Multiplexed Assays for Probing Y Chromosome and Mitochondrial Markers

5th Annual DNA Forensics Meeting
June 27-28, 2002
Peter M. Vallone
National Institute of Standards and Technology

Outline of Presentation

Multiplexing
Assays and Instrumentation
Y Chromosome and Mitochondrial DNA
Primer design strategy
Results
  mtSNP 10 plex
  Y SNP 5 plex
  Y STR multiplexes
Multiplexing

Probing multiple loci/markers simultaneously

Multiple strands of short DNA bind to template DNA and chemistry occurs (PCR, primer extension, hybridization, etc)

What are the Advantages of Multiplexing?

Obtain more information per unit time
Reduce the amount of limited forensic sample used
Save on reagents; enzyme, buffers, DNA oligomers
Reduces labor
Streamlines data analysis
For certain markers it is essential (SNPs, YSTRs)
Coincides with high capacity instrumentation
What are the Challenges of Multiplexing?

- Only guidelines exist for designing multiplexes
- More markers = increased complexity
- Testing a robust multiplex
- Inclusion of useful markers in the multiplex

What Type of Genetic Variation?

- **Length Variation**
  - short tandem repeats (STRs)
  - CTAGTCG\textit{(GATA)(GATA)(GATA)}GCGATCGT

- **Sequence Variation**
  - single nucleotide polymorphisms (SNPs)
  - insertions/deletions
  - GCTAGTCGATGCTC\textit{(G/A)}GCGTATGCTGTAGC
What Assays are we Multiplexing?

Polymerase chain reaction (PCR)
   Amplification of specific region of the human genome
   Typically used for STR and SNP
   Use Capillary Electrophoresis for detection

Primer Extension reaction (minisequencing)
   Typically used for SNP markers
   Use Capillary Electrophoresis and Mass Spectrometry for detection

NIST Goals for Multiplex Assays

Working with collaborators who have markers of forensic interest

By using our multiplex assays collaborators can type markers and evaluate forensic utility

Further understanding of multiplex assays

Publish assay details for others to evaluate (commercial and research)
Multiplexing

Assays and Instrumentation

Y Chromosome and Mitochondrial DNA

Primer design strategy

Results

- mtSNP 10 plex
- Y SNP 5 plex
- Y STR multiplexes

Multiplex PCR

Multiple primer pairs target more than one specific site on the DNA strand

Compatible primers are the key to successful multiplex PCR

Commercial kits are available for targeting and simultaneously amplifying 16 markers
Capillary Electrophoresis Instrumentation

**ABI 310**
single capillary

**ABI 3100**
16-capillary array

---

ABI PRISM® SNaPshot™ Multiplex System

**Primer extension** assay that utilizes fluorescently labeled ddNTPs

Analysis of fragment size and fluorescent label identity by CE allows typing of multiple SNPs

Multiplexed amplicons or pooled singleplex PCR amplicons can be used as templates

**Primer design must be done by user!**
Primer Extension with SNaPshot™

SNP Primer is extended by one base unit

“tail” used to vary electrophoretic mobility

Fluorescently labeled ddNTPs +

Oligonucleotide primer 18-28 bases

PCR Amplified DNA Template

<table>
<thead>
<tr>
<th>ddNTP</th>
<th>Dye label</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>dR6G</td>
<td>Green</td>
</tr>
<tr>
<td>C</td>
<td>dTAMRA</td>
<td>Black</td>
</tr>
<tr>
<td>G</td>
<td>dR110</td>
<td>Blue</td>
</tr>
<tr>
<td>T</td>
<td>dROX</td>
<td>Red</td>
</tr>
</tbody>
</table>

25 Cycles
96°C 10s  
50°C 5s  
60°C 30s

Detection of SNPs with ABI 310/3100

SNaPshot™ CEPH Control Reaction

Multiplexing possible by use of different length primers

Poly(T) tail or non-nucleotide linker to aid separation

Priming site
### Primer Extension for MALDI TOF MS Analysis

**PCR Amplified DNA Template**

- **Oligonucleotide primer 18-28 bases**
  - 5’
  - 3’
  - SNP

**SNP Primer is extended by one base unit**

- **Natural non-labeled ddNTPs**
- **polymerase**

<table>
<thead>
<tr>
<th>ddNTP</th>
<th>Mass (Da)</th>
<th>40 Cycles</th>
<th>Mass difference between SNP primer and single base extension product provides genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>297</td>
<td>96°C 10s</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>273</td>
<td>50°C 20s</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>313</td>
<td>72°C 30s</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>288</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Time-of-Flight Mass Spectrometry (TOF-MS)**

- **DNA Reaction Products** (Size separated and drifting to the detector)
- **Pulsed Laser Beam**
- **High-Density Sample Array**
- **Ion Extractor**
- **Acceleration Region (20 kV)**
- **Drift Region Electric-Field Free**
- **Detector**
- **X-Y sample control**
Multiplexing
Assays and Instrumentation
Y Chromosome and Mitochondrial DNA
Primer design strategy
Results
- mtSNP 10 plex
- Y SNP 5 plex
- Y STR multiplexes

MALDI-TOF MS data can be collected in 5-10 seconds
Markers of Interest

• Mitochondrial DNA (mtDNA)
  – maternally inherited
  – polymorphic control region (D-loop)
  – ~1000’s of copies per cell

• Y chromosome
  – paternally inherited
  – variety of Y STR and Y SNP markers
  – *haplotype rather than genotype*

Require large databases because recombination does not occur

Multiplexing
Assays and Instrumentation
Y Chromosome and Mitochondrial DNA

Primer design strategy

Results
  mtSNP 10 plex
  Y SNP 5 plex
  Y STR multiplexes
Multiplex PCR Primer Selection

Identify markers of interest (collaborations, literature, research)

Obtain reference sequences containing the sites of interest (Genbank) with approximately 500 bases of sequence information upstream and downstream of the marker

Decide upon a desired PCR product size
- Short amplicons for degraded samples, SNPs
- Longer amplicons for STRs

Use software for selecting singleplex primer pairs

Primer3
www-genome.wi.mit.edu/genome_software/other/primer3.html

Multiplex PCR Design

Select singleplex PCR primers for each amplicon using Primer 3 software

```
[OLIGO]

LEFT PRIMER
sequence: 27 20 60.96 50.00 4.00 2.00 CAGTAAACAGCAATTCTA

RIGHT PRIMER
sequence: 174 32 90.91 50.00 6.00 3.00 CGCTCTGAAACTGATACGCT

SEQUENCE SIZE: 205
INCLUDED REGION SIZE: 205

PRODUCT SIZE: 140, PAIR ANY COMPL: 3.00, PAIR 3' COMPL: 2.00
EXCLUDED REGIONS (start, len): 70,58

1 CTTGACCAACGGAACCAAGTTAACCTAGGTTAAGGAACTACTCTCTCTGAC

61 TCAAAGATAGGGTTTTCAGACCTGTTTGGATCAAGGACATCCATUGGTGCAAGUGCT

121 TAAAAGCTCTCTCCTCAAGCTTAAAGCTCTACTGCTCTCTGACGAGAT

101 TTCAGGTCGCTCTGACCGCACTTC
```
Stand Alone Primer3

Sending multiple sequences over the web for primer selection can be tedious

The Primer3 web output is fine for the screen viewing or printing but not for organizing in spreadsheets

Primer3 is publicly available and can be run (in batch!) on a Unix, Linux, or Mac (OSX) computer

Developed a program that formats files for Primer3 input

Reference sequences that are stored in Excel can be quickly formatted for Primer3

Example input format for Primer3

```
PRIMER_SEQUENCE_ID=M9
SEQUENCE=GCAGCATATAAAACTTTCAGGACCCTGAAATACAGAACTG
   CAAAGAAACGCGCTAAGATGTTGAATNCCTTTTATTATTTTTTTTTTATTTAG
   ACATGTTCAAGTGCTCAGGATGTTACATACTTTTTTTGTTATGTAAGTAAGTTAG
   CGCCTACTCATTAGCATTTCAATCTCAAATTCCTTTGTGAAAT
   GTTGAAATATTTTTTCTAATCTGTTTCACGAGCTCATCTCAAAAAATGAGGAAT
   GATTGCAATTTACAGCATTTTGCAGAAAATGGCCTTTTTTTTTTTTATCGATTTATTT
   ACTTAAACATTACGTACATTACGCTAGCAAAGATTTAA
PRIMER_COMMENT=(340 bp); G to C at position 68
PRIMER_LIBRARY=/Users/vallone/Desktop/primer3/misprM9
PRIMER_MAX_MISPRIMING=8
PRIMER_PAIR_MAX_MISPRIMING=20
EXCLUDED_REGION=38,60
PRIMER_PRODUCT_SIZE_RANGE=90-150
PRIMER_PRODUCT_OPT_SIZE=105
PRIMER_MIN_SIZE=18
PRIMER_MAX_SIZE=27
PRIMER_OPT_SIZE=20
PRIMER_OPT_TM=60
PRIMER_MIN_TM=57
PRIMER_MAX_TM=63
PRIMER_NUM_RETURN=1
PRIMER_EXPLAIN_FLAG=1
PRIMER_EXPLAIN_FLAG=1
```

```
PRIMER_SEQUENCE_ID=M42
SEQUENCE=AAAGCGAGAGATTCAATCCAGGATGACAGAATGCGTTCAGCTTAAAGGGATTAAAAGAAGTATAATACAGTCTGTATTATTAGATCACCCAGAGACACACAAAACAAGAACCGTGAATTGAATTAGTGGTATACTAATAGAGTGGTTTTACCTGAAATATTTACACATCAATCCTACTGAATTCTTACAAC
```

Can help utilize all the tools that Primer3 provides

Non-Specific Interactions

Primers that interact with non-specific (undesired) regions of a genome OR with each other can degrade PCR performance

Screening for alternate genomic binding regions can be accomplished using BLAST http://www.ncbi.nlm.nih.gov

Screening for potential primer-dimer interactions is accomplished using in house software - AutoDimer
AutoDimer Check

Screening for potential intramolecular hairpin and intermolecular primer-dimer formation

PCR Assay Design
If primer pairs meet criteria

Obtain primer pairs and test singleplex PCR
(QC all primers with MS/CE/HPLC)

Begin initial testing of multiplex PCR
Start with a PCR mix containing
1.0 µM of each primer pair

Evaluate amplicon yields, presence and balance

Vary primer pair concentrations, [polymerase], number of cycles, [Mg++] , [dNTPs], BSA

Redesign and retest failing loci
Multiplexing
Assays and Instrumentation
Y Chromosome and Mitochondrial DNA
Primer design strategy

Results

mtSNP 10 plex
Y SNP 5 plex
Y STR multiplexes

Coding Region mtSNP 10-plex Assay

Multiplex primer extension with different length SNP primers and fluorescent ddNTPs

mtDNA control region

mtDNA coding region

Collaboration with Tom Parsons (AFDIL) provided sequence variation sites located in the coding region

PCR product sizes kept under 150 bp to enable success with degraded DNA samples

Multiplex PCR used to co-amplify all regions of interest at once
Tailed SNP primers allows for multiplexing in the SNaPshot assay

Sequences for 10 SNP primers

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCAGAAGTGAAAGGGGC</td>
<td>18/na</td>
</tr>
<tr>
<td>TTTTTTTTTGTGGATCAGGACATCCC</td>
<td>19/26</td>
</tr>
<tr>
<td>TTTTTTTTTTTACTAAGAAGAATTTTAGGA</td>
<td>20/30</td>
</tr>
<tr>
<td>TTTTTTTTTTTTTGACCACAGCTACGAAAATC</td>
<td>20/34</td>
</tr>
<tr>
<td>TTTTTTTTTTTTTTTGACGACTACGTGTTAGC</td>
<td>20/38</td>
</tr>
<tr>
<td>TTTTTTTTTTTTTTTTTCCACACACTTCTCGGCT</td>
<td>20/42</td>
</tr>
<tr>
<td>TTTTTTTTTTTTTTTTTTTTGGGGCTATTTAGGCTATG</td>
<td>22/46</td>
</tr>
<tr>
<td>TTTTTTTTTTTTTTTTTTTTTGCAGCCATTCAAGCAATCTCTATA</td>
<td>23/50</td>
</tr>
<tr>
<td>TTTTTTTTTTTTTTTTTTTTTTGGTTAGAAGTTGGAATAAAAAGCTAG</td>
<td>25/64</td>
</tr>
<tr>
<td>TTTTTTTTTTTTTTTTTTTTTTTTTTGACCATAACCAATACTACCAATCA</td>
<td>25/58</td>
</tr>
</tbody>
</table>

Template binding sequence – black
Tailed sequence for fragment separation - red

mtSNP 10-plex run on ABI 3100
(SNapShot™ assay)

Multiplex PCR and Multiplex SNP Detection

mtDNA coding region SNPs

Sizing performed by comparison to GS120 LIZ internal size standard (not shown)
Data for 10 mt Samples

<table>
<thead>
<tr>
<th></th>
<th>mtSNP1</th>
<th>mtSNP2</th>
<th>mtSNP3</th>
<th>mtSNP4</th>
<th>mtSNP5</th>
<th>mtSNP6</th>
<th>mtSNP7</th>
<th>mtSNP8</th>
<th>mtSNP9</th>
<th>mtSNP10</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>G</td>
<td>A</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>T</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>T</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>Q</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>G</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>T</td>
<td>T</td>
<td>A</td>
<td>C</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>O</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>T</td>
<td>T</td>
<td>A</td>
<td>C</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>C</td>
<td>T</td>
</tr>
<tr>
<td>I</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>A</td>
<td>C</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>C</td>
<td>T</td>
</tr>
<tr>
<td>J</td>
<td>A</td>
<td>O</td>
<td>A</td>
<td>A</td>
<td>T</td>
<td>T</td>
<td>A</td>
<td>C</td>
<td>C</td>
<td>T</td>
</tr>
</tbody>
</table>

In this set of 10 samples each possible allele is represented.

16 Test Samples with mtSNP 10-plex on ABI 3100
Multiplexing
Assays and Instrumentation
Y Chromosome and Mitochondrial DNA
Primer design strategy

Results

- mtSNP 10 plex
- Y SNP 5 plex
- Y STR multiplexes

There is a growing interest in the Y-chromosome to aid forensic, paternity, and missing persons testing...

245 Y SNPs typed
74 males (YCC cell lines)
153 haplogroups observed

This paper unifies previous haplogroup nomenclatures

Primers and other information for all 245 markers are included in supplementary material

5 Y SNPs typed define major haplogroups

Multiplex PCR with Y-Chromosome SNP Markers

Rapid CE Separation and Quantitation of Multiplex PCR Products

Intercalating dye used to fluorescently label amplicons

Multiplexed Y SNPs (5-plex) Analyzed by TOF-MS

Multiplexing is possible by using primers with non-overlapping masses.

Y SNP Results with SNaPshot Assay

Poly(T) tails used to space Y SNP alleles

Room for additional Y SNP markers

Data obtained by Gordon Spangler (graduate student at American University)
Y SNP Haplotypes for 16 Test Samples

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>M9 (C/G)</th>
<th>M42 (A/T)</th>
<th>M45 (G/A)</th>
<th>M89 (C/T)</th>
<th>M96 (G/C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male 1</td>
<td>G</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>C</td>
</tr>
<tr>
<td>Male 2</td>
<td>G</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>C</td>
</tr>
<tr>
<td>Male 3</td>
<td>G</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>C</td>
</tr>
<tr>
<td>Male 4</td>
<td>G</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>C</td>
</tr>
<tr>
<td>Male 5</td>
<td>G</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>C</td>
</tr>
<tr>
<td>Male 6</td>
<td>G</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>C</td>
</tr>
<tr>
<td>Male 7</td>
<td>G</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>C</td>
</tr>
<tr>
<td>Male 8</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>G</td>
</tr>
<tr>
<td>Male 9</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>G</td>
</tr>
<tr>
<td>Male 10</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>G</td>
</tr>
<tr>
<td>Male 11</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>G</td>
</tr>
<tr>
<td>Male 12</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>G</td>
</tr>
<tr>
<td>Male 13</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>G</td>
</tr>
<tr>
<td>Male 14</td>
<td>G</td>
<td>T</td>
<td>G</td>
<td>T</td>
<td>C</td>
</tr>
<tr>
<td>Male 15</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>T</td>
<td>C</td>
</tr>
<tr>
<td>Female</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Multiplexing
Assays and Instrumentation
Y Chromosome and Mitochondrial DNA
Primer design strategy

Results

- mtSNP 10 plex
- Y SNP 5 plex
- Y STR multiplexes
Schematic of Loci in NIST Y STR 20-plex

Used 5-dyes

6FAM (blue)

VIC (green)

NED (yellow)

PET (red)

LIZ (orange)

LIZ GS500-internal size standard

Same dyes as new ABI Identifier™ Kits

Utilizes 5-dyes

European "extended haplotype"

Collaboration with Mike Hammer and Alan Redd (U. Arizona)

Y STR 20-plex Assay
Y STR 20-plex & Y STR 11-plex

PCR Product Size (bp)


Y STR 9-plex – Sensitivity Study

DNA Amounts
75pg (28 cycles)
75 pg (32 cycles)
60 pg (28 cycles)
60 pg (32 cycles)

Failed Amplification at 28 Cycles
All Loci Detected at 32 Cycles
High-throughput Y STR Typing on the ABI 3100 (16-capillary array)

7680 data points in 24 hours using Y STR 20plex

Acknowledgments

Funding:
U.S. National Institute of Justice Grant #97-LB-VX-0003
Interagency Agreement between NIJ and NIST Office of Law Enforcement Standards

Collaborators
Thomas Parsons and Mike Coble (AFDIL)
Mike Hammer and Alan Redd (U. of Arizona)

petev@nist.gov