The Value of Additional STR Markers
Mike Coble and Becky Hill
NIST Applied Genetics Group
National Institute of Standards and Technology
Gaithersburg, Maryland

Expanding CODIS Core Loci

Additional STR Loci in the Future?

- Will be needed for more complex kinship analyses and extended applications
  - Example: Y-STRs needed for familial searching

- Immigration testing needs more than 13 STRs

- Larger DNA databases will require more loci

The Copenhagen Forensic Genetic Summer School

June 27-28, 2012

### Proposed Expanded CODIS Core Loci

<table>
<thead>
<tr>
<th>Locus</th>
<th>Description</th>
<th>Locus</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>D3S1358</td>
<td>Required (in order of preference)</td>
<td>DYS389I</td>
<td>Required</td>
</tr>
<tr>
<td>D5S818</td>
<td>Required (in order of preference)</td>
<td>DYS391</td>
<td>Current CODIS 13 loci in red font</td>
</tr>
<tr>
<td>CSF1PO</td>
<td>Required (in order of preference)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D2S1338</td>
<td>Required (in order of preference)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D19S433</td>
<td>Required (in order of preference)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D1S1656</td>
<td>Required (in order of preference)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D12S391</td>
<td>Required (in order of preference)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D2S441</td>
<td>Required (in order of preference)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D10S1248</td>
<td>Required (in order of preference)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PentaE</td>
<td>Required (in order of preference)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYS391</td>
<td>Required (in order of preference)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Criteria for Acceptance of Additional Loci

- **STR Loci**
  - No known association to medical conditions or defects
  - Low mutation rate
  - High level of independence
  - High level of discrimination
  - Used by international forensic DNA community
  - Number of loci vs. discrimination factor
  - Compliance with Quality Assurance Standards (QAS)

- **Kit performance**
  - Balance between loci
  - Reliable
  - Reproducible
  - Sensitive
  - Quality results
  - Adaptable for use by NDIS laboratories (number of amplifications, ability of kit manufacturers to produce)
  - QAS compliant (documentation and availability of validation requirements)

---

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

Adventitious Matches

- The only published account of a false match from a DNA database came in 1999 when the UK database then consisting of 660,000 profiles with only 6 STR loci (SGM assay) lead to a “hit” between two individuals whose 6-locus random match probability was 1 in 37 million (R. Willing, USA Today, Feb 8, 2000, “Mismatch calls DNA test into question”).
- Further testing with four additional STRs (SGM Plus loci) showed that the samples were from different individuals. The UK expanded the number of core loci from 6 to 10 with the adoption of the SGM Plus kit to try and prevent another adventitious match.
- The growth of DNA databases necessitates the inclusion of additional loci to avoid this problem.

International Comparability

| US       | Europe   | ESS = European Standard Set
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>TPOX</td>
<td>FGA</td>
<td>7 ESS loci</td>
</tr>
<tr>
<td>CSF1PO</td>
<td>VWA</td>
<td></td>
</tr>
<tr>
<td>D5S818</td>
<td>D5S819</td>
<td></td>
</tr>
<tr>
<td>D7S820</td>
<td>D5S820</td>
<td></td>
</tr>
<tr>
<td>D13S317</td>
<td>D21S11</td>
<td></td>
</tr>
<tr>
<td>FGA</td>
<td>TH01</td>
<td></td>
</tr>
<tr>
<td>VWA</td>
<td>D16S539</td>
<td></td>
</tr>
<tr>
<td>D6S1043</td>
<td>D10S1248</td>
<td></td>
</tr>
<tr>
<td>TH01</td>
<td>D12S391</td>
<td></td>
</tr>
<tr>
<td>D16S539</td>
<td>D19S433</td>
<td></td>
</tr>
<tr>
<td>D2S1338</td>
<td>D18S51</td>
<td></td>
</tr>
<tr>
<td>D8S1179</td>
<td>D21S11</td>
<td></td>
</tr>
<tr>
<td>TH01</td>
<td>D19S433</td>
<td></td>
</tr>
<tr>
<td>D16S539</td>
<td>D21S11</td>
<td></td>
</tr>
<tr>
<td>D2S1338</td>
<td>D19S433</td>
<td></td>
</tr>
<tr>
<td>D12S391</td>
<td>D18S51</td>
<td></td>
</tr>
<tr>
<td>D1S1656</td>
<td>D22S1045</td>
<td></td>
</tr>
<tr>
<td>D22S1045</td>
<td>SE33</td>
<td></td>
</tr>
<tr>
<td>D22S1045</td>
<td>ESS3</td>
<td></td>
</tr>
</tbody>
</table>

Currently there are 24 autosomal STR markers present in commercial kits

3 miniSTR loci developed at NIST

Locus used in China

D6S1043
Common Forensic STR Loci

European Standard Set: D16S539, D2S1338, D19S433, SE33

U.S. CODIS Core Loci: D2S1338, D19S433, Penta D, Penta E

What can we learn from European Standard Set (ESS) expansion experience?

European Expansion Efforts

More loci added as databases grew...

With expanded loci selections, focus is on casework capabilities not familial searching potential
Lessons from European ESS Expansion

- **Data studies should drive decisions**
  - Interlaboratory study with degraded DNA (Dixon et al. 2006 article was key)
  - Casework capabilities are a primary goal
  - miniSTRs and desire for kits with ability to overcome inhibitors
  - Initial locus selection announced through Letters to the Editor of the leading forensic DNA journal (Gill et al. 2006a, 2006b)
  - Companies responded with prototype kits for evaluation
  - Expanded ESS loci were selected and voted upon after data review by ENFSI labs (4 years after initial recommendations were made)
  - EU adopted recommendations of ENFSI
  - Commercial kits became available to meet expanded ESS requirements
  - Population data gathered and software developed
  - European labs must be compliant by Nov 30, 2011 (2 years after adoption)
  - Casework capabilities not familial searching potential were the intent of the core loci selection

EDNAP Study Showed Value of miniSTRs

“Recently, there has been much debate about what kinds of genetic markers should be implemented as new core loci that constitute national DNA databases. The choices lie between conventional STRs, ranging in size from 100 to 450 bp; miniSTRs, with amplicon sizes less than 200 bp; and single nucleotide polymorphisms (SNPs). Results were collated and analysed and, in general, mini-STR systems were shown to be the most effective…”

Data Driven Decisions

“A recent meeting by the ENFSI and EDNAP groups on the 4–5 April, 2005, in Glasgow, UK, it was unanimously agreed that the process of standardization within Europe should take account of recent work that unequivocally demonstrated that chance of obtaining a result from a degraded sample was increased when small amplicons (mini-STRs) were analysed…”
The Copenhagen Forensic Genetic Summer School  
June 27-28, 2012

Characterizing New STR Loci

Main Points:

- In April 2011, the FBI announced plans to expand the core loci for the U.S. beyond the current 13 CODIS STRs
- Our group is collecting U.S. population data on new loci and characterizing them to aid understanding of various marker combinations
- We are collecting all available information from the literature on the 24 commonly used autosomal STR loci

Presentations/Publications:

- AAFS 2011 presentation
- Hares (2012) Expanding the U.S. core loci... FSI Genetics 6(1): e52-e54

---

Article in the January 2012 issue of Forensic Science Review

Available at http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm

Biology and Genetics of New Autosomal STR Loci Useful for Forensic DNA Analysis


ABSTRACT: Short tandem repeats (STRs) are regions of tandemly repeated DNA segments found throughout the human genome that vary in length through insertion, deletion, or mutation with a core repeated DNA sequence. Forensic laboratories commonly use tandem repeat repeats, containing a four-base pair (4bp) repeat structure such as GATA. In 1997, the Federal Bureau of Investigation (FBI) laboratory selected 13 STR loci that form the backbone of the U.S. national DNA database. Building on the European expansion in 2009, the FBI announced plans in April 2011 to expand the U.S. core loci to as many as 20 STRs to include more global DNA data sharing. Commercial STRs that enable consistency in marker use and allele nomenclature between laboratories and help improve quality control. The STRbase website, maintained by the U.S. National Institute of Standards and Technology (NIST), contains critical information on STR markers used in human identity testing.

Key Words: Autosomal genetic markers, CODIS STRs, core loci, DNA typing, European Standard Set, expanded U.S. core loci, short tandem repeat (STR) STR kits.

Discusses the 24 autosomal STR loci available in commercial kits

---

The 11 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>AAAGT</td>
<td>3 to 49</td>
<td>178</td>
</tr>
<tr>
<td>D6S1043</td>
<td>6q15</td>
<td>AGAT/AGAC</td>
<td>8 to 25</td>
<td>25</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; ‡SE33 alleles have complex repeat structure
The Copenhagen Forensic Genetic Summer School

June 27-28, 2012

25 Alleles Reported in the Literature for D1S1656

<table>
<thead>
<tr>
<th>Allele</th>
<th>Homogen</th>
<th>Promega</th>
<th>Promega</th>
<th>ABI</th>
<th>Repeat Structure</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>133 bp</td>
<td>222 bp</td>
<td>171 bp</td>
<td>T[AGA]_171</td>
<td>Phillips et al. (2010)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>127 bp</td>
<td>265 bp</td>
<td>175 bp</td>
<td>T[AGA]_175</td>
<td>Phillips et al. (2010)</td>
<td></td>
</tr>
<tr>
<td>10 (a)</td>
<td>141 bp</td>
<td>230 bp</td>
<td>179 bp</td>
<td>T[AGA]_179</td>
<td>Larios et al. (1996)</td>
<td></td>
</tr>
<tr>
<td>10 (b)</td>
<td>141 bp</td>
<td>230 bp</td>
<td>179 bp</td>
<td>T[AGA]_179</td>
<td>Phillips et al. (2010)</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>145 bp</td>
<td>234 bp</td>
<td>183 bp</td>
<td>T[AGA]_183</td>
<td>Larios et al. (1996)</td>
<td></td>
</tr>
<tr>
<td>12 (a)</td>
<td>149 bp</td>
<td>238 bp</td>
<td>187 bp</td>
<td>T[AGA]_187</td>
<td>Larios et al. (1996)</td>
<td></td>
</tr>
<tr>
<td>12 (b)</td>
<td>149 bp</td>
<td>238 bp</td>
<td>187 bp</td>
<td>T[AGA]_187</td>
<td>Larios et al. (1996)</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>153 bp</td>
<td>242 bp</td>
<td>191 bp</td>
<td>T[AGA]_191</td>
<td>Larios et al. (1996)</td>
<td></td>
</tr>
<tr>
<td>13 (a)</td>
<td>153 bp</td>
<td>242 bp</td>
<td>191 bp</td>
<td>T[AGA]_191</td>
<td>Phillips et al. (2010)</td>
<td></td>
</tr>
<tr>
<td>13 (b)</td>
<td>153 bp</td>
<td>242 bp</td>
<td>191 bp</td>
<td>T[AGA]_191</td>
<td>Phillips et al. (2010)</td>
<td></td>
</tr>
</tbody>
</table>

NIST U.S. Population Allele Frequencies

<table>
<thead>
<tr>
<th>Allele</th>
<th>African American (N = 341)</th>
<th>Caucasian (N = 381)</th>
<th>Hispanic (N = 236)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.01433</td>
<td>0.00277</td>
<td>0.00630</td>
</tr>
<tr>
<td>11</td>
<td>0.04871</td>
<td>0.07756</td>
<td>0.02731</td>
</tr>
<tr>
<td>12</td>
<td>0.06304</td>
<td>0.11773</td>
<td>0.08824</td>
</tr>
<tr>
<td>13</td>
<td>0.10029</td>
<td>0.06648</td>
<td>0.11655</td>
</tr>
<tr>
<td>14</td>
<td>0.25788</td>
<td>0.07695</td>
<td>0.11765</td>
</tr>
<tr>
<td>14.3</td>
<td>0.00716</td>
<td>0.00277</td>
<td>0.00420</td>
</tr>
<tr>
<td>15</td>
<td>0.15616</td>
<td>0.14820</td>
<td>0.13866</td>
</tr>
<tr>
<td>15.3</td>
<td>0.03009</td>
<td>0.05617</td>
<td>0.05042</td>
</tr>
<tr>
<td>16</td>
<td>0.11032</td>
<td>0.13573</td>
<td>0.17437</td>
</tr>
<tr>
<td>16.3</td>
<td>0.10029</td>
<td>0.06094</td>
<td>0.05462</td>
</tr>
<tr>
<td>17</td>
<td>0.02865</td>
<td>0.04709</td>
<td>0.04202</td>
</tr>
<tr>
<td>17.3</td>
<td>0.05014</td>
<td>0.12596</td>
<td>0.14496</td>
</tr>
<tr>
<td>18</td>
<td>0.00287</td>
<td>0.00554</td>
<td>0.00630</td>
</tr>
<tr>
<td>18.3</td>
<td>0.02436</td>
<td>0.05125</td>
<td>0.02521</td>
</tr>
<tr>
<td>19.3</td>
<td>0.00573</td>
<td>0.01385</td>
<td>0.00420</td>
</tr>
</tbody>
</table>

D1S1656 Characteristics

- 15 alleles observed
- 92 genotypes observed
- >89% heterozygotes (heterozygosity = 0.8934)
- 0.0220 Probability of Identity ($P_I$)

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

These values have been calculated for all 24 STR loci across the U.S. population samples examined

Addendum to expanding the CODIS core loci in the United States

http://www.fbi.gov/about

The Copenhagen Forensic Genetic Summer School

June 27-28, 2012

Determination of Additional CODIS Core Loci


Steps in Adopting Genetic Markers

Role of the NIST Human Identity Project Team

- Assay Construction
- Kit Development
- Populations Study
- Information Gathered
- STRBase website

Research

NIH-Funded or Other

Development

Company (e.g., Promega)

Forensic Application

Forensic Labs

- Kit Testing
- Release to Community
- Internal Validation
- Use in Casework
- Court Presentation/Acceptance

- Loci Described
- Kit Development
- STR Loci present in STR kits ranked ordered by their variability

<table>
<thead>
<tr>
<th>Loci</th>
<th>Observed Genotypes</th>
<th>Hom. (obs)</th>
<th>P value</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEsh</td>
<td>53 292 0.9360 0.0063</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D21S1338</td>
<td>13 68 0.8785 0.0219</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D19S566</td>
<td>15 92 0.9304 0.0220</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D18S51</td>
<td>21 91 0.8689 0.0256</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D12S391</td>
<td>23 110 0.8795 0.0257</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FGA</td>
<td>26 93 0.8742 0.0299</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Penta E*</td>
<td>16 71 0.8754 0.0396</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D21S11</td>
<td>25 81 0.8358 0.0410</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D19S533</td>
<td>16 76 0.8124 0.0561</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D5S179</td>
<td>11 45 0.7878 0.0582</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vWA</td>
<td>11 38 0.8060 0.0622</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D7S602</td>
<td>11 32 0.8070 0.0734</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TH01</td>
<td>8 24 0.7590 0.0794</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D16S539</td>
<td>9 28 0.7825 0.0784</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D13S317</td>
<td>8 29 0.7655 0.0912</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D10S1248</td>
<td>12 39 0.7825 0.0837</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D2S441</td>
<td>14 41 0.7772 0.0855</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D9S1358</td>
<td>11 30 0.7669 0.0873</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D22S1045</td>
<td>11 42 0.7697 0.0933</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSF1PO</td>
<td>9 30 0.7537 0.1071</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D5S818</td>
<td>9 34 0.7164 0.1192</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TPOX</td>
<td>9 28 0.6983 0.1283</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N = 938 (only validated samples used)

- Better for mixtures
- More polymorphic than current CODIS 13 STRs
- Better for kinship
- (low mutation rate)

- There are several loci
- Met
- Enables target goals to be established
- Allows protocols to be performed
- Evaluates if desired goals are met
- Examines if target goals can be met
- Creates tools to meet target goals
- Identifies issues for the community
- Provides updates for DNA analysts and informs manufacturers
- Sets desired target goals
- Establishes target goals
- Identifies recommendations
- Authors provide reports on status of CODIS Core Loci project in meetings
- Determines core locus
- Established
- More polymorphic than current CODIS 13 STRs
- Better for kinship
- (low mutation rate)

- More alleles seen
- Better for mixtures
- More polymorphic than current CODIS 13 STRs

- N = 938

- 23 STR Loci present in STR kits ranked ordered by their variability

- Better for mixtures
- (more alleles seen)

- CSF1PO
- TPOX
- Penta D & Penta E run on subset (N = 658)

- 361 Caucasians
- 341 African Americans
- 236 Hispanics
- 341 African Americans
- 361 Caucasians
- More polymorphic than current CODIS 13 STRs

- 23 STR Loci present in STR kits ranked ordered by their variability

- Better for mixtures
- (more alleles seen)


Recent Court Decision Impacting Sale of STR Typing Kits

Disclaimer: The information contained herein is only as accurate as my understanding of the information available to me at the time this presentation was given. Things are still evolving with this case...

http://www.appliedbiosystems.com

Notice on ABI STR Kits

IMPORTANT NOTICE

The UNITED STATES DISTRICT COURT FOR THE WESTERN DISTRICT OF WISCONSIN ruled that certain products (listed below) sold by Life Technologies Corporation ("Life") can only be used by customers for forensic and paternity uses ("Licensed Use"). Specifically, the Court held that the license Life holds from Promega Corporation ("Promega") does not include the following applications: (1) chimerism (which involves determining the relative amount present of two different types of DNA); (2) classifying molar specimens (which involves determining whether a mole is present and what type it is); (3) cell line authentication (which involves a determination of whether two cell lines are unique); (4) determination of fetal sex; (5) cancer analysis; (6) genetic research; (7) non-casework-related forensic applications such as general research in forensics or teaching and training of persons not employed in a forensic laboratory; (8) maternal cell contamination; and (9) sample tracking. Accordingly, this notice replaces any other label license or use statement for the listed products only as those labels or statements relate to the use of such products under the Promega license. Any other restrictions, such as regulatory restrictions, related to the use of these products are not affected by this notice. If a customer has any question regarding whether their intended use is within or outside the Licensed Use, please contact LicenseQuery@lifetech.com.

The following products are subject to this notice:

- 4322288 AmpFISTR® Identifier® PCR Amplification Kit
The following products are subject to this notice:

- 432288 AmpFSTR® Identifiler® PCR Amplification Kit
- 446781 AmpFSTR® Identifiler® Direct PCR Amplification Kit (1000 tests)
- 447368 AmpFSTR® Identifiler® Plus PCR Amplification Kit
- 447372 AmpFSTR® Minifiler™ PCR Amplification Kit
- 4415021 AmpFSTR® NGM™ PCR Amplification Kit (1000 rxn)
- 4415020 AmpFSTR® NGM™ PCR Amplification Kit (200 rxn)
- 4457890 AmpFSTR® NGM Select™ PCR Amplification Kit (1000 rxn)
- 4457889 AmpFSTR® NGM Select™ PCR Amplification Kit (200 rxn)
- 403038 AmpFSTR® Profiler® PCR Amplification Kit
- 430326 AmpFSTR® Profiler Plus® PCR Amplification Kit
- 430284 AmpFSTR® Profiler Plus® ID PCR Amplification Kit
- 430246 AmpFSTR® Cofiler® PCR Amplification Kit
- 4307133 AmpFSTR® SGM Plus® PCR Amplification Kit
- 4382699 AmpFSTR® SEfiler Plus™ PCR Amplification Kit
- 4305795 AmpFSTR® Profiler Plus® and AmpFLSTR® Cofiler® Kits
- 4330621 AmpFSTR® Profiler Plus® ID Kit and AmpFLSTR® Cofiler® Kit
- 4359513 AmpFSTR® Yfiler® PCR Amplification Kit
- 4382306 AmpFSTR® Sinofiler™ PCR Amplification Kit
- 4382324 AmpFSTR® Sinofiler™ PCR Amplification Kit Primer Set
This Patent was Previously Licensed to Both Promega and Applied Biosystems

United States Patent

[19] Patent Number: 5,564,759

Dnym Typing With Short Tandem Repeat Polymorphism
and Application to Human Identification
[A1] [A2] [A4] [A5] [A6] [A7]

Assignee: Baylor College of Medicine, Houston, Tex.
Appl. No.: 647,535
Filed: Jan. 31, 1991

This Patent expired after 17 years on November 15, 2011

Timeline to Court Case

On October 20, 2009, Life Technologies (LTI = ABI) sent a letter to Promega asserting new interpretation of the 1996 License Agreement which would have required Promega to pay >$50M within 60 days of demand (>20X what has previously been paid)

During January 2010 meeting, Promega and ABI agreed to conduct audits about royalty payments

In a February 10, 2010 letter, LTI conceded it had no documentary evidence to support its novel claim

In a May 4, 2010 letter, LTI demanded arbitration of a supposed royalty BUT ABI had breached the 1996 agreement

In a July 7, 2010 follow-up letter, Promega sought a declaration that LTI and ABI have willfully infringed 5 patents by selling outside permitted fields (in clinical diagnostics, clinical research, & research markets)

Trial Dates and Results

February 6, 7, 8, 9, 10, 13, 14, 15 (2012)

Jury verdict on February 15, 2012

Judgment on February 23, 2012

Promega received $52,009,941 from Life Technologies (Applied Biosystems)
Jury Verdict on February 15, 2012

IN THE UNITED STATES DISTRICT COURT
FOR THE WESTERN DISTRICT OF WISCONSIN

PROMEGA CORPORATION,
Plaintiff,

v.

LIFE TECHNOLOGIES CORPORATION,
INVITROGEN IP HOLDINGS, INC. and
APPLIED BIOSYSTEMS, LLC,
Defendants.

SPECIAL VERDICT

Answer Question No. 1: What is the total dollar amount of worldwide STR kit sales made between August 29, 2006 through the end of January 2012 by defendants Life Technologies Corporation, Invitrogen IP Holdings, Inc. and Applied Biosystems, LLC?

Answer: $ 707,618,247

Answer Question No. 5:

Question No. 5: What profits, if any, did plaintiff lose as a result of defendants’ sales that you found in Question No. 4?

Answer: $ 53,009,941
Forensic DNA Labs

- Forensic & paternity testing DNA laboratories performing casework should not be directly impacted by this court ruling because ABI has a license to sell for casework applications

Potential Impact on NIST

- Judge has narrowly defined that only forensic labs and paternity labs may be sold ABI kits – NOT universities or other research labs
- I have spoken with lawyers from both Promega and Life Technologies (Applied Biosystems)
- The initial plan was for Promega to work with LTI/ABI to develop a permitted purchase list institution by institution
  - Promega wants to take over cell line authentication market and other clinical DNA applications
- Purchase of ABI STR kits for forensic research and training may not be permitted in the future
- Both companies would like to keep their customers happy
New STR Kits

Commercially Available STR Kits

<table>
<thead>
<tr>
<th>Applied Biosystems (17)</th>
<th>Promega Corporation (15)</th>
<th>Qiagen (2010)</th>
</tr>
</thead>
<tbody>
<tr>
<td>• AmpFRSTR Blue (1996)</td>
<td>• PowerPlex 1.1 (1997)</td>
<td>Primary sales kit in Europe</td>
</tr>
<tr>
<td>• AmpFRSTR Green I (1997)</td>
<td>• PowerPlex 1.2 (1998)</td>
<td>Due to patent restrictions</td>
</tr>
<tr>
<td>• Profiler (1997)</td>
<td>• PowerPlex 2.1 (1999)</td>
<td>cannot sell in U.S.</td>
</tr>
<tr>
<td>• Profiler Plus (1997)</td>
<td>• PowerPlex 16 (2000)</td>
<td>• ESSplex</td>
</tr>
<tr>
<td>• COfiler (1998)</td>
<td>• PowerPlex ES (2002)</td>
<td>• ESSplex SE</td>
</tr>
<tr>
<td>• SGM Plus (1999)</td>
<td>• PowerPlex Y (2003)</td>
<td>• Decaplex SE</td>
</tr>
<tr>
<td>• Identifiler (2001)</td>
<td>• PowerPlex S5 (2007)</td>
<td>• IDplex</td>
</tr>
<tr>
<td>• Profiler Plus ID (2001)</td>
<td>• PowerPlex 16 HS (2009)</td>
<td>• Nonaplex ESS</td>
</tr>
<tr>
<td>• SEFilter (2002)</td>
<td>• PowerPlex ESX 16 (2009)</td>
<td>• Hexaplex ESS</td>
</tr>
<tr>
<td>• Yfiler (2004)</td>
<td>• PowerPlex ESX 17 (2009)</td>
<td>• HD (Chimera)</td>
</tr>
<tr>
<td>• MiniFilter (2007)</td>
<td>• PowerPlex ESI 16 (2009)</td>
<td>• Argus X-12</td>
</tr>
<tr>
<td>• SEFilter Plus (2007)</td>
<td>• PowerPlex ESI 17 (2009)</td>
<td>• Argus Y-12</td>
</tr>
<tr>
<td>• SnpFilter (2008)</td>
<td>• PowerPlex 18D (2011)</td>
<td>• DiPlex (30 InDels)</td>
</tr>
<tr>
<td>• Identifiler Direct (2009)</td>
<td>• PowerPlex 21 (2012)</td>
<td></td>
</tr>
<tr>
<td>• NGM (2009)</td>
<td>• PowerPlex ESI 17 Pro (2012)</td>
<td></td>
</tr>
<tr>
<td>• Identifiler Plus (2010)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• NGM SElect (2010)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

~1/3 of all STR kits were released in the last three years

Same DNA Sample Tested with Five STR Kits
STR 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

STR Kit Concordance Testing

\[
\begin{array}{c}
\text{PCR primer set 1} \\
\text{STR repeat region} \\
\text{PCR primer set 2} \\
\text{PCR primer set 1} \\
\text{STR repeat region} \\
\text{PCR primer set 2} \\
\end{array}
\]

\[
\begin{array}{c}
\text{allele a} \\
\text{Set 1 Amplicons} \\
\text{Set 2 Amplicons} \\
\text{allele b} \\
\text{Set 1 Amplicons} \\
\text{Set 2 Amplicons} \\
\end{array}
\]

If no primer binding site mutations

If a primer binding site mutation (≠) exists

\[
\begin{array}{c}
\text{a b} = \text{a b} \\
\text{Set 1 Amplicons} \\
\text{Set 2 Amplicons} \\
\text{a b} \neq \text{b} \\
\text{Set 1 Amplicons} \\
\text{Set 2 Amplicons} \\
\end{array}
\]
The Copenhagen Forensic Genetic Summer School

STR Kit Performance

21plex (PowerPlex 21) vs 16plex (Identifiler Plus)

Three fully heterozygous (except PT83 at Penta D) pristine DNA samples were examined in a dilution series with PowerPlex 21 and Identifiler Plus. Results are ordered by amplion size and dye color.

<table>
<thead>
<tr>
<th>PowerPlex 21</th>
<th>Identifiler Plus</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 cycles (Sa@3kV)</td>
<td>28 cycles (10s@3kV)</td>
</tr>
</tbody>
</table>

Having 5 additional loci did not adversely impact success rates.

Total alleles possible = 875
Total alleles present = 805

12% detected

STR Kit Comparisons Searching for Primer Binding Site Mutations

Kits compared

<table>
<thead>
<tr>
<th>Samples</th>
<th>Loci compared</th>
<th>Comparisons</th>
<th># Differences</th>
<th>Concordance (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SGM-ID</td>
<td>1436</td>
<td>11</td>
<td>15,796</td>
<td>1</td>
</tr>
<tr>
<td>Id-ProPlus</td>
<td>1427</td>
<td>10</td>
<td>14,270</td>
<td>1</td>
</tr>
<tr>
<td>ID-ISplex</td>
<td>669</td>
<td>16</td>
<td>19,704</td>
<td>19</td>
</tr>
<tr>
<td>ID-PP16</td>
<td>662</td>
<td>14</td>
<td>9,268</td>
<td>4</td>
</tr>
<tr>
<td>ID-MiniFiler</td>
<td>1308</td>
<td>9</td>
<td>11,772</td>
<td>27</td>
</tr>
<tr>
<td>SGM-NGM</td>
<td>1436</td>
<td>11</td>
<td>15,796</td>
<td>4</td>
</tr>
<tr>
<td>ID-NGM</td>
<td>1449</td>
<td>1427</td>
<td>1445</td>
<td>1445</td>
</tr>
<tr>
<td>ProPlus-NGM</td>
<td>1427</td>
<td>1436</td>
<td>1427</td>
<td>1427</td>
</tr>
<tr>
<td>SGM-ESI</td>
<td>1436</td>
<td>1427</td>
<td>1427</td>
<td>1427</td>
</tr>
<tr>
<td>ProPlus-ESS</td>
<td>1436</td>
<td>1427</td>
<td>1427</td>
<td>1427</td>
</tr>
<tr>
<td>ESI-ESS</td>
<td>1445</td>
<td>1445</td>
<td>1445</td>
<td>1445</td>
</tr>
<tr>
<td>ESI-ESSplex</td>
<td>1445</td>
<td>1445</td>
<td>1445</td>
<td>1445</td>
</tr>
<tr>
<td>ESI-NGMSelect</td>
<td>715</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

128 kit-to-kit comparisons
1,104,031 allele comparisons
1224 differences observed
~99.9% concordance
(many corrected now)

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

Random Match Probability for Various Combinations (assuming unrelated individuals)

<table>
<thead>
<tr>
<th>STR Marker Combinations</th>
<th>RMP*</th>
<th>1 in ...</th>
</tr>
</thead>
<tbody>
<tr>
<td>13 CODIS STRs (+D2S1338, D19S433)</td>
<td>6.0E-16</td>
<td>1.7E+15</td>
</tr>
<tr>
<td>15 STRs (+D2S441, D10S1248, D22S1045)</td>
<td>7.3E-19</td>
<td>1.4E+18</td>
</tr>
<tr>
<td>18 STRs (+D15S1656, D12S391)</td>
<td>4.9E-22</td>
<td>2.0E+21</td>
</tr>
<tr>
<td>20 STRs (+D15S1656, D12S391)</td>
<td>2.8E-25</td>
<td>3.6E+24</td>
</tr>
<tr>
<td>23 STRs (+SE33, Penta D, Penta E)</td>
<td>1.2E-30</td>
<td>8.4E+29</td>
</tr>
</tbody>
</table>

*RMP values calculated by combining Probability of Identity values for each locus

More Loci are Useful in Situations Involving Relatives

- **Missing Persons** and Disaster Victim Identification (kinship analysis)
- Immigration Testing (often limited references)
  - Recommendations for 25 STR loci
- Deficient Parentage Testing
  - Often needed if only one parent and child are tested

Relationship testing labs are being pushed to answer more difficult genetic questions... and we want to make sure the right tools are in place.

**In February 25, 2011 issue of Forensic Science International...**

Examples of kinship analysis where Profiler Plus™ was not discriminatory enough for the identification of victims using DNA identification

D. Hartman, M. L. Benton, L. Meerson, J. Breyer, M. Spindler, A. Stock

Disaster victim identification from the 2009 Victorian bushfires relied on DNA (62% involved kinship associations rather than direct matching)

They advocate additional autosomal STR loci to aid kinship associations

How do 13 loci perform for kinship analysis?

The degree of overlap corresponds with possible values for false positive or false negative results.

Parent-offspring comparisons: No overlap between unrelated and related LR distributions

Full sibling comparisons:
- False positive rate = 0.027
- False negative rate = 0.033

Half sibling comparisons:
- False positive rate = 0.155
- False negative rate = 0.168
Do additional loci improve kinship determination?

- Additional autosomal STR loci exist in new STR kits and are being studied at NIST in U.S. population sample sets
- To avoid potential adventitious matches with large DNA databases, enable greater international data sharing, and aid missing persons applications, it is highly likely that additional loci will be added to the U.S. core in the future

Summary

Contact Information

Becky Hill
Research Biologist
becky.hill@nist.gov
301-975-4275

Mike Coble
Research Biologist
michael.coble@nist.gov
301-975-4330

Acknowledgments: NIJ & FBI Funding

Thank you for your attention

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