., only sons were used from father-son samples)

Additional DNA results available on subsets of these samples
- mtDNA control region/whole genome (AFDIL)
- mtGenome
- mtDNA sequencing to
- STRs in PowerPlex Y23 with global population comparison

http://www.cstl.nist.gov/strbase/strbase.html
Example Primer Binding Site Mutation that Causes a Null Allele

G → A

This region could potentially represent where the reverse primer is located to include the primer binding site mutation.

Applied Biosystems does not publish their primer sequences.

The 4 “S”s of Concordance

- NIST Standard Samples
  - Run same samples with multiple kits to compare results
- Concordance Software
  - Allows comparison of data sets using NIST developed software
- DNA Sequencing
  - To validate and determine the exact cause for the null allele
- STRBase website
  - To report verified null alleles and discordant results to the forensic community

STR Kit Concordance Testing

Profiles in DNA Article Published April 2010

NIST Concordance Testing Steps

Kit 1

PP-ESX17 Sample Set

Results

Comparisons with software

Discordant Results

Sequence differences

Sequence differences

Vendor changes final kit to correct for primer binding site mutation

Vendor decides not to change the final configuration of the kit

Applied Biosystems AmpFlSTR Kits

- Identifier
- Minifiler
- Profiler Plus
- SGM Plus
- NGM
- NGM SElect
- Yfiler Plus

GlobalFiler only examined with 50 bloodstains

Completed Concordance Studies

http://www.cstl.nist.gov/strbase/training.htm
Promega PowerPlex Systems
- PowerPlex 16/16HS
- **PowerPlex ESX 17** (& Fast)
- **PowerPlex ESI 17** (& Fast)
- PowerPlex ESI 17 Pro
- PowerPlex 18D (rapid and direct kit)
- PowerPlex 21
- PowerPlex Fusion
- PowerPlex Y23

Qiagen Investigator HID Kits
- ESSplex
- ESSplex Plus
- ESSplex SE
- ESSplex SE Plus
- Hexaplex ESS
- IDplex
- IDplex Plus
- 24plex
- 24plex GO!

Completed Concordance Studies

<table>
<thead>
<tr>
<th>Sample Type</th>
<th>Loci Compared</th>
<th># Allele Comparisons</th>
<th># Differences</th>
<th>Concordance (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>1,404,031</td>
<td>1,373</td>
<td>99.9%</td>
<td></td>
</tr>
</tbody>
</table>

1,404,031 allele comparisons
1,373 total differences
99.9% concordance

Kits (except identifier) were kindly provided by Promega, Qiagen and Applied Biosystems for concordance testing performed at NIST

Final Concordance Results
- All up-to-date results can be found on STRBase:
  - ISFG poster (Vienna, Austria), 8/31-9/2, 2011, “Concordance Testing Comparing STR Multiplex Kits with a Standard Data Set”
  - Promega ISHI (National Harbor, MD), 10/4-10/5, 2011, “Concordance Testing Comparing STR Multiplex Kits with a Standard Data Set”

NIST SRM 2391b/2391c
PCR-Based Profiling Standard
- The first set of samples run with new STR multiplex kits is SRM 2391b/SRM 2391c
- All new kits tested have been completely concordant with the certified values of all markers for each component for SRM 2391b and SRM 2391c
- One exception for SRM 2391b: MiniFiler – Genomic 8 with D16S539

SRM 2391b Genomic 8 with D16S539
All allele calls with MiniFiler for CSF1PO, D7S820, D13S317, D18S51, D21S11, FGA, and D16S539 (with the exception noted below) match previously certified values.

MiniFiler
- Null Allele
- D16S539

PowerPlex 16
- 24plex
- 24plex GO!

*Due to primer binding site mutation

http://www.cstl.nist.gov/strbase/training.htm
Components A through D are DNA extracts in liquid form. Certified values are for STR alleles based on length.

Reference Values

Certified Values

Components E and F are DNA spotted on 903 paper or FTA paper.

http://www.cstl.nist.gov/strbase/training.htm

Certified, Reference, & Information Values of SRM 2391c

- **Certified Values**: Value for which NIST has the highest confidence in its accuracy in that all known or suspected sources of bias have been investigated or taken into account
  - 2 or more methods are used to compare values (i.e., Sanger sequencing, genotyping using multiple sets of primers)
- **Reference Values**: High-confidence estimate of the true value but where all possible sources of bias have not been fully investigated by NIST
  - Genotyping with only 2 sets of primers to compare
- **Information Values**: Data that may be of interest and use to the SRM user, but insufficient information is available to access the confidence of the assignment
  - Genotyping of only 1 kit is available

STR Typing Kits and Primer Mixes

D16S539 SRM 2391b Genomic 8

T→C mutation 34 bp downstream of the repeat

Position of the T→C probably affects the reverse primer of Miniliter and is the 3rd base found the 5’end of the Reverse PP16 primer. This could explain the imbalance of the allele seen when using PP16.

SRM 2391c:

PCR-Based DNA Profiling Standard

- Components A through D are DNA extracts in liquid form
- Components E and F are DNA spotted on 903 paper or FTA paper
- Certified values are for STR alleles based on length

Certification with SRMs enables confidence in comparisons of results between laboratories

Current Values for STR Loci with SRM 2391c

Updated Values for STR Loci

New Y-STR loci in commercial kits (Yfiler Plus & PPy23)

Update to be completed by Oct. 2014

http://www.cstl.nist.gov/strbase/training.htm
Concordance Testing at NIST

- Concordance testing is valuable when different sets of primers are used to amplify the same markers.
- Null alleles and discordant results are reported on STRBase:
  
  http://www.cstl.nist.gov/biotech/strbase/NullAlleles.htm

- NIST plays an important role in concordance testing to aid the community:
  - SRM 2391b&c concordance
  - Several null alleles have been fixed before the final release of new STR multiplex kits

---

Characterization of STR Loci

Available in Commercial Kits

---

The 10 STR Loci Beyond the CODIS 13

<table>
<thead>
<tr>
<th>STR Locus</th>
<th>Location</th>
<th>Repeat Motif</th>
<th>Allele Range*</th>
<th># Alleles*</th>
</tr>
</thead>
<tbody>
<tr>
<td>D2S1338</td>
<td>2q35</td>
<td>TGGG/TTGC</td>
<td>10 to 31</td>
<td>40</td>
</tr>
<tr>
<td>D19S433</td>
<td>1q12</td>
<td>AAGG/TTAGG</td>
<td>5.2 to 20</td>
<td>36</td>
</tr>
<tr>
<td>Penta D</td>
<td>2q22.3</td>
<td>AAAGA</td>
<td>1.1 to 19</td>
<td>50</td>
</tr>
<tr>
<td>Penta E</td>
<td>15q26.2</td>
<td>AAAGA</td>
<td>5 to 32</td>
<td>53</td>
</tr>
<tr>
<td>D1S1656</td>
<td>1q42</td>
<td>TAGA</td>
<td>8 to 20.3</td>
<td>25</td>
</tr>
<tr>
<td>D12S391</td>
<td>12p13.2</td>
<td>AGATA/GAGAC</td>
<td>13 to 27.2</td>
<td>52</td>
</tr>
<tr>
<td>D2S441</td>
<td>2p14</td>
<td>TCTA/TCTAA</td>
<td>8 to 17</td>
<td>22</td>
</tr>
<tr>
<td>D10S1248</td>
<td>10q26.3</td>
<td>GGAA</td>
<td>7 to 19</td>
<td>13</td>
</tr>
<tr>
<td>D22S1045</td>
<td>22q12.3</td>
<td>ATT</td>
<td>7 to 20</td>
<td>14</td>
</tr>
<tr>
<td>SE33</td>
<td>6q14</td>
<td>AAAGA</td>
<td>3 to 49</td>
<td>178</td>
</tr>
</tbody>
</table>

*Allele range and number of observed alleles from Appendix 1, J.M. Butler (2012) Advanced Topics in Forensic DNA Typing: Methodology; 1SE33 alleles have complex repeat structure

---

NIST U.S. Population Allele Frequencies for D1S1656 (15 different alleles)

<table>
<thead>
<tr>
<th>Allele</th>
<th>African American (n=342)</th>
<th>Asian (n=97)</th>
<th>Caucasian (n=361)</th>
<th>Hispanic (n=236)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.0146</td>
<td>0.0000</td>
<td>0.0028</td>
<td>0.0064</td>
</tr>
<tr>
<td>11</td>
<td>0.0453</td>
<td>0.0309</td>
<td>0.0776</td>
<td>0.0275</td>
</tr>
<tr>
<td>12</td>
<td>0.0643</td>
<td>0.0464</td>
<td>0.1163</td>
<td>0.0890</td>
</tr>
<tr>
<td>13</td>
<td>0.1009</td>
<td>0.1340</td>
<td>0.0665</td>
<td>0.1144</td>
</tr>
<tr>
<td>14</td>
<td>0.2573</td>
<td>0.0619</td>
<td>0.0789</td>
<td>0.1165</td>
</tr>
<tr>
<td>14.3</td>
<td>0.0073</td>
<td>0.0000</td>
<td>0.0028</td>
<td>0.0042</td>
</tr>
<tr>
<td>15</td>
<td>0.1579</td>
<td>0.2784</td>
<td>0.1496</td>
<td>0.1377</td>
</tr>
<tr>
<td>15.3</td>
<td>0.0292</td>
<td>0.0000</td>
<td>0.0058</td>
<td>0.0508</td>
</tr>
<tr>
<td>16</td>
<td>0.1096</td>
<td>0.2010</td>
<td>0.1357</td>
<td>0.1758</td>
</tr>
<tr>
<td>16.3</td>
<td>0.1023</td>
<td>0.0155</td>
<td>0.0069</td>
<td>0.0508</td>
</tr>
<tr>
<td>17</td>
<td>0.0278</td>
<td>0.0722</td>
<td>0.0471</td>
<td>0.0424</td>
</tr>
<tr>
<td>17.3</td>
<td>0.0497</td>
<td>0.0876</td>
<td>0.1330</td>
<td>0.1483</td>
</tr>
<tr>
<td>18</td>
<td>0.0029</td>
<td>0.0155</td>
<td>0.0055</td>
<td>0.0064</td>
</tr>
<tr>
<td>18.3</td>
<td>0.0234</td>
<td>0.0515</td>
<td>0.0499</td>
<td>0.0254</td>
</tr>
<tr>
<td>19.3</td>
<td>0.0073</td>
<td>0.0052</td>
<td>0.0152</td>
<td>0.0042</td>
</tr>
</tbody>
</table>

N=1036 (only unrelated samples used; fathers removed from this sample set)

15 different alleles

15 different alleles

---

D1S1656 Characteristics

- 15 alleles observed
- 93 genotypes observed
- >89% heterozygotes (heterozygosity = 0.889)
- 0.0224 Probability of Identity ($P_i$)

$$P_i = \sum (\text{genotype frequencies})^2$$

These values have been calculated for all 29 STR loci across the U.S. population samples examined

---

http://www.cstl.nist.gov/strbase/training.htm
NIST U.S. Population Data

- The data from our 1036 U.S. population samples is currently available on STRBase:
  http://www.cstl.nist.gov/biotech/strbase/NISTpop.htm

- A summary of the NIST 1036 data set has been published in Profiles in DNA for autosomal and YSTR loci

- Population Data announcements have been published in FSI: Genetics for
  - 29 autosomal STR loci (Hill et al)
  - 23 Y-STR loci (Coble et al)

- Added to CODIS PopStats software in Sept 2013

**Summary**

- Additional STR loci are important as DNA databases grow larger each year: the power of discrimination increases as new loci are added
  - Adding seven new loci (CODIS 13 vs CODIS 20) adds approximately 8 orders of magnitude improvement

- Commercial companies are continuing to release larger STR multiplex kits to meet the needs of the forensic community

- NIST has a set of 1036 unrelated U.S. population samples that have been used to fully characterize 29 autosomal STR loci available in commercial STR multiplex kits

**Acknowledgments**

- National Institute of Justice, FBI, and NIST
- Promega, Life Technologies, Qiagen for kits
- John Butler (NIST Office of Special Programs)
- Mike Coble (NIST Applied Genetics Group)
- Pete Vallone (NIST Applied Genetics Group)

Contact info:
becky.hill@nist.gov
301-975-4275

Final version of this presentation available at:
http://www.cstl.nist.gov/strbase/strbase/ISHIWorkshop.htm

October 2nd, 2014
Experience with PowerPlex Fusion

Jeffrey Nye
Michigan State Police

Final version of this presentation will be available at:
http://www.cstl.nist.gov/strbase/training.htm

Michigan State Police
STR Kit Chemistry History

- Transitioned from RFLP directly to STRs
- Initial casework kits included AMPF/STR Blue, Green I and II
  - Provided 9 STR Loci
- Casework moved toProfiler Plus/Cofiler
  - Provided 13 STR Loci
- Database group utilized PowerPlex 16
  - Provided 13 STR Loci plus Penta E and Penta D
- Casework moved to PowerPlex 16 HS in 2011
  - Interest in capturing Penta E and Penta D markers
  - Detailed comparison of Identifiler Plus and PowerPlex 16 HS

PowerPlex Fusion

- Interest in looking at extended STR Loci panels
- Currently use 3130 and 3130xl Genetic Analyzer platforms
  - Contracts to purchase 3500s being established
- Offender database of 300,000+ include Penta loci

INTERNAL VALIDATION

- Casework and Database units
- Studies completed include:
  - Precision
  - Sensitivity
  - Baseline noise evaluation
  - Contamination Assessment
  - Mock Casework
  - Mixtures
  - Concordance
  - NIST SRM

Precision

Ladders were run multiple times over a few days. All capillaries were tested.
A table in GMID was created and exported as a tab-delimited file for import into Microsoft Excel
The average bp size and standard deviation were calculated for each allele of each locus

In general:
Smaller loci are averaging differences approximately 0.03 to 0.05bp
Larger loci (FGA, PentaE, PentaD) are averaging 0.08-0.09bp
Well below 0.5bp window
Contamination Assessment and Baseline Noise Evaluation

- 40 Reagent Blanks were amplified
  - Stain Extraction Blanks
  - Maxwell Extraction Blanks
  - Epithelial Extraction Blanks
  - Sperm Fraction Blanks
- Run on 3130
- Analyzed at 5 rfu
- GMID table created and exported to Excel
- Average Peak height was 5.8rfu +/- 3.2 rfu
- A few peaks were labeled at approximately 25 rfu

Sensitivity

- Samples less than 0.5 ng were amplified
  - 12 times each
- Data was reviewed for imbalance and dropout
- Full profiles consistently down to 125pg
- Increase in concentration = better PHR

Sensitivity

- 3 samples of extracted DNA were used
  - Many heterozygous loci
- Samples were quantified using Plexor HY
- Serial dilutions were made
- Each dilution was quantified in triplicate
- Replicates were amplified. (Range of >1ng to ~ 8 pg)
Mixtures prepared from extracted reference sample DNA

Ratios from 100:0 to 0:100 were prepared with an input template amount of 0.75ng

Amplified product was ran on a 3130 with a standard 5 second injection time

Minor donor detected at all mixture ratios

Fortunate to be able to use actual casework samples

More than 40 cases have been analyzed with Fusion-only able to show a small portion here

Concordant data overall
Swab of a water bottle

PP16HS: Swab of Hat

Fusion: Swab of Hat

PP16HS: Fingernail Scrapings
5 total markers interpretable

Fusion: Fingernail Scrapings
10 total markers interpretable
NIST sample set 2931b was amplified with Fusion chemistry. All profiles obtained were concordant with the published profiles for the various components provided in the kit.

INTERNAL VALIDATION
• NIST Standards Concordance

Internal Validation

INTERNAL VALIDATION
• NIST Standards Concordance

INTERNATIONAL VALIDATION
• Genomic DNA 3

INTERNAL VALIDATION
• NIST Standards Concordance
– Genomic DNA 10 (GM09948)

INTERNAL VALIDATION
• NIST Standards Concordance
– Genomic DNA 10 (GM09948)

INTERNAL VALIDATION
Direct Amplification of Knowns
• Vary amount of swab
  o Cotton Swab – 1 swab, ½ swab, ¼ swab
  o Omni Swab – 2 teeth, 1 tooth, ½ tooth

INTERNATIONAL VALIDATION
Direct Amplification of Knowns
• Cycle Number Optimization
  • 30 Cycles / 27 Cycles / 25 Cycles

INTERNATIONAL VALIDATION
Direct Amplification of Knowns
• Procedure:
  o Add 1 mL of swab solution to swab amounts listed above.
  o Place samples in 70 C hot plate for 30 minutes.
  o Use 1 ul of solution for amp 1 and 2 ul solution for amp 2.
  o Amp samples at same cycle / extension as evidence cycle parameters. (30 cycles)
  o Additional Amps performed at 27, 26, and 25 Cycles

http://www.cstl.nist.gov/strbase/training.htm
Artifact Assessments

Additional Studies
- Comparison of artifacts

Additional Studies
- Comparison of artifacts

Additional Studies
- Comparison of artifacts

Additional Studies
- Comparison of artifacts
Considerations for Selecting a Kit

- One kit vs. two kits
- 5 dye vs. 6 dye
- Familiarity
- Software changes
- Ease of Direct Amplification
- Support
- Use of Pentas
- Cost
- Required CODIS loci (No changes yet)

ACKNOWLEDGEMENTS

Thank you to Promega Corporation for asking us to participate and all of the other collaborators on the project.

A special thanks to Kirk DeLeuw, Josh Strong, Donald Yet and Kristin Schelling for all of the effort in conducting the internal validation.
Internal Validation of PowerPlex® Fusion

Hope Olson
ND Office of Attorney General
Crime Laboratory Division

Final version of this presentation will be available at:
http://www.cstl.nist.gov/strbase/training.htm

History of Kits in ND

- 2000 AmpFISTR® Cofiler/Profiler Casework and Database, 310 Genetic Analyzer
- 2007 AmpFISTR® Identifiler Database, 3130 Genetic Analyzer
- 2008 AmpFISTR® Identifiler Casework, 3130 Genetic Analyzer
- 2008 Y-Filer™ Casework, 3130 Genetic Analyzer
- 2013 Direct Amp Fusion Database, 3500 Genetic Analyzer
- 2014 PowerPlex® Fusion Casework, 3500 Genetic Analyzer, and ArmedXpert™

Considerations in Selecting a Kit

- Cost
- Training
- Changes in Loci for CODIS (Not Effective Yet)
- Sensitivity
- Robustness
- One Kit or Two Kits
- Ease of Direct Amplification
- NDIS Approval

Software Changes
Instruments
5 dyes vs. 6 dyes

Change?

- ND’s IT policy mandates that every computer had to be compatible with Windows 7 – April, 2014
- Platforms had to change while still processing casework (3130 Genetic Analyzers and GMID to 3500 Genetic Analyzer with GMID-X)
- Purchased one 3500 Genetic Analyzer and started training two new analysts last year
- We just purchased a second 3500 Genetic Analyzer in June, 2014

Internal Validation

- Database and Casework Studies completed:
  - Optimized Cycle Number
  - Precision and Reproducibility
  - Known and Mock Case Samples
  - Sensitivity and Stochastic
  - Mixture Studies
  - Contamination Assessment

Optimized Cycle Number Database

- Direct Amplification Procedure
  - Add 400 µl Swab Solution
  - Incubate for 60 minutes at 90°C
  - Amplify 1 µl extract
  - Half Reaction Volume
  - Amplify using 25 cycles
    - We tried 25, 26, and 27 cycles
Optimized Cycle Number

Casework

- Started out validating 30 cycles
- Evaluated data and reduced cycle number to 29
- Full Reaction Volume

Precision and Reproducibility

Ladders

- The average standard deviation across all loci was less than 0.064 bps
- Generally, the loci with the largest sizes: TPOX D10S1248, FGA, Penta E, Penta D, and DYS391 had alleles with the greatest standard deviations (0.048, 0.051, 0.051, 0.052, 0.060, and 0.063)
- All well below 0.5 bp sizing window

Database Known Samples

- NIST Standard 2391 B was injected at 8, 12, 18, and 24 seconds
  - Average std deviation across all loci was 0.027
- Positive Control 2800 injected 8 times over different time periods
  - Average std deviation across all loci was 0.105 bps
  - Range of values 0.036 to 0.156 bps
  - Values taken from two different columns
- Both Sample Sets demonstrated precision and reproducibility
- 46 Previously Analyzed Known Samples were compared with Identifiler were concordant at the loci examined

Database Analytical Threshold

Direct Amplification

- 41 amplification blanks were analyzed at 50 rfus
- Average baseline was 63 rfus
- Standard Deviation 20 rfus
- Minimum Analytical Threshold set at 150 rfus
Casework Analytical Threshold

17 amplification blanks were analyzed at 25 rfus

Each Dye Channel was evaluated:
- Blue (37 rfus average; 12 rfus std dev) = 73 (3 std dev + mean)
- Green (42 rfus average; 23 rfus std dev) = 111 (3 std dev + mean)
- Red (41 rfus average; 27 rfus std dev) = 122 (3 std dev + mean)
- Yellow (37 rfus average; 16 rfus std dev) = 85 (3 std dev + mean)

The analytical threshold for analyzing casework was set to 150 rfus to encompass three standard deviations plus the mean for each of the dye channels.

Database Stochastic

Direct Amplification
- NIST Standard 2391 A, B, and C were amplified in triplicate using 7.8, 15.6, 31.25, 63.5, 125, 250, 500, and 1000pg
- Samples run at 8, 12, 18, and 24 second injections
- Report Manger in GeneMapper®ID-X was used and a tab delimited file was imported into Excel to determine peak height ratios and drop-out for the sister allele.

Database Stochastic

• NIST Standard 2391 Components A, B, and C had maximum stochastic values of 305, 320, and 420
• Peak height ratios were set at 60%
• An additional study was performed by varying the amount of the buccal swab in duplicate which yielded a stochastic level of 484
  — Stochastic level was set at 550 rfus

Casework Stochastic

• Three known samples were amplified in triplicate using 31.25, 65.5, 125, 250, 500, and 1000 picograms
• Run on a 3500 Genetic Analyzer utilizing four different injection parameters (8, 12, 18, and 24 seconds)

Casework Stochastic

Reviewing each individual dye channel the following levels were identified:
- Blue 878 rfus
- Green 1142 rfus
- Yellow 1106 rfus
- Red 1315 rfus

Level was set at 1350 rfus, PHR set at 63%
Target Amount 250 pg to 1000 pg

Casework Stochastic

Peak Height Ratio vs Input DNA

- 68.7% 63.3%
- 53.9%
- 45.5%
- 36.0%
- 21.7%
Casework Stochastic

Looking at the heat maps full profiles were consistently obtained at 125 picograms. (Some even at 65.5 picograms)

- 31.25 pg
- 65.5 pg
- 125 pg
- 250 pg

Casework Mixtures

Mixtures were prepared in duplicate using the following ratios with 1.0 ng of Input DNA:

- 1:19, 1:9, 1:4, 1:2, 1:1, 2:1, 4:1, 9:1, 19:1 Male:Male
- 1:19, 1:9, 1:4, 1:2, 1:1, 2:1, 4:1, 9:1, 19:1 Female:Male
- Each amplified product was subjected to four injection times (8, 12, 18, and 24 seconds)

ArmedXpert™ was used to separate the mixtures and handle the data.

Full profiles at 65.5 picograms

Casework Mixtures

Alleles from the minor contributor were detected in all mixture proportions:

Casework Mixtures

Known and Non-Probative Samples

Casework

- 36 Known samples
- 24 Non-probative mock casework samples
- NIST 2391 A
- 2 proficiency tests

All samples yielded concordant results for the loci compared.
Mock Casework

- 24 Non-Probative samples
  - Cigarette butt
  - Baseball cap
  - T-shirt swab
  - Urine swab
  - Feces swab
  - Steering wheel swab (2)
  - Blood swab (8 samples)
  - Swab of a coat hanger
  - Swab of ignition
  - Swab of shoe
  - Swab of a knife handle
  - Semen samples (2)
  - Fingernail scrapings

Mock Casework

Comparing AmpFISTR® Identifier with PowerPlex® Fusion

Contamination Assessment

Two methods were used to assess contamination for both Direct Amplification and Casework

- Samples were set up using a BiomekNXp in a Striped and Checkerboard configuration
- 64 known samples were distributed across two 96 well plates
- No detectable alleles above 150 rfu's were present in the blank wells when the samples were amplified with PowerPlex® Fusion

PowerPlex® Fusion Casework

PowerPlex® Fusion Casework
PowerPlex® Fusion Casework

Acknowledgements
ND Office of Attorney General
Crime Laboratory Division
DNA Unit:
  Amy Gebhart
  Jennifer Penner
  Stephanie Maier
  Shannon Johnson
  Alexandria Gibbs
  Kyle Splichal
  Emily Verstraete
Alabama’s Internal Validation of the GlobalFiler STR Kit for Forensic Casework

Jason Kokoszka, PhD
Forensic Biology Section Chief
Alabama Department of Forensic Sciences
Mobile Regional Laboratory

FBI’s Quality Assurance Standards for Forensic DNA Testing Laboratories

For Internal Validation Studies:
1. Have all internal validation studies been documented and summarized?
2. Have all internal validation studies conducted on or after July 1, 2000, included, as applicable:
   - Known and non probative evidence samples or mock evidence samples?
   - Reproducibility and precision?
   - Sensitivity and stochastic studies?
   - Mixture studies?
   - Contamination assessment?

Internal Validation Studies - Background

Current Kit (Casework):
- Identifiler
- Detection Platform: 3130 Genetic Analyzer
- 3130 Analytical Threshold: 75 rfu

Test Kit:
- GlobalFiler
- Detection Platform: 3500 Genetic Analyzer (8 capillary)
- 3500 Analytical Threshold: 175 rfu – starting point

First and Foremost…

PCR Cycle Number and Post-Amplification Conditions Evaluation

- 27, 28, and 29 amplification cycles were evaluated:
  - Manufacturer’s recommendation (29 cycles)
  - a dilution series of the 007 human male DNA control
  - a range of mixed DNA samples consisting of varying amounts of the 007 and 9947A human DNA controls

CRITERIA EVALUATED
- overall peak heights
- peak morphology
- presence of artifacts
- ability to detect the minor component of a mixture
- the overall success rate of each sample

First and Foremost…

PCR Cycle Number and Post-Amplification Conditions Evaluation

- Post-Amplification conditions were evaluated on the above samples using:
  - 1 µl and 2 µl of amplified product
  - 15 and 20 second injection times

CRITERIA EVALUATED
- overall peak heights
- peak morphology
- ability to detect the minor component of a mixture
- presence of artifacts
- the overall success rate of each sample
The following amounts of human DNA control 007 were tested:

- 1.5, 1.25, 1.0, 0.5, 0.1, 0.05 and 0.01 ng

Sensitivity Studies

- Each sample was amplified for 27, 28, and 29 PCR cycles, in triplicate
- 1 µl and 2 µl of amplified product were injected for 15 secs and 20 secs

Sensitivity Studies-Alleles Detected

<table>
<thead>
<tr>
<th>Sample Amount</th>
<th>27 Cycles 1 µl</th>
<th>27 Cycles 2 µl</th>
<th>28 Cycles 1 µl</th>
<th>28 Cycles 2 µl</th>
<th>29 Cycles 1 µl</th>
<th>29 Cycles 2 µl</th>
<th>Average Number of 007 Alleles Detected</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5 ng</td>
<td>43</td>
<td>43</td>
<td>43</td>
<td>43</td>
<td>43</td>
<td>43</td>
<td>43</td>
</tr>
<tr>
<td>1.25 ng</td>
<td>43</td>
<td>43</td>
<td>43</td>
<td>43</td>
<td>43</td>
<td>43</td>
<td>43</td>
</tr>
<tr>
<td>1.0 ng</td>
<td>43</td>
<td>43</td>
<td>43</td>
<td>43</td>
<td>43</td>
<td>43</td>
<td>43</td>
</tr>
<tr>
<td>0.5 ng</td>
<td>43</td>
<td>43</td>
<td>43</td>
<td>43</td>
<td>43</td>
<td>43</td>
<td>43</td>
</tr>
<tr>
<td>0.1 ng</td>
<td>10</td>
<td>17.3</td>
<td>10</td>
<td>20.3</td>
<td>34</td>
<td>39.3</td>
<td>32.3</td>
</tr>
<tr>
<td>0.05 ng</td>
<td>0.3</td>
<td>1</td>
<td>0.7</td>
<td>1.3</td>
<td>9.7</td>
<td>16.3</td>
<td>8.7</td>
</tr>
<tr>
<td>0.01 ng</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

Average Number of 007 Alleles Detected (43 unique alleles)

Goal: Maximize the data and Minimize the artifacts

PCR Cycle Number and Post-Amplification Conditions Evaluation

The human DNA controls 9947A and 007 were mixed in the following ratios:


Mixture Studies

- One (1) nanogram of each mixture was amplified for 27, 28, and 29 PCR cycles, in triplicate
- 1 µl and 2 µl of amplified product were injected for 15 secs and 20 secs on the 3500 Genetic Analyzer

Mixture Studies-Alleles Detected

<table>
<thead>
<tr>
<th>Sample Ratio</th>
<th>27 Cycles 1 µl</th>
<th>27 Cycles 2 µl</th>
<th>28 Cycles 1 µl</th>
<th>28 Cycles 2 µl</th>
<th>29 Cycles 1 µl</th>
<th>29 Cycles 2 µl</th>
<th>Average Number of 007 Alleles Detected</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>99:1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>79:1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>59:1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39:1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19:1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9:1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4:1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:79</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0:1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Number of Artifacts Detected (21 samples tested)

- One (1) nanogram of each mixture was amplified for 27, 28, and 29 PCR cycles, in triplicate
- 1 µl and 2 µl of amplified product were injected for 15 secs and 20 secs on the 3500 Genetic Analyzer

Mixture Studies-Artifacts Detected

<table>
<thead>
<tr>
<th>Sample Ratio</th>
<th>27 Cycles 1 µl</th>
<th>27 Cycles 2 µl</th>
<th>28 Cycles 1 µl</th>
<th>28 Cycles 2 µl</th>
<th>29 Cycles 1 µl</th>
<th>29 Cycles 2 µl</th>
<th>Number of Artifacts Detected</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>99:1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>79:1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>59:1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39:1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19:1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9:1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4:1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:79</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0:1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Mixture Studies-Alleles Detected**

Average Number of 007 Alleles Detected (28 unique alleles)

<table>
<thead>
<tr>
<th>Mixture</th>
<th>27 cycles</th>
<th>28 cycles</th>
<th>29 cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 : 1</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>1 : 2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>1 : 3</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>1 : 4</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
</tbody>
</table>

**Goal:**

Maximize the data and Minimize the artifacts

---

**Mixture Studies-Artifacts Detected**

Number of Artifacts Detected (57 samples tested)

<table>
<thead>
<tr>
<th>Mixture</th>
<th>27 cycles</th>
<th>28 cycles</th>
<th>29 cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 : 2</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1 : 3</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1 : 4</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

---

**Results of the PCR Cycle Number and Post-Amplification Conditions Evaluation**

- **Optimal Number of PCR cycles:** 28
- **Post-Amplification Conditions:**
  - 2 µl of amplified product with 20 second injections

28 PCR cycles with 2 µl of amplified product injected for 20 seconds maximized the data obtained while minimizing artifacts that would complicate downstream interpretations (i.e. Mixtures)

These analysis parameters were then used to assess Stochastic Effects (PHR and allele dropout) in the overall sensitivity and mixture studies

---

**Sensitivity Studies-Allele Dropout**

Peak Height of Detected Allele at a Locus with Dropout

- *analytical threshold of 178 RFUs utilized*
- 399 total heterozygous loci
- all 54 of the 399 loci all detected alleles were less than 500 RFUs
- all 24 of the 84 loci one allele dropped-out

---

**007 PHR in mixtures (007:9947A)-Stochastic Effects**

- peak height ratio > 70%
- peak height ratio < 70%
- drop-out of one allele
- drop-out of both alleles

---

**PHR Heat Map-Stochastic Effects**

- *CODIS core locus*

---

**0.0%**

**10.0%**

**20.0%**

**30.0%**

**40.0%**

**50.0%**

**60.0%**

**70.0%**

**80.0%**

**90.0%**

**100.0%**

---

**Hi-ST mainly at FGA - 11.55% cutoff**
Known and Non-Probative Evidence Samples

Previously extracted non-probative evidence samples, which mimic other sample types routinely encountered in forensic casework, were tested with the GlobalFiler STR Kit. The results obtained using GlobalFiler were compared to those previously obtained with the Identifiler STR Kit.

The sample types included in this study included:
- sexual assault samples (VS, condom, bedding, and panties)
- wearer items
- cigarette butts
- bottle/can swabs
- blood swabs
- degraded/inhibited samples
- reference samples (buccal swabs and blood cards)

GlobalFiler:
- tests 24 loci
- 21 loci are autosomal STRs
- 12 loci < 250 bp

Identifiler:
- tests 16 loci
- 15 loci are autosomal STRs
- 9 loci < 250 bp

variables:
- input volume differences (10ul vs 15ul)
- chemistry differences between the GF and ID kits
- 3500 and 3130 differences
- injection parameters
- analysis algorithms

Great concordance between GlobalFiler and Identifiler

More information from GlobalFiler due to the additional dye and the number of loci < 250 bp

GlobalFiler much more effective in overcoming inhibition and degradation

1) Semen stain from a quilt
- 0.041 ng/µl
- 0.41 ng amplified with Identifiler
- maximum injection parameters on the 3130 (2 µl/ 9 secs)
- partial male DNA profile
  (10 loci of data with 7 of the CODIS core)

- 0.62 ng amplified with GlobalFiler
- full male DNA profile (13 CODIS core)
Known and Non-Probative Evidence Samples

2) Semen stain from a pair of panties
   - 0.0091 ng/µl
   - 0.09 ng amplified with Identifiler
   - maximum injection parameters on the 3130 (2 µl/9 secs)
   - partial male DNA profile
   - 12 loci of data with 9 of the CODIS core

Known and Non-Probative Evidence Samples

3) Swabbing of a concrete block
   - 0.17 ng/µl
   - 1 ng amplified with Identifiler
   - standard injection on the 3500 (1 µl/15 secs)
   - no results
   - inhibition suspected due to extract color

The GlobalFiler™ Kits: Configuration

GlobalFiler

- 0.14 ng amplified with GlobalFiler
- partial male DNA profile
- 23 loci of data and 12 CODIS core

GlobalFiler

- 1 ng amplified with GlobalFiler
- partial DNA profile (19 loci of data; 10 CODIS core)

Mixture Studies-Male Gender Marker

- mixture containing at least 1 male
- mixture containing at least 2 males
Male Gender Marker Redundancy-Identifiler

DNA reference sample from a self-identified male

Male Gender Marker Redundancy-GlobalFiler

Reproducibility, Precision, and Accuracy

**Precision**: characterizes the amount of agreement among a series of individual (or repeated) measurements.

- **Precision** does not relate to the **accuracy** of the measurements.

**Accuracy**: describes how close the measured values are to the true value.

Reproducibility and Precision Study Design

GlobalFiler allelic ladder was prepared in one column and injected eight (8) times on the 3500 Genetic Analyzer.

- Size determinations were conducted on all alleles of the ladder for each injection.

- Standard deviations in base pair sizing were calculated for each allele in the ladder.

GlobalFiler allelic ladder – 343 alleles across 24 loci
2,744 alleles per 8-cap injection
21,952 alleles per run of 8 injections

Reproducibility and Precision of GlobalFiler Allelic Ladder

- Standard deviation was averaged across all the alleles at each locus.
Accuracy Study Design

One (1) ng of the 007 Human DNA control was amplified concurrently four times with each amplified product prepared for electrophoresis in triplicate for a total of twelve (12) samples. The twelve (12) samples were injected on three separate days generating data for thirty-six (36) samples across three separate runs.

**Accuracy**: describes how close the measured values are to the true value.

The measurement of each allele in the 007 sample was compared to the measurement of the corresponding allele in the GlobalFiler allelic ladder.

Accuracy Study Day 1

Accuracy Study Day 2

Accuracy Study Day 3

Final Thoughts

**Mixture studies**:  
- nearly complete minor component could be detected in the 1:9 and 9:1 mixture samples  
- significant minor component detected in the 1:19 and 19:1 mixture samples.

**Sensitivity and stochastic studies**:  
- nearly complete DNA profiles were obtained from 100 pg of template  
- significant data was obtained from as low as 50 pg of template.

**Final Thoughts**

**Casework samples**:  
- Great concordance between DNA profiles previously obtained with Identifiler  
- GlobalFiler very successful in overcoming both inhibition and degradation.

**Reproducibility and precision of GlobalFiler on the ABI 3500**:  
- the standard deviation across multiple allelic ladder injections was < 0.05 bp.

**Accuracy of the GlobalFiler STR Kit on the ABI 3500**:  
- all alleles detected from the 007 sample across three separate 3500 runs were within ± 0.25 bp of the allelic ladder measurement.
Acknowledgements

Jacquelyn M. Jenkins, PhD, MSFS
Forensic Scientist
Birmingham Regional Laboratory

Angelo Della Manna, MSFS
Assistant Director
Chief of Forensic Biology & DNA
A.Dellamanna@adfs.alabama.gov

Alabama Department of Forensic Sciences
NIST Studies with New Y-STR Loci & Kits

Michael D. Coble
NIST Applied Genetics Group

What has happened in the past decade…

- Selection of core Y-STR loci (SWGDAM Jan 2003)
- “Full” Y-chromosome sequence became available in June 2003; over 700 Y-STR loci identified (only ~20 in 2000)
- Commercial Y-STR kits released
  - PowerPlex Y (2000-03), Powerplex Y (8/03), Yfiler (12/04), PPY23 (6/12)
  - Yfiler Plus (coming soon)
- Many population studies performed and online databases generated with thousands of Y-STR haplotypes
- Forensic casework demonstrations showing value of Y-STR testing along with court acceptance
- Renewed interest in Y-STRs to aid familial searching

Product Disclaimer

- I will mention commercial STR kit names and information, but I am in no way attempting to endorse any specific products.

NIST Disclaimer: Certain commercial equipment, instruments and materials are identified in order to specify experimental procedures as completely as possible. In no case does such identification imply a recommendation or it imply that any of the materials, instruments or equipment identified are necessarily the best available for the purpose.

Points of view are mine and do not necessarily represent the official position of the National Institute of Standards and Technology or the U.S. Department of Justice.

Our group receives or has received funding from the FBI Laboratory and the National Institute of Justice.

STR Marker Layouts for Y-STR Kits

Yfiler Plus Kit

http://www.cstl.nist.gov/strbase/training.htm
Improved Allelic Ladder (DYS437)

- Yfiler
  - 5 Alleles
- Yfiler-Plus
  - 9 Alleles
- PPY-23
  - 9 Alleles

PowerPlex Y-23

Promega New PowerPlex® Y23 STR System Reveals More Y-STR Loci in Half the Time

PPY-23 Kit Performance

Sensitivity

Sensitivity Study – Sample C (rep 1)

0.125ng QC2 (Tube C)

Sensitivity Study – Sample C (rep 2)

0.125ng QC2 (Tube C)
Sensitivity Study – Sample D (rep 1)

0.0625ngQC2 (Tube D)

Sensitivity Study – Sample D (rep 2)

0.0625ngQC2 (Tube D)

Sensitivity Study – Sample E (rep 1)

0.0312ngQC2 (Tube E)

Sensitivity Study – Sample E (rep 2)

0.0312ngQC2 (Tube E)

Performance with related males

Relatives in the NIST Population

Samples

12 sets of 2 individuals
2 sets of 3 individuals
Adding 6 markers from PP-Y23

<table>
<thead>
<tr>
<th>Sample Name</th>
<th>DYS576</th>
<th>DYS570</th>
<th>DYS643</th>
<th>DYS533</th>
<th>DYS481</th>
<th>DYS549</th>
</tr>
</thead>
<tbody>
<tr>
<td>OT07753</td>
<td>18</td>
<td>17</td>
<td>10</td>
<td>12</td>
<td>23</td>
<td>13</td>
</tr>
<tr>
<td>UC10177</td>
<td>18</td>
<td>18</td>
<td>10</td>
<td>13</td>
<td>22</td>
<td>13</td>
</tr>
<tr>
<td>Y16</td>
<td>14</td>
<td>19</td>
<td>14</td>
<td>11</td>
<td>20</td>
<td>11</td>
</tr>
<tr>
<td>Y17</td>
<td>14</td>
<td>19</td>
<td>14</td>
<td>11</td>
<td>20</td>
<td>11</td>
</tr>
<tr>
<td>ZT79994</td>
<td>15</td>
<td>17</td>
<td>14</td>
<td>11</td>
<td>20</td>
<td>11</td>
</tr>
<tr>
<td>ZT79995</td>
<td>15</td>
<td>17</td>
<td>14</td>
<td>11</td>
<td>20</td>
<td>11</td>
</tr>
</tbody>
</table>

6 pairs still unresolved

Two Samples from the Population Plates ZT79994 and ZT79995

- Match exactly over 23 Y-STRs

<table>
<thead>
<tr>
<th>ZT79994</th>
<th>Pos</th>
<th>Seq</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>172</td>
<td>T - C</td>
</tr>
<tr>
<td>16</td>
<td>219</td>
<td>A - G</td>
</tr>
<tr>
<td>16</td>
<td>311</td>
<td>T - C</td>
</tr>
<tr>
<td>16</td>
<td>344</td>
<td>C - T</td>
</tr>
<tr>
<td>73</td>
<td>172</td>
<td>T - C</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ZT79995</th>
<th>Pos</th>
<th>Seq</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>172</td>
<td>T - C</td>
</tr>
<tr>
<td>16</td>
<td>219</td>
<td>A - G</td>
</tr>
<tr>
<td>16</td>
<td>311</td>
<td>T - C</td>
</tr>
<tr>
<td>16</td>
<td>344</td>
<td>C - T</td>
</tr>
<tr>
<td>73</td>
<td>172</td>
<td>T - C</td>
</tr>
</tbody>
</table>

Same mtDNA haplotype (U6a)

Autosomal STR Analysis

24 STR Markers

Kinship Analysis

- Combined
  - Parent-Child: 0
  - Full Sib: 56327
  - Half-Sib*: 5692
  - 1st Cousin: 264

*Is also the same stat for Aunt/Uncle/Niece/Nephew and Grandparent/Grandchild relationships

Performance with related males

Father-Son Mutation Rates

<table>
<thead>
<tr>
<th>Population</th>
<th>Meioses</th>
<th>Mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>AfAm</td>
<td>85</td>
<td>10</td>
</tr>
<tr>
<td>Asian</td>
<td>101</td>
<td>16</td>
</tr>
<tr>
<td>Caucasian</td>
<td>100</td>
<td>6</td>
</tr>
<tr>
<td>Hispanic</td>
<td>100</td>
<td>12</td>
</tr>
<tr>
<td>Sum</td>
<td>386</td>
<td>44</td>
</tr>
</tbody>
</table>
Father to Son

<table>
<thead>
<tr>
<th>+1 repeat</th>
<th>-1 repeat</th>
<th>+2 repeat</th>
<th>-2 repeat</th>
</tr>
</thead>
<tbody>
<tr>
<td>DYS439</td>
<td>2</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>DYS388H1</td>
<td>3</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>DYS481</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>DYS570</td>
<td>3</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>DYS576</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>DYS589</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

Y-GATA-H4

<table>
<thead>
<tr>
<th></th>
<th>+1</th>
<th>+1</th>
<th>+1</th>
<th>+1</th>
</tr>
</thead>
<tbody>
<tr>
<td>DYS442</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>DYS459</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYS525</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYS591</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYS603</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYS643</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYS658</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Yfiler Plus Kit

Released August 2014

Yfiler-Plus Kit Performance

Sensitivity

Sensitivity Experiment

Full reaction volumes (25 ul):
- 0.5 ng single amp
- 0.25 ng single amp
- 0.125 ng duplicate amp
- 0.0625 ng duplication amp
- 0.03125 ng duplicate amp

Half reaction volumes (12.5 ul):
- 0.5 ng duplicate amp
- 0.25 ng duplicate amp
- 0.125 ng duplicate amp
- 0.0625 ng duplication amp
- 0.03125 ng duplicate amp

http://www.cstl.nist.gov/strbase/training.htm
### Sensitivity Experiment

<table>
<thead>
<tr>
<th>Marker</th>
<th>RFU 0.0625 ng</th>
<th>RFU 0.125 ng</th>
<th>RFU 0.25 ng</th>
<th>RFU 0.5 ng</th>
</tr>
</thead>
<tbody>
<tr>
<td>DYS576</td>
<td>84</td>
<td>404</td>
<td>740</td>
<td>484</td>
</tr>
<tr>
<td>DYS389I</td>
<td>1080</td>
<td>1285</td>
<td>2442</td>
<td>1064</td>
</tr>
<tr>
<td>DYS635</td>
<td>331</td>
<td>911</td>
<td>1540</td>
<td>1950</td>
</tr>
<tr>
<td>DYS389II</td>
<td>1055</td>
<td>717</td>
<td>2561</td>
<td>1026</td>
</tr>
<tr>
<td>DYS627</td>
<td>1175</td>
<td>876</td>
<td>1340</td>
<td>2054</td>
</tr>
<tr>
<td>DYS460</td>
<td>785</td>
<td>605</td>
<td>1469</td>
<td>2961</td>
</tr>
<tr>
<td>DYS458</td>
<td>394</td>
<td>262</td>
<td>2099</td>
<td>2616</td>
</tr>
<tr>
<td>DYS19</td>
<td>835</td>
<td>351</td>
<td>1283</td>
<td>1886</td>
</tr>
<tr>
<td>GATA_H4</td>
<td>245</td>
<td>475</td>
<td>1690</td>
<td>1766</td>
</tr>
<tr>
<td>DYS448</td>
<td>594</td>
<td>1842</td>
<td>1131</td>
<td>1611</td>
</tr>
<tr>
<td>DYS391</td>
<td>199</td>
<td>398</td>
<td>1140</td>
<td>398</td>
</tr>
<tr>
<td>DYS456</td>
<td>122</td>
<td>121</td>
<td>920</td>
<td>1280</td>
</tr>
<tr>
<td>DYS390</td>
<td>214</td>
<td>673</td>
<td>771</td>
<td>1428</td>
</tr>
<tr>
<td>DYS438</td>
<td>321</td>
<td>155</td>
<td>656</td>
<td>717</td>
</tr>
<tr>
<td>DYS392</td>
<td>153</td>
<td>1306</td>
<td>446</td>
<td>1853</td>
</tr>
<tr>
<td>DYS518</td>
<td>367</td>
<td>262</td>
<td>844</td>
<td>280</td>
</tr>
<tr>
<td>DYS570</td>
<td>196</td>
<td>592</td>
<td>376</td>
<td>1165</td>
</tr>
<tr>
<td>DYS437</td>
<td>145</td>
<td>1152</td>
<td>404</td>
<td>325</td>
</tr>
<tr>
<td>DYS385</td>
<td>93</td>
<td>613</td>
<td>630</td>
<td>1672</td>
</tr>
<tr>
<td>DYS385</td>
<td>197</td>
<td>726</td>
<td>585</td>
<td>1434</td>
</tr>
<tr>
<td>DYS449</td>
<td>79</td>
<td>112</td>
<td>420</td>
<td>1727</td>
</tr>
<tr>
<td>DYS393</td>
<td>77</td>
<td>824</td>
<td>719</td>
<td>621</td>
</tr>
<tr>
<td>DYS439</td>
<td>286</td>
<td>841</td>
<td>1266</td>
<td>1449</td>
</tr>
<tr>
<td>DYS481</td>
<td>83</td>
<td>618</td>
<td>689</td>
<td>1368</td>
</tr>
<tr>
<td>DYF387S1</td>
<td>51</td>
<td>296</td>
<td>842</td>
<td>999</td>
</tr>
<tr>
<td>DYF387S1</td>
<td>448</td>
<td>499</td>
<td>1447</td>
<td>1655</td>
</tr>
<tr>
<td>DYS533</td>
<td>527</td>
<td>494</td>
<td>4033</td>
<td>2288</td>
</tr>
</tbody>
</table>

#### High Female [DNA] Experiment

- **Looking for:**
  - Artifacts
  - Inhibition

- Three male samples (B, C, F) were each combined with same female
- Input female DNA was constant at approximately 200 ng
- Input male DNA at four levels in the range of approx. 1 ng – 0.05 ng
- Male only samples for comparison
- 24 samples total

<table>
<thead>
<tr>
<th>Ratio</th>
<th>Male DNA</th>
<th>Female DNA</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>200 : 1</td>
<td>200 ng + 1 ng</td>
<td>1 ng</td>
<td>Artifacts?</td>
</tr>
<tr>
<td>700 : 1</td>
<td>200 ng + 0.3 ng</td>
<td>0.3 ng</td>
<td>Inhibition</td>
</tr>
<tr>
<td>2000 : 1</td>
<td>200 ng + 0.1 ng</td>
<td>0.1 ng</td>
<td>Artifacts?</td>
</tr>
<tr>
<td>4000 : 1</td>
<td>200 ng + 0.05 ng</td>
<td>0.05 ng</td>
<td>Artifacts?</td>
</tr>
</tbody>
</table>

#### High Level Female Experiment

- **Male Sample B**

### Yfiler-Plus Kit Performance

- **F:M Mixtures**

- **High Female Experiment**
  - Male Sample B
  - Artifacts at 4000 : 1 (approx. 200 ng female + 0.05 ng male)

- **High Level Female Experiment**
  - Male Sample B
  - No artifacts at 200 ng female + 0.05 ng male

---

http://www.cstl.nist.gov/strbase/training.htm
Discrimination Capacity

- is a measure of the number of unique haplotypes in a given population

\[
DC = \frac{\#H}{N}
\]

\# of Haplotypes

Population size

Results so far (582 individuals)

582 Males

Yfiler Plus 27 Y-STRs

580 Unique Haplotypes

1 Shared Haplotype

DC = 0.998

<table>
<thead>
<tr>
<th># haplotypes</th>
<th>PPY</th>
</tr>
</thead>
<tbody>
<tr>
<td>816</td>
<td>0.8608</td>
</tr>
<tr>
<td>930</td>
<td>0.9810</td>
</tr>
<tr>
<td>945</td>
<td>0.9968</td>
</tr>
<tr>
<td>946</td>
<td>0.9979</td>
</tr>
</tbody>
</table>

Kinship stats over 24 aSTRs

<table>
<thead>
<tr>
<th>Relatedness</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent-Child</td>
<td>0.000</td>
</tr>
<tr>
<td>Full Sib</td>
<td>0.000</td>
</tr>
<tr>
<td>Half-Sib</td>
<td>0.121</td>
</tr>
<tr>
<td>1st Cousin</td>
<td>0.814</td>
</tr>
</tbody>
</table>

Two Caucasians
Same Haplotype at Yfiler Plus Loci

http://www.cstl.nist.gov/strbase/training.htm
<table>
<thead>
<tr>
<th>N = 948 males</th>
<th>Yfiler</th>
<th>New Loci*</th>
<th>Yfiler Plus*</th>
</tr>
</thead>
<tbody>
<tr>
<td># haplotypes</td>
<td>930</td>
<td>945</td>
<td>946</td>
</tr>
<tr>
<td>discrimination capacity</td>
<td>0.9810</td>
<td>0.9842</td>
<td>0.9979</td>
</tr>
<tr>
<td># times haplotype observed</td>
<td>Yfiler (17 loci)</td>
<td>New Loci* (10 loci)</td>
<td>Yfiler Plus* (27 loci)</td>
</tr>
<tr>
<td>1</td>
<td>916</td>
<td>918</td>
<td>944</td>
</tr>
<tr>
<td>2</td>
<td>11</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>14</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>16</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>17</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>18</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>19</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>20</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

The new loci alone perform slightly better than Yfiler

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)
- Not as informative as autosomal STR results
  - More like addition (10 + 10 + 10 = 30) than multiplication (10 x 10 x 10 = 1,000)
- 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

Rapidly Mutating (RM) Y-STRs

Trying to separate close male relatives

Using Y-STRs with a higher mutation rate, father-son and brother pairs can sometimes be distinguished

A new frontier of forensic Y-chromosome analysis: Rapidly mutating Y-STRs for differentiating male relatives and paternal lineages

Manfred Kayser


Rapidly Mutating Y-STRs

Mutability of Y-Chromosomal Microsatellites: Rates, Characteristics, Molecular Basis, and Forensic Implications

Manfred Kayser


http://www.cstl.nist.gov/strbase/training.htm
Why do these markers mutate “rapidly”?

<table>
<thead>
<tr>
<th>Locus</th>
<th>(average mutation rate)</th>
<th>&quot;Large&quot; number of repeats</th>
</tr>
</thead>
<tbody>
<tr>
<td>DYS449</td>
<td>(1.2%)</td>
<td></td>
</tr>
<tr>
<td>DYS518</td>
<td>(1.8%)</td>
<td></td>
</tr>
<tr>
<td>DYS547</td>
<td>(2.4%)</td>
<td></td>
</tr>
<tr>
<td>DYS570</td>
<td>(1.2%)</td>
<td></td>
</tr>
<tr>
<td>DYS576</td>
<td>(1.4%)</td>
<td></td>
</tr>
<tr>
<td>DYS612</td>
<td>(1.4%)</td>
<td></td>
</tr>
<tr>
<td>DYS626</td>
<td>(1.2%)</td>
<td></td>
</tr>
<tr>
<td>DYS627</td>
<td>(1.2%)</td>
<td></td>
</tr>
<tr>
<td>DYF387S1</td>
<td>(1.6%)</td>
<td></td>
</tr>
<tr>
<td>DYF399S1</td>
<td>(7.7%)</td>
<td></td>
</tr>
<tr>
<td>DYF403S1</td>
<td>(3.1/1.2%)</td>
<td></td>
</tr>
<tr>
<td>DYF404S1</td>
<td>(1.3%)</td>
<td></td>
</tr>
<tr>
<td>DYF526 a/b</td>
<td>(1.3%)</td>
<td></td>
</tr>
</tbody>
</table>

Markers in Yfiler Plus

DYF458 (0.64%) is highest in Yfiler loci where average is ~0.2%

Gene Diversity

- is a measure of the uniqueness of a particular marker in a given population

\[
GD = \left(1 - \sum x_i^2\right)
\]

Relative frequency of each allele

Gene Diversity of the YFP Markers

<table>
<thead>
<tr>
<th>Locus</th>
<th>Gene Diversity</th>
</tr>
</thead>
<tbody>
<tr>
<td>DYF387S1a/b</td>
<td>0.919</td>
</tr>
<tr>
<td>DYF388a/b</td>
<td>0.919</td>
</tr>
<tr>
<td>DYF449</td>
<td>0.8584</td>
</tr>
<tr>
<td>DYF461</td>
<td>0.82</td>
</tr>
<tr>
<td>DYF518</td>
<td>0.8196</td>
</tr>
<tr>
<td>DYF576</td>
<td>0.7954</td>
</tr>
<tr>
<td>DYF570</td>
<td>0.7852</td>
</tr>
<tr>
<td>DYF458</td>
<td>0.7671</td>
</tr>
<tr>
<td>DYF530</td>
<td>0.7645</td>
</tr>
<tr>
<td>DYF635</td>
<td>0.7457</td>
</tr>
<tr>
<td>DYF499I</td>
<td>0.7335</td>
</tr>
<tr>
<td>DYF546</td>
<td>0.7002</td>
</tr>
<tr>
<td>DYF456</td>
<td>0.7015</td>
</tr>
<tr>
<td>DYF548</td>
<td>0.695</td>
</tr>
<tr>
<td>DYF19</td>
<td>0.6681</td>
</tr>
<tr>
<td>DYF439</td>
<td>0.6553</td>
</tr>
<tr>
<td>DYF531</td>
<td>0.6372</td>
</tr>
<tr>
<td>DYF437</td>
<td>0.6305</td>
</tr>
<tr>
<td>GATA_H4</td>
<td>0.6026</td>
</tr>
<tr>
<td>DYF502</td>
<td>0.6001</td>
</tr>
<tr>
<td>DYF490</td>
<td>0.5736</td>
</tr>
<tr>
<td>DYF389I</td>
<td>0.548</td>
</tr>
<tr>
<td>DYF391</td>
<td>0.5352</td>
</tr>
<tr>
<td>DYF393</td>
<td>0.4749</td>
</tr>
</tbody>
</table>

N = 100

\[GD = \left(1 - \sum x_i^2\right)\]

Marker Y

1 type = 100%

\[GD = 0\]

\[DC = 1/100 = 0.01\]

\[N = 100\]

\[GD = \left(1 - \sum x_i^2\right)\]

Marker Y

4 types = 25%

\[GD = 0.75\]

\[DC = 4/100 = 0.04\]

\[N = 100\]

\[GD = \left(1 - \sum x_i^2\right)\]

Marker Y

100 types = 0%

\[GD = 0.99\]

\[DC = 100/100 = 1.0\]

http://www.cstl.nist.gov/strbase/training.htm
Interpretational Issues

- We will need to move away from simply “excluding” based upon a number of discordant markers.
- A Likelihood Ratio can provide weight to the evidence based upon competing propositions.
- This will require information on the haplotype frequency and mutation rate data.

Relating two deep-rooted pedigrees from Central Germany by high-resolution Y-STR haplotyping
Manfred Kayser<sup>*</sup>, Mark Vermeulen<sup>**</sup>, Hans Knaubt<sup>**</sup>, Herbert Schurer<sup>**</sup>, Michael Krawczak<sup>**</sup>, Latc Roesler<sup>***</sup>

Mutation Rate Information

<table>
<thead>
<tr>
<th>Meioses</th>
<th>Mutations</th>
<th>Group</th>
<th>Marker</th>
<th># of Mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>63</td>
<td>15</td>
<td>AfAm</td>
<td>DYF399S1</td>
<td>15</td>
</tr>
<tr>
<td>89</td>
<td>25</td>
<td>Asian</td>
<td>DYF403S1a/b</td>
<td>11</td>
</tr>
<tr>
<td>91</td>
<td>11</td>
<td>Caucasian</td>
<td>DYF5627</td>
<td>7</td>
</tr>
<tr>
<td>88</td>
<td>20</td>
<td>Hispanic</td>
<td>DY5612</td>
<td>7</td>
</tr>
<tr>
<td>331</td>
<td>71</td>
<td>total</td>
<td>DY5518</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DY5570</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DY5626</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DY5547</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DY5526a/b</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DY5576</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DY5449</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DYF404S1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DYF387S1</td>
<td>1</td>
</tr>
</tbody>
</table>

Summary

- Rapidly Mutating Y-STRs are highly diverse markers that can discriminate common haplotypes and close relatives.
- These markers may create interpretational issues for paternity/missing persons cases, but LRs can be useful for evaluating these situations.
- The Yfiler Plus kit is sensitive and provides improved haplotype discrimination.

Acknowledgments

Promega Corporation
Thermo Fisher Life Technologies
Jonathan Tabak, Andrea Carbonaro, Ariana Wheaton

Contact Info:
mcoble@nist.gov
301-975-4330

NIST Disclaimer: Certain commercial equipment, instruments and materials are identified in order to specify experimental procedures as completely as possible. In no case does such identification imply that any of the materials, instruments or equipment identified are necessarily the best available for the purpose.

Points of view are those of the presenters and do not necessarily represent the official position of the National Institute of Standards and Technology or the U.S. Department of Justice.

http://www.cstl.nist.gov/strbase/training.htm
STRBase Resources & Additional Information

John M. Butler
NIST Office of Special Programs

A Brief History of the STRBase Website

- Initial information was collected on STR markers while working on my PhD dissertation in 1993-1995
- Started a review article in 1996 while a NIST postdoc but wanted to create a dynamic rather than an out-of-date resource
- Created hundreds of individual web pages that were hyperlinked together
- Website launched in July 1997 (discussed at ISHI 1997)
- Became a NIST Standard Reference Database (SRD 130) because of its high visibility
- I continue to update the website (via an HTML editor)...
- I have more information than I have had time to upload (i.e., there is additional information in development)

Core STRBase Information

- STR101: Brief Introduction to STRs
- Core Loci: FRJ/CODIS Core STR Loci and European Core Loci
- STR Fact Sheets (observed alleles and PCR product sizes)
- Multiplex STR kits
- Sequence Information (associated)
- Variant Allele Reports
- Tri-Allelic Patterns
- Mutation Rates for Common Loci
- Published PCR primers
- Y-chromosome STRs
- Low-Template DNA Information
- Marker Interpreters
- Kaplan Analyses
- miniSTRs (short amplicons)
- Null Alleles - discordance observed between STR kits
- STR Reference List - now 3657 references

Multiplex STR Kit Information

http://www.cstl.nist.gov/strbase/multiplex.htm

http://www.cstl.nist.gov/strbase/training.htm
We collect contributions from all over the world where unusual results have been observed with STR data.

Enables laboratories to check if others have seen a specific variant allele or tri-allelic pattern.

Currently (as of Jul 29, 2014 update)

- 714 variants at 43 loci
- 353 tri-allelic patterns at 36 loci

Information on Variant Alleles

- We collect contributions from all over the world where unusual results have been observed with STR data.
- Enables laboratories to check if others have seen a specific variant allele or tri-allelic pattern.
- Currently (as of July 29, 2014 update)

714 variants at 43 loci
353 tri-allelic patterns at 36 loci

From D25S1336 Variants Table (http://www.cstl.nist.gov/strbase/var_D25S1336.htm)
http://www.cstl.nist.gov/strbase/var_tab.htm
We have analyzed 112 D12S391 alleles (126 total) reported so far in STRBase (data provided based on 123 NGM Select, 1 EST16, 1 NGM, and 1 PP21).

- We have analyzed 1036 unrelated samples with 29 autosomal STRs and 23 Y-STRs (all current STR and Y-STR kit loci).
- Becky Hill and Mike Coble have described this data set in their presentations.

From NIST 1036 data set

<table>
<thead>
<tr>
<th>Allele</th>
<th>#</th>
<th>AFAm</th>
<th>AfAm</th>
<th>Ave</th>
<th>Ghee</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>5</td>
<td>0.5</td>
<td>0.1</td>
<td>0.1</td>
<td>0.02</td>
<td>0.40%</td>
</tr>
<tr>
<td>15</td>
<td>105</td>
<td>5.1</td>
<td>7.7</td>
<td>4.1</td>
<td>3.2</td>
<td>4.4</td>
</tr>
<tr>
<td>16</td>
<td>848</td>
<td>4.1</td>
<td>6.7</td>
<td>1.0</td>
<td>2.2</td>
<td>4.2</td>
</tr>
<tr>
<td>17</td>
<td>2568</td>
<td>12.5</td>
<td>16.8</td>
<td>9.2</td>
<td>12.7</td>
<td>17.4</td>
</tr>
<tr>
<td>18</td>
<td>432</td>
<td>20.8</td>
<td>25.3</td>
<td>26.3</td>
<td>17.2</td>
<td>17.4</td>
</tr>
<tr>
<td>19</td>
<td>100</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>20</td>
<td>262</td>
<td>12.6</td>
<td>10.4</td>
<td>19.6</td>
<td>11.1</td>
<td>15.5</td>
</tr>
<tr>
<td>21</td>
<td>250</td>
<td>2.0</td>
<td>0.1</td>
<td>0.3</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>22</td>
<td>269</td>
<td>10.1</td>
<td>6.4</td>
<td>9.8</td>
<td>12.9</td>
<td>18.0</td>
</tr>
<tr>
<td>23</td>
<td>141</td>
<td>7.0</td>
<td>0.4</td>
<td>0.9</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td>24</td>
<td>102</td>
<td>4.9</td>
<td>2.9</td>
<td>2.6</td>
<td>6.9</td>
<td>5.7</td>
</tr>
<tr>
<td>25</td>
<td>253</td>
<td>2.6</td>
<td>1.3</td>
<td>1.0</td>
<td>4.7</td>
<td>1.7</td>
</tr>
<tr>
<td>26</td>
<td>104</td>
<td>1.0</td>
<td>0.0</td>
<td>0.5</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>27</td>
<td>24</td>
<td>1.2</td>
<td>0.9</td>
<td>1.5</td>
<td>1.7</td>
<td>0.6</td>
</tr>
<tr>
<td>28</td>
<td>7</td>
<td>0.3</td>
<td>1.0</td>
<td>0.3</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>29</td>
<td>6</td>
<td>0.3</td>
<td>0.5</td>
<td>0.1</td>
<td>0.4</td>
<td>0.4</td>
</tr>
</tbody>
</table>

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


Theoretical heterozygotes (2pq)
2 x 0.013 x 0.013 = 0.0002696 (0.0027%)
2 x 0.013 x 0.013 x 0.013 = 0.000000349 (0.00000035%)

Observed heterozygotes with a single nucleotide difference
9 out of 1036 = 0.87%

17, 17.1
17.3, 18 (3x)
18, 18.1
18.3, 19 (2x)
19, 19.1
19.3, 20

http://www.cstl.nist.gov/strbase/training.htm
Stutter Observations vs Stutter Filters

Data from GlobalFiler Kit User Manual, p. 79

Recent NIST Publications on “New” STR Loci


Additional Information Needed/Planned

• **Mutation rate information** to aid kinship analysis
  – More father/son studies are needed with D12S391, D1S1656, D2S441, D10S1248, and D22S1045

• A complete summary of **flanking region variation** and null alleles produced from primer binding site mutations

• Future plans for STRBase: listing of **full sequences for detected STR alleles** (repeats and flanking regions) to aid next-generation sequencing efforts
  – Will enable nomenclature and classification of sub-allele variation for STR markers

Summary

• The U.S. forensic DNA community will soon expand to additional STR loci and new kits are now available to help with this effort

• New information is being developed at NIST, published, and added to STRBase to aid understanding of these additional STR loci

• STRBase can be a model for other forensic disciplines in sharing information with the forensic science community

Acknowledgments

$ National Institute of Justice, FBI, and NIST
Becky Hill and Mike Coble (NIST Applied Genetics)
Jeff Nye (Michigan State Police)
Hope Olson (North Dakota Office of Attorney General)
Jason Kokoszka & Angelo Della Manna (Alabama DFS)

Contact info:
john.butler@nist.gov
+1-301-975-4049

Final version of this presentation available at:
http://www.cstl.nist.gov/strbase/NISTpub.htm

http://www.cstl.nist.gov/strbase/training.htm
Autosomal STR Information

European and U.S. Core Loci Expansion Efforts


Hares, D.R. (2012b). Addendum to expanding the CODIS core loci in the United States. *Forensic Science International: Genetics, 6*, e135


Information on STR Kits or New Assays


Muller, K., et al. (2010). Casework testing of the multiplex kits AmpFISTR SEfiler Plus PCR amplification kit (AB), PowerPlex S5 System (Promega) and AmpFISTR MiniFiler PCR amplification kit (AB). *Forensic Science International: Genetics*, 4, 200-205.


**Population Data on NIST U.S. Samples**


**On-line Population Databases**

AllST*R Autosomal Database for Short Tandem Repeats: http://allstr.de/


ENFSI DNA WG STR Population Database: http://strbase.org/


popSTR: http://spsmart.cesga.es/popstr.php


**Population Data on New Autosomal STRs**


**Information on STR Loci**


**Concordance Studies**


Li, F., et al. (2014). Identification of new primer binding site mutations at TH01 and D13S317 loci and determination of their corresponding STR alleles by allele-specific PCR. *Forensic Science International: Genetics, 8*, 143-146.


Next-Generation Sequencing of STR Alleles


Amelogenin Sex Typing and Anomalies Observed


Closely-Spaced Loci: D12S391 and vWA Linkage Disequilibrium Studies


**Previous Concerns over Potential Disease Linkage (either raised or addressed)**


SWGDAM (2012), SWGDAM Executive Board Considerations for Claims that the CODIS Core Loci are ‘Associated’ with Medical Conditions/Diseases. [http://www.swgdam.org/SWGDAM_State_v_Abernathy.pdf](http://www.swgdam.org/SWGDAM_State_v_Abernathy.pdf)

Y-STR Information

SWGDAM Interpretation Guidelines for Y-Chromosome STR Typing by Forensic DNA Laboratories: http://swgdam.org/SWGDAM_YSTR_Guidelines_APPROVED_01092014_v_02112014_FINAL.pdf

Y-STR Haplotype Databases


U.S. Y-STR Database: http://www.usystrdatabase.org/


Y-Chromosome Haplotype Reference Database (YHRD): http://www.yhrd.org
(Version 4.0, August 2014): 192,553 minimal haplotypes; 90,792 PPY; 78,530 Yfiler; 21,909 PPY23; 873 Yfiler Plus

Yfiler Haplotype Database: http://www6.appliedbiosystems.com/yfilerdatabase/
(August 29, 2014): 11,393 Yfiler haplotypes

Ysearch (Genetic Genealogy): http://www.ysearch.org/
(August 29, 2014): 141,947 records including 108,402 unique haplotypes and 95,231 surnames

Y-STR Kits


PowerPlex Y23 Population Data

Calderon, S., et al. (2013). Phylogenetic and forensic studies of the Southeast Florida Hispanic population using the next-generation forensic PowerPlex® Y23 STR marker system. Legal Medicine, 15(6), 289-292.


Rapidly Mutating (RM) Y-STRs


Early Y-STR Work at NIST


Impact of Additional Y-STR Loci


**Y-STR Mutations**


**Y-STR Profile Anomalies**


