16 Going on 24: GlobalFiler and the New STR Loci

Becky Hill
National Institute of Standards and Technology

Mid-Atlantic Association of Forensic Scientists
Annual Conference
State College, PA
May 20, 2014
GlobalFiler STR Kit

Launched Friday, September 14, 2012

Introducing the world’s most powerful STR kit

Around the world, forensic labs are being asked to do more with less. That’s why the new GlobalFiler™ STR Kit combines reduced amplification time with maximum data recovery power. As part of the only fully integrated and validated forensic workflow, this breakthrough 6-dye, 24-loci technology is designed to deliver unprecedented lab performance. And, it’s backed by Life Technologies best-in-class training, service, and support.
• 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

Two separate 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>TGCC/TTCC</td>
<td>10 to 31</td>
<td>40</td>
</tr>
<tr>
<td>D19S433</td>
<td>19q12</td>
<td>AAGG/TAGG</td>
<td>5.2 to 20</td>
<td>36</td>
</tr>
<tr>
<td>Penta D</td>
<td>21q22.3</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>AGAT/AGAC</td>
<td>13 to 27.2</td>
<td>52</td>
</tr>
<tr>
<td>D2S441</td>
<td>2p14</td>
<td>TCTA/TCAA</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>AAAG‡</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 (2011) Advanced Topics in Forensic DNA Typing: Methodology; ‡SE33 alleles have complex repeat structure
Loci sorted on Probability of Identity ($P_I$) values

<table>
<thead>
<tr>
<th>Locus</th>
<th>Alleles Observed</th>
<th>Genotypes Observed</th>
<th>Het (obs)</th>
<th>$P_I$ Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE33</td>
<td>52</td>
<td>304</td>
<td>0.9353</td>
<td>0.0066</td>
</tr>
<tr>
<td>Penta E</td>
<td>23</td>
<td>138</td>
<td>0.8996</td>
<td>0.0147</td>
</tr>
<tr>
<td>D2S1338</td>
<td>13</td>
<td>68</td>
<td>0.8793</td>
<td>0.0220</td>
</tr>
<tr>
<td>D1S1656</td>
<td>15</td>
<td>93</td>
<td>0.8890</td>
<td>0.0224</td>
</tr>
<tr>
<td>D18S51</td>
<td>22</td>
<td>93</td>
<td>0.8687</td>
<td>0.0258</td>
</tr>
<tr>
<td>D12S391</td>
<td>24</td>
<td>113</td>
<td>0.8813</td>
<td>0.0271</td>
</tr>
<tr>
<td>FGA</td>
<td>27</td>
<td>96</td>
<td>0.8745</td>
<td>0.0308</td>
</tr>
<tr>
<td>D6S1043</td>
<td>27</td>
<td>109</td>
<td>0.8494</td>
<td>0.0321</td>
</tr>
<tr>
<td>Penta D</td>
<td>16</td>
<td>74</td>
<td>0.8552</td>
<td>0.0382</td>
</tr>
<tr>
<td>D21S11</td>
<td>27</td>
<td>86</td>
<td>0.8330</td>
<td>0.0403</td>
</tr>
<tr>
<td>D8S1179</td>
<td>11</td>
<td>46</td>
<td>0.7992</td>
<td>0.0558</td>
</tr>
<tr>
<td>D19S433</td>
<td>16</td>
<td>78</td>
<td>0.8118</td>
<td>0.0559</td>
</tr>
<tr>
<td>vWA</td>
<td>11</td>
<td>39</td>
<td>0.8060</td>
<td>0.0611</td>
</tr>
<tr>
<td>F13A01</td>
<td>16</td>
<td>56</td>
<td>0.7809</td>
<td>0.0678</td>
</tr>
<tr>
<td>D7S820</td>
<td>11</td>
<td>32</td>
<td>0.7944</td>
<td>0.0726</td>
</tr>
<tr>
<td>D16S539</td>
<td>9</td>
<td>28</td>
<td>0.7761</td>
<td>0.0749</td>
</tr>
<tr>
<td>D13S317</td>
<td>8</td>
<td>29</td>
<td>0.7674</td>
<td>0.0765</td>
</tr>
<tr>
<td>TH01</td>
<td>8</td>
<td>24</td>
<td>0.7471</td>
<td>0.0766</td>
</tr>
<tr>
<td>Penta C</td>
<td>12</td>
<td>49</td>
<td>0.7732</td>
<td>0.0769</td>
</tr>
<tr>
<td>D2S441</td>
<td>15</td>
<td>43</td>
<td>0.7828</td>
<td>0.0841</td>
</tr>
<tr>
<td>D10S1248</td>
<td>12</td>
<td>39</td>
<td>0.7819</td>
<td>0.0845</td>
</tr>
<tr>
<td>D3S1358</td>
<td>11</td>
<td>30</td>
<td>0.7519</td>
<td>0.0915</td>
</tr>
<tr>
<td>D22S1045</td>
<td>11</td>
<td>44</td>
<td>0.7606</td>
<td>0.0921</td>
</tr>
<tr>
<td>F13B</td>
<td>7</td>
<td>20</td>
<td>0.6911</td>
<td>0.0973</td>
</tr>
<tr>
<td>CSF1PO</td>
<td>9</td>
<td>31</td>
<td>0.7558</td>
<td>0.1054</td>
</tr>
<tr>
<td>D5S818</td>
<td>9</td>
<td>34</td>
<td>0.7297</td>
<td>0.1104</td>
</tr>
<tr>
<td>FESFPS</td>
<td>12</td>
<td>36</td>
<td>0.7230</td>
<td>0.1128</td>
</tr>
<tr>
<td>LPL</td>
<td>9</td>
<td>27</td>
<td>0.7027</td>
<td>0.1336</td>
</tr>
<tr>
<td>TPOX</td>
<td>9</td>
<td>28</td>
<td>0.6902</td>
<td>0.1358</td>
</tr>
</tbody>
</table>

29 STR Loci present in STR kits rank ordered by their variability

N=1036 (only unrelated samples used)

There are several loci more polymorphic than the current CODIS 13 STRs

- Better for mixtures (more alleles seen)
- Better for kinship (low mutation rate)

361 Caucasians
342 African Americans
236 Hispanics
97 Asians
GlobalFiler: What has been done at NIST?

- We received a prototype GlobalFiler Express (direct amp) kit in July 2012
- All requested “VTS” experiments performed
  - Sensitivity, Reproducibility, and Performance
- 50 NIST bloodstains on FTA
- 50 NIST bloodstains on 903 (same DNA samples)
  - Prep-n-Go treatment
- 2 buccal samples collected with Whatman EasiCollect
- Degraded DNA and concordance studies
Samples Tested

**Bloodstains on FTA paper**

- Whatman EasiCollect
- Worked well

**Buccal swab on FTA paper**

- Positive control (AB001)
- Negative control
- Paper punches with no sample applied
- No result
  - (not enough cells collected?)

**Automated Harris puncher used**
Samples Following Direct PCR

Bloodstain punch samples

Buccal punch samples

NTC

AB001

007

Liquid DNA sample
The J6 6-Dye Spectral Calibration Worked Well
Cycle Number Experiments

*Selected 26 cycles for all additional studies*
Baseline Comparison in Negative Control

25 cycles

26 cycles

27 cycles
GlobalFiler Allelic Ladder

343 alleles across these 24 loci
Positive Control 007

1 punch, 26 cycles
Successful buccal swab (Whatman collector)
Negative Control
Degraded DNA Study

- Tested DNA fragments < 250 bp with GlobalFiler Casework
- GlobalFiler: 10 loci dropped out (>250 bp)
- MiniFiler: Full profile, 7 loci gained (D21S11 and Amel overlap)
- Total of 21 loci included in the profile with combined kits (GF & MF)
- Only 3 loci did not give a signal: TPOX, DYS391 and SE33
Concordance Evaluation

- **50 NIST bloodstain sample results were 99.8% concordant** with previous Identifiler Plus, PowerPlex 16 HS, PowerPlex 18D, and PowerPlex ESI 17 Pro results

- **Two loci out of 1100 comparisons** (22 loci x 50 samples) were discordant
  - A large **D8S1179 “24” allele** produced a false D21S11 tri-allele in bloodstain #26 (ID/NGM: 25, 29, 30 and GF: 26, 29, 30)
  - **D22S1045 allele 15 dropout/severe imbalance** in one sample (bloodstain #29)
A large D8S1179 allele runs into the adjacent D21S11 locus
Allele 12: [TCTA]_2 TCTG [TCTA]_9

Allele “24”: [TCTA]_2 TCTG [TCTA]_9 duplication of the 48 bases
10 bases downstream of the repeat
Result with This Large D8S1179 Allele
Using European STR Kits

NGM SElect

False D21S11 tri-allele

PP ESX 17

False FGA tri-allele

PP ESI 17

No extra peak

Reverse primer internal to duplicated flanking region

Previously presented at CODIS Conference (November 2011)
Likely Primer Binding Site Mutation at D22S1045

This sample has not yet been sequenced to discover the underlying sequence around the flanking region.
SRM 2391c Concordance

• All SRM 2391c components run with GlobalFiler Casework were **concordant** at all loci
  – Exception: Y indel was not included in the comparison because no other kits use this marker

• Component D at D12S391 shows 1 bp resolution:
GlobalFiler Express Results

• **Sensitivity Study** (to optimize cycle number)
  – 26 cycles is optimal number in our lab

• **Reproducibility Study**
  – All replicates at 26 cycles gave similar peak height values and ratios

• **Performance Study**
  – All 50 blood samples produced full, well-balanced profiles

• **Concordance Study**
  – Only 2 discordant results observed with 1100 allele comparisons (99.8% concordance)

• **Degraded DNA Study (Casework kit)**
  – Expected results observed with each fragment size
Rapid DNA Typing: How Fast is Erica Butts?

Experimental Design

**Steps Involved**
- Collection
- Extraction
- Quantitation
- Multiplex PCR
- STR Typing
- Interpretation of Results

**GlobalFiler Express**
Single Source Reference Samples

- None
  - (1.2 mm Blood Punch)
- GlobalFiler Express
  - (9700)
- ABI 3500 Genetic Analyzer
  - GeneMapper IDX v1.2

8 unique samples were typed in parallel

Testing was timed from collection through data interpretation of results, to include all sample transfer steps
Total time from swab to answer: 1 hour 25 min
Includes set up times

1.2 mm blood punch
GlobalFiler STR kit
FBI Consortium Validation Project

• The CODIS database will potentially increase from 13 to 20 core loci
• Two larger STR multiplex kits were developed to meet these requirements
  – GlobalFiler and PowerPlex Fusion
• Purpose of the CVP
  – To have several crime labs across the U.S. validate these two new kits
  – To determine if these new loci are appropriate to add to the CODIS core set
  – To implement the new loci into the U.S. National DNA Database
Implementation Plan

Plan Published in February 2012


• Selection of laboratories to participate in validation studies
• **Validation of proposed new CODIS core loci**
• Selection of CODIS core loci
• Implementation of new CODIS core loci into NDIS operations
Validation Plan

- CODIS Core Loci Group and NIST are currently working with U.S. laboratories participating in the consortium validation effort
  - Manufacturer grade kits were provided for validation effort
  - Specific plans distributed to Database, Casework and Missing Person Laboratories
  - 11 laboratories involved in the consortium effort
  - Data is currently being collected and evaluated from both GlobalFiler and PowerPlex Fusion kits
Summary

• GlobalFiler is a 6 dye STR multiplex kit that includes the required new loci for upload to CODIS

• GlobalFiler performed well in our initial studies, including concordance, degraded DNA, sensitivity, reproducibility, and performance studies

• GlobalFiler results can be generated in less than 2 hours
  – Overall time includes: collection, sample handling, and liquid transfer steps

• GlobalFiler is currently being validated as part of the FBI Consortium Validation Project and results from these studies will eventually be published and presented
Thank you for your attention!

Our team publications and presentations are available at:
http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm

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

Questions?
becky.hill@nist.gov
301-975-4275

Acknowledgements:
Erica Butts for GF rapid work

Funding from the National Institute of Justice (NIJ) through NIST Office of Law Enforcement Standards