The Future of Forensic DNA Typing

John M. Butler
NIST Applied Genetics Group
National Institute of Standards and Technology
Gaithersburg, Maryland

http://www.ojp.usdoj.gov/nij/pubs-sum/183697.htm

- Report published in Nov 2000
- Asked to estimate where DNA testing would be 2, 5, and 10 years into the future

Conclusions
STR typing is here to stay for a few years because of DNA databases that have grown to contain millions of profiles

My Thoughts on the Future

- Near-term future
  - New autosomal STR loci for expanded core loci
  - Expanded use of databases (e.g., familial searching)
  - Rapid DNA testing

- More distant future
  - Next-generation DNA sequencing?
  - Loci besides STRs for identity testing?
  - Phenotyping capabilities?

STRs vs SNPs Article

- SNPs are unlikely to replace STRs for routine forensic DNA testing due to challenges with high-level multiplexing and mixture detection/interpretation
- Most likely use of SNPs will be as ancestry-informative markers (AIMs) for sample ethnicity estimation

National Academies Report on Forensic Science

- Released February 18, 2009
- Entitled “Strengthening Forensic Science in the United States: A Path Forward”
- 13 recommendations provided to Congress
- Recommends establishing a National Institute of Forensic Science (NIFS)
- NIST will have a role in NIFS and our group has been asked to contribute expertise regarding validation and testing of DNA systems as a model for other forensic disciplines

Forensic Science Review Article

- 2009 review article covers 160 DNA articles published in 2007-2008

National Commission on the Future of DNA Evidence

- Conclusions
  - STR typing is here to stay for a few years because of DNA databases that have grown to contain millions of profiles

- Near-term future
  - New autosomal STR loci for expanded core loci
  - Expanded use of databases (e.g., familial searching)
  - Rapid DNA testing

- More distant future
  - Next-generation DNA sequencing?
  - Loci besides STRs for identity testing?
  - Phenotyping capabilities?
Value of a Historical Review

“If you want to understand today, you have to search yesterday.”

– Attributed to Pearl Buck

Historical Perspective on DNA Typing

2012: DNA is an important part of the criminal justice system

Stages of Forensic DNA Progression

<table>
<thead>
<tr>
<th>Stages</th>
<th>Time Frame</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Exploration</strong></td>
<td>1985-1995</td>
<td>Beginnings, different methods tried (RFLP and early PCR)</td>
</tr>
<tr>
<td><strong>Stabilization</strong></td>
<td>1995-2005</td>
<td>Standardization to STRs, selection of core loci, implementation of Quality Assurance Standards</td>
</tr>
<tr>
<td><strong>Growth</strong></td>
<td>2005-2012</td>
<td>Rapid growth of DNA databases, extended applications pursued</td>
</tr>
<tr>
<td><strong>Sophistication</strong></td>
<td>The Future</td>
<td>Expanding tools available, confronting privacy concerns</td>
</tr>
</tbody>
</table>

Basis of DNA Profiling

The genome of each individual is unique (with the exception of identical twins) and is inherited from parents

Probe subsets of genetic variation in order to differentiate between individuals (statistical probabilities of a random match are used)

DNA typing must be performed efficiently and reproducibly (information must hold up in court)

Current standard DNA tests DO NOT look at genes – little/no information about race, predisposition to disease, or phenotypical information (eye color, height, hair color) is obtained

Expanding the Forensic Core Competency

Type of Match

- **Direct**
- **Indirect**

Type of Search

- **1-to-1**
- **1-to-many**

Type of Typing

- **Standard STR Typing**
- **Paternity Testing**
- **Kinship Analysis**
- **Standard STR Typing for Database Search**
- **Familial Searching/Missing Persons**

Level of Certainty

- **High**
- **Low**

Expanding the Forensic Core Competency

- **Direct** Type of Search
- **Indirect** Type of Search

- **Standard STR Typing**
- **Paternity Testing**
- **Kinship Analysis**
- **Standard STR Typing for Database Search**
- **Familial Searching/Missing Persons**

The FBI has selected 13 core STR loci that must be run in all DNA tests in order to provide a common currency with DNA profiles
Growth of DNA Databases

- Within the U.S., we have benefited from significant federal funding over the past seven years
- Expanded laws now enable more offenders to be included
- Have effectively locked technology with core STR markers used to generate DNA profiles that now number greater than 10 million profiles

Half of the U.S. Requires Arrestee DNA Testing

Growth in Numbers of U.S. States Requiring DNA Collection for Various Offenses

<table>
<thead>
<tr>
<th>Offenses</th>
<th>1999</th>
<th>2004</th>
<th>2008</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex crimes</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>All violent crimes</td>
<td>36</td>
<td>48</td>
<td>50</td>
</tr>
<tr>
<td>Burglary</td>
<td>14</td>
<td>47</td>
<td>50</td>
</tr>
<tr>
<td>All felons</td>
<td>5</td>
<td>37</td>
<td>47</td>
</tr>
<tr>
<td>Juveniles</td>
<td>24</td>
<td>32</td>
<td>32</td>
</tr>
<tr>
<td>Arrestees/suspects</td>
<td>1</td>
<td>4</td>
<td>14</td>
</tr>
</tbody>
</table>


Starting initially with sex crimes, each category has grown in the past decade… burglary, all felons, arrestees…

National DNA Index System (NDIS)

- Launched in October 1998 and now links all 50 states
- Used for linking serial crimes and unsolved cases with repeat offenders
- Convicted offender and forensic case samples along with a missing persons index
- Requires 13 core STR markers
- >170,000 investigations aided nationwide as of April 2012
- Contains more than 11 million DNA profiles

Sources: http://www.fbi.gov/hq/lab/codis/index1.htm
Growth in Numbers of DNA Profiles Present in Various NDIS Indices (cumulative totals by year)

<table>
<thead>
<tr>
<th>Year ending Dec 31</th>
<th>Forensic</th>
<th>ConvictedOffender</th>
<th>Arrestee</th>
<th>TotalOffender*</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>21,625</td>
<td>441,181</td>
<td>--</td>
<td>441,181</td>
</tr>
<tr>
<td>2001</td>
<td>27,897</td>
<td>750,929</td>
<td>--</td>
<td>750,929</td>
</tr>
<tr>
<td>2002</td>
<td>46,177</td>
<td>1,247,163</td>
<td>--</td>
<td>1,247,163</td>
</tr>
<tr>
<td>2003</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td>248,943</td>
<td>6,398,874</td>
<td>140,719</td>
<td>6,539,919</td>
</tr>
<tr>
<td>2009</td>
<td>298,369</td>
<td>7,389,917</td>
<td>351,926</td>
<td>7,743,329</td>
</tr>
<tr>
<td>2010</td>
<td>351,951</td>
<td>8,559,841</td>
<td>668,849</td>
<td>9,233,554</td>
</tr>
</tbody>
</table>

Source: FBI Laboratory’s CODIS Unit

In the last two years (2009, 2010):
- 103,008 forensic samples added
- 2,693,635 offender samples added

Expanding the CODIS Core Loci

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

Currently there are 24 autosomal STR markers present in commercial kits

- 13 CODIS loci
- 7 ESS loci

Possible scenarios for extending sets of genetic markers to be used in national DNA databases

- (a) Extra loci would be included (due to large PCR multiplexes)
- (b) Some loci may be dropped to enable replacement with better loci
- (c) Maintaining connection to legacy data is essential
- (d) Highly unlikely to start over with new loci
Proposed Expanded CODIS Core Loci


<table>
<thead>
<tr>
<th>Locus</th>
<th>Section A</th>
<th>Section B (in order of preference)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSF1PO</td>
<td>D5S818</td>
<td>D16S539</td>
</tr>
<tr>
<td>D19S433</td>
<td>D3S1358</td>
<td>D7S820</td>
</tr>
<tr>
<td>D2S441</td>
<td>CSF1PO</td>
<td>D19S433</td>
</tr>
<tr>
<td>D3S1338</td>
<td>D7S820</td>
<td>D2S441</td>
</tr>
<tr>
<td>D12S391</td>
<td>D3S1358</td>
<td>D2S441</td>
</tr>
<tr>
<td>D12S391</td>
<td>D7S820</td>
<td>D3S1358</td>
</tr>
<tr>
<td>D13S317</td>
<td>TH01</td>
<td>D12S391</td>
</tr>
<tr>
<td>D18S51</td>
<td>D13S317</td>
<td>TH01</td>
</tr>
<tr>
<td>Amelogenin</td>
<td>D18S51</td>
<td>D13S317</td>
</tr>
<tr>
<td>D21S11</td>
<td>D18S51</td>
<td>Amelogenin</td>
</tr>
<tr>
<td>FGA</td>
<td>D21S11</td>
<td>D18S51</td>
</tr>
<tr>
<td>D7S820</td>
<td>FGA</td>
<td>D21S11</td>
</tr>
<tr>
<td>D13S317</td>
<td>D7S820</td>
<td>FGA</td>
</tr>
<tr>
<td>D16S539</td>
<td>D13S317</td>
<td>D7S820</td>
</tr>
<tr>
<td>D19S433</td>
<td>D16S539</td>
<td>D13S317</td>
</tr>
<tr>
<td>D19S433</td>
<td>D16S539</td>
<td>D13S317</td>
</tr>
<tr>
<td>D19S433</td>
<td>Amelogenin</td>
<td>D16S539</td>
</tr>
<tr>
<td>D19S433</td>
<td>D16S539</td>
<td>Amelogenin</td>
</tr>
<tr>
<td>D19S433</td>
<td>Amelogenin</td>
<td>D16S539</td>
</tr>
<tr>
<td>D19S433</td>
<td>Amelogenin</td>
<td>D16S539</td>
</tr>
<tr>
<td>D19S433</td>
<td>Amelogenin</td>
<td>D16S539</td>
</tr>
<tr>
<td>D19S433</td>
<td>Amelogenin</td>
<td>D16S539</td>
</tr>
<tr>
<td>D19S433</td>
<td>Amelogenin</td>
<td>D16S539</td>
</tr>
</tbody>
</table>

20 required loci

Amelogenin for gender (appears)
11 STR (D7S820)

Determination of Additional CODIS Core Loci


Steps in Adopting Genetic Markers

Role of the NIST Human Identity Project Team

Loci Described

- Assay Construction
- Kit Development
- Kit Testing
- Population Study
- Internal Validation
- Release to Community
- Use in Casework
- Court
- Acceptance

Steps in Adopting Genetic Markers

- Assay Construction
- Kit Development
- Kit Testing
- Population Study
- Internal Validation
- Release to Community
- Use in Casework
- Court
- Acceptance

NIST 26plex Demonstration

- Our group at NIST has demonstrated that 25 autosomal STRs and amelogenin (26plex with 52 PCR primers) can be co-amplified with sensitivities similar to commercial STR kits
  - See also http://www.cstl.nist.gov/biotech/strbase/str26plex.htm

NIST 26plex

- Gender identification + 23 autosomal STR loci in a single amplification


NIST 26plex

- Gender identification + 23 autosomal STR loci in a single amplification


NIST 26plex

- Gender identification + 23 autosomal STR loci in a single amplification

PowerPlex 21

- Promega STR kit was released in January 2012
  - NIST has been working with this kit since spring 2011 primarily for concordance testing and has permission from Promega to discuss results
- Contains 20 autosomal STRs + amelogenin
- Enables examination of performance characteristics similar to a future U.S. megaplex containing at least 20 loci

DNA Dilution Series with PowerPlex 21

As expected with any STR kit/assay, allele dropout occurs below 100 pg.

Measurement of Allele Dropout and Extreme Peak Height Imbalance for 2 STR Kits

Three fully heterozygous cases at low to pristine DNA samples were examined in a dilution series with PowerPlex 21 and Identifiler Plus. Results are ordered by amplicon size and dye color.

Going Beyond the Core Competencies of Forensic DNA Testing

Core Competency

Standard STR Typing (DNA Profile)

- Direct Matching (or Parentage)
- Sufficient DNA quantity (ng)

Challenging kinship search questions

Familial Searching Attempts (looking for brothers or other relatives)

Solution: Additional Markers (Y-chromosomes, more STRs) and Multiple Reference Samples

Lower amounts of DNA being tested

Touch DNA Attempts (poor quality, mixtures, low-level stochastic effects)

Solution: Replicate Testing

Be very cautious when outside the box...

PowerPlex 21 (released by Promega in early 2012)

<table>
<thead>
<tr>
<th>PowerPlex 21</th>
<th>Identifiler Plus</th>
</tr>
</thead>
<tbody>
<tr>
<td>D3S1358</td>
<td>D3S1656</td>
</tr>
<tr>
<td>D3S1338</td>
<td>D12S391</td>
</tr>
<tr>
<td>D3S1338</td>
<td>D19S433</td>
</tr>
<tr>
<td>D6S1043</td>
<td>Penta E</td>
</tr>
<tr>
<td>D6S1043</td>
<td>Penta D</td>
</tr>
<tr>
<td>TH01</td>
<td>vWA</td>
</tr>
<tr>
<td>TH01</td>
<td>D21S11</td>
</tr>
<tr>
<td>D7S820</td>
<td>D5S818</td>
</tr>
<tr>
<td>D8S1179</td>
<td>TPX</td>
</tr>
<tr>
<td>D16S539</td>
<td>D13S391</td>
</tr>
<tr>
<td>D19S433</td>
<td>D1S1656</td>
</tr>
<tr>
<td>D2S1338</td>
<td>Penta E</td>
</tr>
<tr>
<td>D2S1338</td>
<td>Penta D</td>
</tr>
</tbody>
</table>

Promega 5-dye kit

- 13 CODIS STRs + amelogenin
- Penta D & Penta E (PP16 loci)
- D2S391 & D1S1656 (best new European loci)
- D6S1043 (previously only used in China – ABI STRfitter kit – highly polymorphic)

Promega 5-dye kit

- 13 CODIS STRs + amelogenin
- Penta D & Penta E (PP16 loci)
- D2S391 & D1S1656 (best new European loci)
- D6S1043 (previously only used in China – ABI STRfitter kit – highly polymorphic)

Measurement of Allele Dropout and Extreme Peak Height Imbalance for 2 STR Kits

Three fully heterozygous cases at low to pristine DNA samples were examined in a dilution series with PowerPlex 21 and Identifiler Plus. Results are ordered by amplicon size and dye color.

Going Beyond the Core Competencies of Forensic DNA Testing

Core Competency

Standard STR Typing (DNA Profile)

- Direct Matching (or Parentage)
- Sufficient DNA quantity (ng)

Challenging kinship search questions

Familial Searching Attempts (looking for brothers or other relatives)

Solution: Additional Markers (Y-chromosomes, more STRs) and Multiple Reference Samples

Lower amounts of DNA being tested

Touch DNA Attempts (poor quality, mixtures, low-level stochastic effects)

Solution: Replicate Testing

Be very cautious when outside the box...

PowerPlex 21 (released by Promega in early 2012)

<table>
<thead>
<tr>
<th>PowerPlex 21</th>
<th>Identifiler Plus</th>
</tr>
</thead>
<tbody>
<tr>
<td>D3S1358</td>
<td>D3S1656</td>
</tr>
<tr>
<td>D3S1338</td>
<td>D12S391</td>
</tr>
<tr>
<td>D3S1338</td>
<td>D19S433</td>
</tr>
<tr>
<td>D6S1043</td>
<td>Penta E</td>
</tr>
<tr>
<td>D6S1043</td>
<td>Penta D</td>
</tr>
<tr>
<td>TH01</td>
<td>vWA</td>
</tr>
<tr>
<td>TH01</td>
<td>D21S11</td>
</tr>
<tr>
<td>D7S820</td>
<td>D5S818</td>
</tr>
<tr>
<td>D8S1179</td>
<td>TPX</td>
</tr>
<tr>
<td>D16S539</td>
<td>D13S391</td>
</tr>
<tr>
<td>D19S433</td>
<td>D1S1656</td>
</tr>
<tr>
<td>D2S1338</td>
<td>Penta E</td>
</tr>
<tr>
<td>D2S1338</td>
<td>Penta D</td>
</tr>
</tbody>
</table>

Promega 5-dye kit

- 13 CODIS STRs + amelogenin
- Penta D & Penta E (PP16 loci)
- D2S391 & D1S1656 (best new European loci)
- D6S1043 (previously only used in China – ABI STRfitter kit – highly polymorphic)

Measurement of Allele Dropout and Extreme Peak Height Imbalance for 2 STR Kits

Three fully heterozygous cases at low to pristine DNA samples were examined in a dilution series with PowerPlex 21 and Identifiler Plus. Results are ordered by amplicon size and dye color.

Going Beyond the Core Competencies of Forensic DNA Testing

Core Competency

Standard STR Typing (DNA Profile)

- Direct Matching (or Parentage)
- Sufficient DNA quantity (ng)

Challenging kinship search questions

Familial Searching Attempts (looking for brothers or other relatives)

Solution: Additional Markers (Y-chromosomes, more STRs) and Multiple Reference Samples

Lower amounts of DNA being tested

Touch DNA Attempts (poor quality, mixtures, low-level stochastic effects)

Solution: Replicate Testing

Be very cautious when outside the box...

PowerPlex 21 (released by Promega in early 2012)

<table>
<thead>
<tr>
<th>PowerPlex 21</th>
<th>Identifiler Plus</th>
</tr>
</thead>
<tbody>
<tr>
<td>D3S1358</td>
<td>D3S1656</td>
</tr>
<tr>
<td>D3S1338</td>
<td>D12S391</td>
</tr>
<tr>
<td>D3S1338</td>
<td>D19S433</td>
</tr>
<tr>
<td>D6S1043</td>
<td>Penta E</td>
</tr>
<tr>
<td>D6S1043</td>
<td>Penta D</td>
</tr>
<tr>
<td>TH01</td>
<td>vWA</td>
</tr>
<tr>
<td>TH01</td>
<td>D21S11</td>
</tr>
<tr>
<td>D7S820</td>
<td>D5S818</td>
</tr>
<tr>
<td>D8S1179</td>
<td>TPX</td>
</tr>
<tr>
<td>D16S539</td>
<td>D13S391</td>
</tr>
<tr>
<td>D19S433</td>
<td>D1S1656</td>
</tr>
<tr>
<td>D2S1338</td>
<td>Penta E</td>
</tr>
<tr>
<td>D2S1338</td>
<td>Penta D</td>
</tr>
</tbody>
</table>

Promega 5-dye kit

- 13 CODIS STRs + amelogenin
- Penta D & Penta E (PP16 loci)
- D2S391 & D1S1656 (best new European loci)
- D6S1043 (previously only used in China – ABI STRfitter kit – highly polymorphic)
Victims of the Grim Sleeper


The Grim Sleeper’s Victims
1) Debra Jackson (age 29) – August 10, 1985
2) Henrietta Wright (age 35) – August 12, 1986
3) Thomas Steele (age 36) – August 14, 1986
4) Barbara Ware (age 23) – January 10, 1987
5) Bernita Sparks (age 26) – April 15, 1987
6) Mary Lowe (age 26) – October 31, 1987
7) Lachrica Jefferson (age 22) – January 30, 1988
8) Monique Alexander (age 18) – September 11, 1988
9) Enetra Washington (raped but survived) – November 1988
10) Princess Berthomieux (age 14) – March 19, 2002

Ballistics on bullets recovered from the victim’s bodies matched DNA evidence recovered

Over a 13 year gap in detected crimes, hence the “Sleeper” nickname

Identifying the Grim Sleeper

• Given that the murders spanned at least 25 years, the paternal relationship was likely father-son

• Undercover police shadowed C. Franklin’s father, Lonnie David Franklin, Jr., who lived in the vicinity of the murders

• Police collected a DNA sample from Lonnie Franklin – Direct match between L. Franklin and the Grim Sleeper

Putative Relative Is Found

• June 30, 2010: Second familial search of the California database yielded one likely relative

• Database profile belonged to Christopher Franklin (31 years old)
  – Profile added to the database in 2009 after a felony weapons possession charge

• Grim Sleeper profile matched C. Franklin’s profile with one allele at all 15 loci

• Both individuals shared the same Y-STR profile, indicating a possible paternal relationship

California Familial DNA Search Team

Familial DNA Testing Scores A Win in Serial Killer Case


Familial Searching in the U.S.

High-profile success in the Grim Sleeper case has led other states to consider familial searching

Experts say Texas might solve Twilight Serial Rapist cases with family DNA

Miami police on hunt for serial killer linked to 7 deaths

Familial DNA hunt sought in East Coast rape case

March 21, 2011 Virginia announced familial searching capability

Putative Relative Is Found

• June 30, 2010: Second familial search of the California database yielded one likely relative

• Database profile belonged to Christopher Franklin (31 years old)
  – Profile added to the database in 2009 after a felony weapons possession charge

• Grim Sleeper profile matched C. Franklin’s profile with one allele at all 15 loci

• Both individuals shared the same Y-STR profile, indicating a possible paternal relationship

Common PCR Thermal Cycling Times

Can we reduce PCR cycling times? What are the effects or limitations?

<table>
<thead>
<tr>
<th>Year</th>
<th>Run on a 9700 thermal cycler</th>
<th>Hot start</th>
<th>Time per cycle</th>
<th>Cycles</th>
<th>Post soak</th>
<th>Total time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997/98</td>
<td>Profiler Plus/Cofiler</td>
<td>11 min</td>
<td>3 min</td>
<td>28</td>
<td>60 min</td>
<td>2:52</td>
</tr>
<tr>
<td>1999</td>
<td>GGM Plus</td>
<td>11 min</td>
<td>3 min</td>
<td>28</td>
<td>45 min</td>
<td>2:35</td>
</tr>
<tr>
<td>2000</td>
<td>PowerPlex 16</td>
<td>12 min</td>
<td>1 min 45 s</td>
<td>32</td>
<td>30 min</td>
<td>3:00</td>
</tr>
<tr>
<td>2001</td>
<td>Identifier</td>
<td>13 min</td>
<td>3 min</td>
<td>28</td>
<td>60 min</td>
<td>2:34</td>
</tr>
<tr>
<td>2003</td>
<td>PowerPlex Y</td>
<td>12 min</td>
<td>1 min 45 s</td>
<td>32</td>
<td>30 min</td>
<td>3:18</td>
</tr>
<tr>
<td>2004</td>
<td>Hiler</td>
<td>11 min</td>
<td>3 min</td>
<td>20</td>
<td>60 min</td>
<td>2:45</td>
</tr>
<tr>
<td>2007</td>
<td>PowerPlex SS</td>
<td>2 min</td>
<td>4 min</td>
<td>30</td>
<td>60 min</td>
<td>3:23</td>
</tr>
<tr>
<td>2007</td>
<td>MiniFiler</td>
<td>11 min</td>
<td>1 min 20 s</td>
<td>30</td>
<td>45 min</td>
<td>3:16</td>
</tr>
<tr>
<td>2009</td>
<td>STR 16, 17 EX 16,17</td>
<td>2 min</td>
<td>4 min</td>
<td>30</td>
<td>45 min</td>
<td>3:22</td>
</tr>
<tr>
<td>2009</td>
<td>PowerPlex 16 HS</td>
<td>2 min</td>
<td>1 min 45 s</td>
<td>32</td>
<td>30 min</td>
<td>2:42</td>
</tr>
<tr>
<td>2009</td>
<td>PowerPlex 18D</td>
<td>2 min</td>
<td>1 min 30 s</td>
<td>29</td>
<td>30 min</td>
<td>2:35</td>
</tr>
<tr>
<td>2009</td>
<td>PowerPlex 18D</td>
<td>2 min</td>
<td>1 min 30 s</td>
<td>29</td>
<td>30 min</td>
<td>2:35</td>
</tr>
<tr>
<td>2009</td>
<td>Identifier Direct</td>
<td>11 min</td>
<td>3 min</td>
<td>26</td>
<td>25 min</td>
<td>2:34</td>
</tr>
<tr>
<td>2010</td>
<td>PowerPlex 18D</td>
<td>2 min</td>
<td>1 min 30 s</td>
<td>29</td>
<td>30 min</td>
<td>2:35</td>
</tr>
<tr>
<td>2011</td>
<td>PowerPlex 18D</td>
<td>2 min</td>
<td>1 min 15 s</td>
<td>27</td>
<td>30 min</td>
<td>2:15</td>
</tr>
</tbody>
</table>
Thermal Cylers

1. GeneAmp 9700 (Applied Biosystems)
2. Mastercycler Pro S (Eppendorf)
3. Rotor-Gene Q (Qiagen)
4. SmartCycler (Cepheid)

- Cycling for most STR kits is run in the ‘9600 emulation mode’ (1°C/s)
- Intended for real-time PCR

PCR Thermal Cycling Profile

<table>
<thead>
<tr>
<th>Temp</th>
<th>Time</th>
<th>Temp</th>
<th>Time</th>
<th>Temp</th>
<th>Time</th>
<th>Temp</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>95°C</td>
<td>10 min</td>
<td>95°C</td>
<td>1 min</td>
<td>59°C</td>
<td>1 min</td>
<td>72°C</td>
<td>1 min</td>
</tr>
<tr>
<td>95°C</td>
<td>5 s</td>
<td>58°C</td>
<td>10 s</td>
<td>72°C</td>
<td>1 min</td>
<td>60°C</td>
<td>60 min</td>
</tr>
</tbody>
</table>

Maximum heating/cooling rate of ~2 to 6°C/s (cycler dependent)

Rapid PCR Conditions

- 1 X Takara PCR mastermix, 1 U SpeedStar polymerase
- Premix Ex Taq™ (Perfect Real Time)
- 10 μL total reaction in a thin walled tube (8-strip)
- 2 μL of Identifier PCR primer mix
- ~1 ng of template DNA
- Utilize maximum ramp rate on thermal cyclers
- GeneAmp 9700 = 1.6°C/s (36 min)
- Rotor-Gene Q = 1.6°C/s (36 min)
- SmartCycler = 5.8°C/s (20 min)
- Mastercycler Pro S = 6.8°C/s (19 min)

Potential Applications with Rapid PCR Capabilities

- Improve overall laboratory throughput
  - Multiplex PCR amplification is already in many situations the longest part of the DNA analysis process (depending on DNA extraction and DNA quantitation methods)
  - With increased use of robotic sample preparation and expert system data analysis, bottleneck for sample processing will shift to time for PCR amplification...

- Enable new potential DNA biometric applications (because the overall DNA analysis process is faster)
  - Permit analysis of individuals at a point of interest such as an embassy, an airport, or a country border

A “Crystal Ball” to the Future?
Compromised Sample Improvements

- Better DNA extraction/recovery
- Continued use of miniSTRs
  - to improve success rates for recovery of information from compromised DNA evidence
- Replicate results for reproducibility
  - to improve reliability with low-template DNA testing

Highly degraded DNA

SNP genotyping in an extreme degradation case
Corpse half buried in a forest for ten years
- Uncovered by a forest fire
- Calcinated remains

Identifier success 0%

Highly degraded DNA

Carlos Vullo’s group from Argentina has published similar results with both SNPforID 52plex and IPATIMUP 38plex InDel reaction on common graves from Argentinian dictatorship period

Geographical Origin Prediction


Phenotypic Trait Prediction

Traits of interest

- Traits whose variation may be classified on discreet categories.
- Regulated by a relatively low number of genes.
- Fine example: Iris and hair pigmentation.
Phenotypic trait prediction


For this particular marker, predictive values for blue iris are high, but intermediate or brown phenotypes are not predicted.

Next Generation Sequencing

- High throughput or ultra-high throughput sequencing
- Thousands or millions of sequencing reads in parallel
- DNA sequencing, RNA expression, clinical diagnostics, microbial forensics, ...

1. Library preparation (genomic DNA or PCR amplicons)
2. Sequencing
3. Data analysis (assembly of reads)

Nature Biotechnology 2012

Performance comparison of benchtop high-throughput sequencing platforms

| Table 1 Price comparison of benchtop instruments and sequencing runs |
|--------------------------|--------------------|-----------------|-------------|
| Platform                  | List price         | Minimum throughput (read length) | Run time | Cost/Mb | Mth |
| 454 GS Junior (Roche)     | $108,000           | 35 Mb (400 bases)              | 8 h      | $31     | 4.4 |
| Ion Torrent PGM (Life Tech) | $80,490         | 100 Mb (100 bases)             | 14 h     | $31     | 5.5 |
| 454 GS Junior (Roche)     | $60,490           | 200 Mb (100 bases)             | 28 h     | $31     | 11.25 |

NGS Platforms

- Roche
  - 454 FLX
  - 454 GS Junior
- PacificBio
- PacBio RS
- Illumina
  - GA1x
  - HiSeq
  - HiScanSG
- MiSeq
- Life Tech
  - 5500 series
  - Ion torrent Proton
  - Ion torrent PGM (personal genome machine)

Generalized NGS Workflow

- 500 ng of genomic DNA
- Fragment to +200 bp
- Liquit PCR adaptation
- PCR
- Sequencing
- Hours to days
- One template per bead/droplet/spot
- Target specific genes or regions
- CODIS STRs, SNPs
- 500 ng of PCR product
Next Generation Sequencing
Forensic Applications

- Going in depth into STR loci and beyond
  - STRs are useful for legacy (databases)
  - SNPs within STRs identify ‘sub-alleles’
  - Millions of bases of sequence variants (SNPs)

- Opens up new human identity applications: biogeographical ancestry, externally visible traits, complex kinship, degraded samples, mixtures, other applications

Specific issues with STRs

- Typically comprised of tetra nucleotide repeats
- Range 70 - 450+ bp regions
- Longer STRs can be difficult to assemble based on read length
- Illumina GAIIx (read length 150 bp)
  - Generated 1000-2500 bp amplicons (13 core loci)
  - Problems detecting D21S11 32.2 and 34.2 alleles
  - Issues detecting D18S51
  - Custom informatics tools for assembling STRs

Assembling STRs

- Challenges
  - Repeating sequences (STRs) and read lengths
  - Sample amount requirements (10 ng to 5 µg)
  - Cost and time per unit of information
  - Data analysis (storage, assembly, interpretation)
  - Policy, privacy, disease related markers
  - Validation
  - Standards/reference materials
    - Nomenclature
    - Accuracy of sequence information
    - Errors, platform and bioinformatics-based bias
Next Generation Sequencing Workshop

- Interagency Workshop on the use of Next-Generation DNA Sequencing for Human Identification and Characterization (Jan 31 2012)
- Discussion of forensic applications of NGS (NIST, DoD, FBI, DHS) – materials can be found at:
- We are in the process of looking at platforms to characterize forensic markers (mitochondrial, STRs, SNPs)
- Evaluate accuracy, reproducibility, identify initial requirements for a NGS forensic reference material

Some Thoughts on the Future…

- PCR amplification
  - Faster enzymes to enable rapid PCR
  - More robust enzymes and master mixes to overcome inhibition
- Instrumentation
  - More dye colors to aid higher levels of multiplexing
  - Rapid, integrated devices
  - Alternatives to capillary electrophoresis: PLEX-ID and NGS
- Quantitative information
  - qPCR and digital PCR
- Marker systems
  - Expanding sets of STR loci for growing DNA databases
  - Other marker systems: SNPs, InDels, X-STRs, RM Y-STRs
  - Body fluid identification with mRNA, miRNA, and DNA methylation
  - Phenotyping for external visible characteristics
  - Challenges with potential whole genome information
- Data interpretation
  - Probabilistic genotyping for low-level DNA and mixture interpretation
  - Probability of dropout

Effective Training is Needed in All Areas

Improvements in Forensic DNA Analysis

- Biology
  - Improved DNA extraction with automation
  - New capabilities for recovery of information from degraded DNA samples (e.g., miniSTRs)
- Technology
  - Parallel processing of DNA with capillary arrays
  - Expert systems for automated data interpretation
- Genetics
  - Ethnicity estimations (with STRs and/or SNPs)
  - Larger Y-STR and mtDNA population databases

AAFS 2009 Topics Regarding Forensic DNA

- Improved DNA extraction
- Predicting hair color and ancestry with SNPs
- X-chromosome STRs
- Familial searching
- Y-STRs and mixtures
- Low level DNA samples
- miniSTRs
- DNA screening assays
- Optimizing database labs
- Microfluidic biochip systems
- Use with property crimes
- Recovery from handguns
- DNA from IEDs
- Expert systems
- Automation with robotics
- DNA quantitation – qPCR
- PCR directly from blood
- mtDNA
- RNA
- Non-human DNA (dogs & cows)
- Mixture interpretation

From abstracts of presentations at AAFS meeting in Denver, CO (Feb 2009)

www.DNA.gov Website

Summary of NIJ-Funded Research

http://www.dna.gov/research/
The DNA Field Moves Forward…

The Future
- More Robotics
- Expert Systems
- Animal & Plant DNA
- Physical Characteristics
- Ethnicity Estimation

The Future of Forensic DNA is Similar to the Olympic Motto of “Swifter, Higher, Stronger”

Acknowledgments
- Slides from Pete Vallone (NIST) and Manuel Fondevila (NIST, USC)
- Funding from National Institute of Justice and FBI Biometrics Center of Excellence for work performed within the NIST Applied Genetics Group

Contact Information
John Butler
NIST Fellow
Group Leader of Applied Genetics
john.butler@nist.gov
301-975-4049
http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm

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