Development of Rapid Multiplex PCR Amplification Techniques

Dr. Peter M. Vallone
National Institute of Standards and Technology
Biochemical Science Division
Applied Genetics Group - Human Identification Project

Lockheed Martin BEACON Lecture Series
January 7th, 2009

Outline

• NIST
• Forensic DNA Testing
• PCR
• Rapid PCR
  1. miniSTR 3plex
  2. Commercial Kits
  3. Larger Custom Multiplexes
  4. Alternative Thermal Cyclers
### NIST History and Mission

- **National Institute of Standards and Technology (NIST)** was created in 1901 as the National Bureau of Standards (NBS). The name was changed to NIST in 1988.
- **NIST** is part of the U.S. Department of Commerce with a mission to develop and promote measurement, standards, and technology to enhance productivity, facilitate trade, and improve the quality of life.
- NIST supplies over 1,300 Standard Reference Materials (SRMs) for industry, academia, and government use in calibration of measurements.
- NIST defines time for the U.S.

### NIST Human Identity Project Team

- **John Butler** (Project Leader)
- **Margaret Kline**
- **Pete Vallone**
- **Dave Duewer**
- **Jan Redman**
- **Amy Decker**
- **Becky Hill**

**Funding:** Interagency Agreement 2003-IJ-R-029 between National Institute of Justice (NIJ) and NIST Office of Law Enforcement Standards (OLES)
Current Areas of Effort with Forensic DNA

• Standards
  – Standard Reference Materials
  – Standard Information Resources (STRBase website)
  – Interlaboratory Studies

• Technology
  – Research programs in SNPs, miniSTRs, Y-STRs, mtDNA, qPCR, rapid PCR
  – Assay and software development, expert system review

• Training Materials
  – Review articles and workshops on STRs, CE, validation
  – PowerPoint and pdf files available for download

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

Standard Reference Materials (SRMs)
http://www.nist.gov/srm

Traceable standards to ensure accurate and comparable measurements between laboratories

SRM 2391b – autosomal STRs
SRM 2392 &-I – mtDNA sequencing
SRM 2395 – Y-STRs
SRM 2372 – DNA quantitation
SRM 2394 – mtDNA heteroplasmy

Calibration with SRMs enables confidence in comparisons of results between laboratories

Helps meet ISO 17025 needs for traceability to a national metrology institute
Support to the Forensic Community

...Bringing traceability and technology to the scales of justice...

- Perform beta-testing of new human identity testing products
- Provide input to
  - Scientific Working Group on DNA Analysis Methods (SWGDAM)
  - Department of Defense Quality Assurance Oversight Committee for DNA Analysis
  - American Prosecutor’s Research Institute (APRI) DNA Forensics Program “Course-in-a-Box” for training lawyers
  - WTC Kinship and Data Analysis Panel (KADAP)
  - 2005 Hurricane Victim DNA Identification Expert Group (HVDIEG)
  - NIJ Expert System Testbed (NEST) Project

Forensic Applications of DNA Typing

- Forensic cases: matching suspect with evidence
- Paternity testing: identifying father
- Missing persons investigations
- Military DNA “dog tag”
- Convicted felon DNA databases
- Mass disasters: putting pieces back together
- Historical investigations
- Genetic genealogy

>3 million tests performed per year
Forensic DNA Testing

Probe subsets of genetic variation in order to differentiate between individuals

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

Typically, we are not looking at genes – little/no information about race, predisposal to disease, or phenotypical information (eye color, height, hair color) is obtained

Short Tandem Repeat (STR) Markers

An accordion-like DNA sequence that occurs between genes

TCCCAAGCTTTCTTCCCTTCCTCCCTAGATCAATACAGACAGAAGACA
GGTGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAG
ACATGCTTACAGATGCACAC

= 12 GATA repeats ("12" is reported)

Target region (short tandem repeat)

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

The number of consecutive repeat units can vary between people

7 repeats
8 repeats
9 repeats
10 repeats
11 repeats
12 repeats
13 repeats
DNA Testing Requires a Reference Sample

A DNA profile by itself is fairly useless because it has no context…

DNA analysis for identity only works by comparison – **you need a reference sample**

<table>
<thead>
<tr>
<th>Crime Scene Evidence</th>
<th>compared to</th>
<th>Suspect(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child</td>
<td>compared to</td>
<td>Alleged Father</td>
</tr>
<tr>
<td>Victim’s Remains</td>
<td>compared to</td>
<td>Biological Relative</td>
</tr>
<tr>
<td>Soldier’s Remains</td>
<td>compared to</td>
<td>Direct Reference Sample</td>
</tr>
</tbody>
</table>

Steps in Forensic DNA Analysis

*Usually 1-2 day process (a minimum of ~8 hours)*

1. Sample Collection & Storage
2. DNA Extraction
3. DNA Quantitation
4. Multiplex PCR Amplification
5. DNA separation and sizing
6. STR Typing
7. Interpretation of Results
8. Statistics Calculated
   - DNA Database search
   - Paternity test
   - Reference sample
9. Applied Use of Information
Identifiler STR Kit (Applied Biosystems)
Information is tied together with multiplex PCR and data analysis

The Random Match Probability (RMP) is over 1 in 800 trillion for unrelated individuals
The chance of someone else having this exact same profile

Adenylation Efficiency or Non-Template Addition
Taq polymerase adds an extra base (dA) to the end of each PCR product
A ‘soak’ at 60°C for extended periods promotes full adenylation
PCR

- Polymerase Chain Reaction
- In vitro enzymatic replication
- Targets a specific region of a genome
- $2^N$ amplification ($N =$ number of cycles)
- 50 – 10,000 base pair fragments
- Products can be used for downstream applications

*A means to create billions of exact copies of a specific region of the genome*

---

**PCR Mechanism**

1. **95°C**  
   **30 – 60s**  
   Denature the genomic DNA template

2. **~60°C**  
   **1 – 2 min**  
   PCR primers bind to the DNA template; defines the product size
PCR Mechanism

Taq primer complex forms and the dNTPs are incorporated into the new strand(s)

$72^\circ\text{C}$
$1 - 2 \text{ min}$

Processivity rate of Taq complex $\sim$20 nucleotides/s

After 30 cycles $2^{30} = 1,073,741,824$ copies are produced
PCR Applications

• Microbiology and Molecular Biology
  – DNA cloning, Southern blotting, DNA sequencing, Next-gen sequencing, DNA methylation assays

• Genotyping
  – forensics, pathogen detection, clinical and diagnostic applications, disease association studies, pharmacogenetics

• Real-time PCR
  – RT-PCR (reverse transcriptase) gene expression
  – Quantitation (qPCR), Genetically mod organisms

Advantages of Multiplex PCR?

• Beneficial with limited sample
  – 500 to 1000 pg of DNA
  – ~83 to 166 copies of the human genome

• Obtain more information per unit time
• Save on reagents; enzyme, buffers, labor
• Streamlines data analysis
• For forensic markers it is essential
• Coincides with high capacity instrumentation
Things That Are Rapid

3h PCR?

Why go Faster?
Applications for Rapid PCR

• Integrated devices (‘Lab on a Chip’)

• Screening at a point of interest (airport, border, crime scene, intelligence community)

• Rapid STR typing ‘in the field’
  – Potential for situations/cases when a quick result is needed
  – Provide initial screening information

• Decrease overall time required for STR typing

Growing Interest in DNA for Biometrics
DNA within the Biometric Model

Creating the reference sample...
Present Biometric → Capture → Process → Store

Testing the “evidence”...
Present Biometric → Capture → Process

Compare

Match of 13 points (each with 2 variable alleles) within DNA

String of 26 numbers (order of listing DNA results would have to be standardized)
16, 17-17, 18-21, 22-12, 14-28, 30-14, 16-12, 13-11, 14-9, 9-11, 13-6, 6-8, 8-10, 10

Match
Deny Entry
“Exonerated”
No Match
“Implicated”
Permit Entry

Efforts towards Portable/Mobile DNA Devices

- NEC (Japan)
  - Poster at Promega meeting in Hollywood, CA (Oct 1-4, 2007)
  - Press release on October 15, 2007
- Mathies group at UC-Berkeley and Microchip Biotech
  - Publications… in Analytical Chemistry, FSI Genetics, etc.
  - http://www.microchipbiotech.com
- Landers group at UVA and MicroLab Diagnostics
  - Publications… Proc Natl Acad Sci USA 2006; 103:19272-19277
  - http://www.microlabdiagnostics.com
- Network Biosystems (based on Dan Ehrlich’s work at Whitehead)
  - http://www.netbio.com
- NIST Michael Gaitan

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Field Trial Results
7:00 a.m. Arrived and set up mock crime scene
7:10-7:30 a.m. Samples collected by CSI
7:30-9:30 a.m. DNA extraction
9:30-10:00 a.m. PCR set-up
10 a.m. – 12 p.m. PCR performed (2 h)
12 – 12:30 p.m. DNA separation
12:30-12:50 p.m. CODIS search of local database

6 hours from sample collection to the generation of the CODIS hit (for one sample)
In the months ahead, we will also challenge the private sector to speed up DNA fingerprinting so that when DNA is left behind, officers can identify suspects more quickly and avoid wrongful arrests. And to do this, we will establish a six-figure prize for anyone who can invent a device tailored to the NYPD that analyzes DNA right at the crime scene. It's just one more way we are trying to bring private sector innovation into the public sector.

Typical STR Typing Workflow

*Can the time required for PCR thermal cycling be reduced?*

- Sample Extraction ~2 h
- Quantitation ~1.5 h
- PCR ~3 h
- CE Run ~1.5 h
- Data Review ?

Alter thermal cycling parameters
Evaluate faster polymerases
Evaluate faster thermal cyclers
Test commercial STR typing kits

Goal: cycling in less than 45 minutes
Trying simple things first…
**Thermal Cyclers**

- **RapidCycler 2 Instrument**
  - Ramp rate = 10°C/sec
  - http://www.idahotech.com

- **AB 9800**
  - http://www.appliedbiosystems.com
  - Ramp rate = 4°C/sec

- **Eppendorf Mastercycler ep**
  - Ramp rate = 6°C/sec
  - http://www.eppendorfnla.com

- **AB 9700**
  - 96 well block
  - 0.2 μL tubes
  - http://www.appliedbiosystems.com

---

**PCR Thermal Cycling Profile (Traditional)**

- 28 - 32 cycles of PCR
- ~3 hour run time

- 95°C 10 min
- 95°C 1 min
- 58°C 1 min
- 72°C 1 min
- 60°C 60 min

Heating and cooling of reaction takes place in a thermal block (peltier heating/cooling)
Heating/Cooling rate of ~1°C/s
**PCR Thermal Cycling Profile**

28-32 cycles of PCR

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Unit</th>
<th>Trad</th>
<th>Rapid</th>
<th>Difference (min)</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hot Start</td>
<td>Min</td>
<td>10</td>
<td>1</td>
<td>9.0</td>
<td>6.3</td>
</tr>
<tr>
<td>Hold</td>
<td>Sec</td>
<td>60</td>
<td>5/10</td>
<td>72.3</td>
<td>50.6</td>
</tr>
<tr>
<td>Soak</td>
<td>Min</td>
<td>60</td>
<td>1</td>
<td>59.0</td>
<td>41.2</td>
</tr>
<tr>
<td>Ramp rate</td>
<td>(deg/sec)</td>
<td>1</td>
<td>4</td>
<td>22.4</td>
<td>15.7</td>
</tr>
</tbody>
</table>

- Cycles: 28
- Time: 2:58:41
- Difference: 2:23:03

Heating/Cooling rate of ~4°C/s

---

**Thermal Cycling**

Evaluate robustness and reproducibility (electropherograms)
Comparison of Thermal Cycling Profiles

Rapid cycling is completed by the ~4th cycle of the traditional cycling profile

Initial Work/Assumptions

- Using common materials/conditions
  - AB 9700 (10 μL volume)
  - Standard plastics
  - Commercial Polymerases
  - Final primer concentration ~0.2 μM
  - ~250 μM dNTPs, 2 mM Mg++
  - 4 + 1 color dye chemistry for labeling primers
  - Separation on AB 3130 (Capillary Electrophoresis)
  - Not sample limited (>500 pg of DNA)
Loci for Testing

- STR Loci present available in commercial kits (CODIS)

- 26 autosomal loci characterized in our laboratory
  – Small 3plex panels
  – Larger 26plex

- Existing commercial STR typing kits are not optimized for rapid PCR

- Challenge for miniaturized/integrated STR typing platforms – since they have to use a commercial kits or develop their own...


DNA Polymerases for Evaluation

<table>
<thead>
<tr>
<th>Polymerase</th>
<th>Vendor</th>
<th>MasterMix</th>
<th>Hot Start</th>
</tr>
</thead>
<tbody>
<tr>
<td>TaqGold</td>
<td>Applied Biosystems</td>
<td>no</td>
<td>10 min</td>
</tr>
<tr>
<td>GeneAmp</td>
<td>Applied Biosystems</td>
<td>yes (2x)</td>
<td>1 min</td>
</tr>
<tr>
<td>SpeedSTAR</td>
<td>Takara</td>
<td>no</td>
<td>1 min</td>
</tr>
<tr>
<td>PyroStart</td>
<td>Fermentas</td>
<td>yes (2x)</td>
<td>1 min</td>
</tr>
<tr>
<td>Qiagen Fast Cycling</td>
<td>Qiagen</td>
<td>yes (2x)</td>
<td>5 min</td>
</tr>
</tbody>
</table>

Brief survey of ‘fast’ commercial polymerases

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Initial Testing with miniSTR 3plex

- 3 loci labeled with 3 fluorescent dyes (FAM, VIC, NED)

- MiniSTR loci (Amplicon size range 65-140 bp)
  - D2S441, D10S1248 and D22S1045
  - ‘European loci’ (contained in the next generation of forensic kits)

- These loci were previously tested under standard cycling conditions in a miniSTR multiplex in our lab

miniSTR 3plex Standard Cycling

750 pg template, 28 cycles, TaqGold, ~3 hour cycling time
Initial Rapid Testing Results

- TaqGold
- Qiagen
- SpeedSTAR
- PyroStart
- GeneAmp

- A
- adeylation

D2S441
D22S1045
D10S1248

Improved adeylation

D22 and D10 not detected

w/o 9800 and specific plastics

Rapid PCR 28 cycles in 36 minutes

D2S441
D10S1248

Peak Balance?

750 pg template, 28 cycles, PyroStart, 36 minute cycling time

D22S1045
Adenylation…
miniSTR 3plex Concordance

• 3plex run on a plate of samples (n = 95)

• Concordance was checked with genotypes obtained with Standard Cycling and TaqGold

• 2/285 (0.7%) of the genotype calls were discordant

• Both cases due to D10S1248 heterozygote peak height imbalance
Standard Cycling
PyroStart
PHR = 0.09

Rapid Cycling
PyroStart + 1 U SpeedSTAR
PHR = 0.94

Peak Height Imbalance for D10S1248

- Standard Cycling (3h) did not solve the problem
- 18 allele was successfully amplified with addition of SpeedSTAR polymerase

Peak Height Ratios for 16 Samples

<table>
<thead>
<tr>
<th>Cycling Sample Name</th>
<th>Normal TagGold</th>
<th>Rapid Pyro</th>
<th>Rapid Pyro+SS</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>MT94859</td>
<td>0.70</td>
<td>0.87</td>
<td>14,19</td>
<td>14,18</td>
</tr>
<tr>
<td>PT84230</td>
<td>0.68</td>
<td>0.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PT84243</td>
<td>0.63</td>
<td>0.28</td>
<td>0.73</td>
<td>14,17</td>
</tr>
<tr>
<td>PT84230</td>
<td>0.66</td>
<td>0.30</td>
<td>0.69</td>
<td>14,17</td>
</tr>
<tr>
<td>WT51354</td>
<td>0.67</td>
<td>0.37</td>
<td>0.79</td>
<td>13,16</td>
</tr>
<tr>
<td>MT97172</td>
<td>0.70</td>
<td>0.37</td>
<td>0.87</td>
<td>13,16</td>
</tr>
<tr>
<td>WT51342</td>
<td>0.71</td>
<td>0.40</td>
<td>0.99</td>
<td>13,16</td>
</tr>
<tr>
<td>WT51355</td>
<td>0.73</td>
<td>0.41</td>
<td>0.88</td>
<td>13,16</td>
</tr>
<tr>
<td>ZT80865</td>
<td>0.74</td>
<td>0.41</td>
<td>0.91</td>
<td>13,16</td>
</tr>
<tr>
<td>UT57310</td>
<td>0.75</td>
<td>0.42</td>
<td>0.88</td>
<td>14,16</td>
</tr>
<tr>
<td>PT84242</td>
<td>0.78</td>
<td>0.42</td>
<td>0.95</td>
<td>12,16</td>
</tr>
<tr>
<td>PT84241</td>
<td>0.78</td>
<td>0.46</td>
<td>0.77</td>
<td>13,16</td>
</tr>
<tr>
<td>GT37862</td>
<td>0.78</td>
<td>0.47</td>
<td>0.87</td>
<td>13,16</td>
</tr>
<tr>
<td>WT51362</td>
<td>0.78</td>
<td>0.50</td>
<td>0.85</td>
<td>14,16</td>
</tr>
<tr>
<td>ZT80863</td>
<td>0.81</td>
<td>0.51</td>
<td>0.90</td>
<td>12,15</td>
</tr>
</tbody>
</table>

- 2 samples were typed as "homozygous"
- 16 samples with lowest PHR values were amplified with extra polymerase
- Balance was improved with the SpeedStar polymerase

Samples with larger allele spreads for D10S1248 exhibited greater imbalance e.g. 14,16 better balance than 14,19
Peak Height Ratios for 3plex Loci

D10S1248 Peak Imbalance

- Imbalance is not solely related to amplicon size
- Improved with additional SpeedStar polymerase
- Not an artifact of rapid thermal cycling conditions
- Other reasons
  - Repeat motif?
  - Primer $T_m$?
Testing Commercials Kits

- Tested various commercial STR kits
  - 10 μL volume
  - 2 μL primer mix (commercial)
    - PyroStart +1 U SpeedSTAR polymerase
  - 1 ng of template DNA
  - 28 cycles (rapid cycling parameters)
  - 36 min

Promega Corporation → PowerPlex16
Applied Biosystems → Identifiler

Identifiler Rapid Cycling

*without SpeedStar polymerase*

6/16 loci exhibited relatively low signal
All 16 loci amplified, good balance, strong signal (except D19, D21)
All 16 loci amplified, good balance, strong signal, efficient adenylation

Rapid PCR Article

Complete concordance of STR allele calls (for 60 samples) between the rapid and standard thermal cycling protocols were observed although there was incomplete adenylation at several of the loci examined and some PCR artifacts were detected. Using less than 750 pg of template DNA and 28 cycles, STR peaks for all loci were above a 150 relative fluorescent unit (RFU) detection threshold with fully adequate inter-locus balance and heterozygote peak height ratios of greater than 0.84.
Further Evaluation of NIST Loci

- We currently have an autosomal 26plex assay working in our lab
- Amplified (25/26) each locus in singleplex under rapid cycling conditions
- Evaluate each locus for signal intensity, full adenylation and non-specific artifacts
- Rank and test candidate loci in a rapid multiplex

Autosomal STR 26plex

Singleplex Evaluation

Amplified using the rapid PCR protocol

4/25 peak balance?
3/25 below 1000 RFUs

17 of the 25 loci pass the initial singleplex screen

Testing 4 Multiplexes

• After singleplex evaluation 4 multiplexes were tested (empirical balancing)
  – 17plex
  – 14plex
  – 12plex
  – 7plex

  Subset of the 17plex

• Run under rapid cycling conditions
• 1 ng DNA, 28 cycles, PyroStart + 1 U SpeedSTAR
Rapid Assays Developed Using NIST Loci

- N = 16 samples
- D4S2364 adenylation issues/artifacts
- D9S2157 severe peak imbalance – allele drop out in 2 samples
- Further evidence that peak imbalance does not directly track with amplicon size
- ‘Troublesome loci’ can be screened out

Cepheid SmartCycler and Stratagene RoboCyler 96

- Working with Dr. Daniele Podini (GWU)
  - Ms. Michelle Burns (NIST/GWU)
- Identifiler with rapid PCR protocols
  - Increased ramp rate
  - Shorter hold times
  - Testing other fast polymerases
  - Improved thermal transfer unique to the SmartCycler cell design
Cepheid SmartCycler

- 16 independent ‘cells’ per block
- Heat 10 °C/s
- Cool 2.5 °C/s

http://www.cepheid.com/systems-and-software/smartcycler-system/

RoboCycler 96

- Robotic arm moves sample tubes move from station to station
- No temperature ramping

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
**SmartCycler 5s/10s/10s**

25 minutes

10 μL reaction

Low signal for D21 and D19 loci

Optimized polymerase cocktail

Improved signal for D21 and D19 loci

1 ng DNA
AB 9700 5s/10s/10s
36 minutes

Improved signal for D21 and D19 loci
Optimized polymerase cocktail

AB 9700 1s/1s/1s
29 minutes

Optimized polymerase cocktail
Final Conclusions

- Fast multiplex PCR amplification is possible
  - Compatible with commercial STR typing kits
  - Provides same genotypes as standard cycling
  - Some artifacts, signal imbalance, poor performing loci
- Fast (optimized) polymerases are needed
- Further work
  - Applying techniques to integrated platforms
  - Formal validation of technique
  - Sharing results with PCR community
  - Understanding the kinetics of PCR
Acknowledgements

John Butler
(Project Leader)

Becky Hill

Margaret Kline

Dr. Daniele Podini
Ms. Michelle Burns
(GWU)

Funding: Interagency Agreement 2003-IJ-R-029 between National Institute of Justice (NIJ) and NIST Office of Law Enforcement Standards (OLES)