Forensic DNA Typing: Application to Mass Disaster Investigations, Paternity Testing and Human Identification

Dr. Peter M. Vallone

*NIST Biotechnology Division*

Methods for Human Identification

Fingerprints have been used since 1901

DNA since 1986
Forensic DNA Testing

The genome of each individual is unique (with the exception of identical twins)

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

Applications for Human Identity Testing

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

As DNA analysis has shown its usefulness, the number of samples gathered for testing purposes has gone up dramatically...
Anthrax Detection

• Bacterial samples grown in culture
• If turbid after 6 hours, they are plated out and a fluorescent antibody test is done (3 hr)
• DNA detection with TaqMan assay (2-2.5 hr)
  – 5 probes performed in duplicate
  – RAPID PCR cycling
• If positive for PCR, then redone at 24 hours

Numerous samples to analyze; public pressure to get results quickly; accuracy better be good…

Sources of Biological Evidence

• Blood
• Semen
• Saliva
• Urine
• Hair
• Teeth
• Bone
• Tissue
Only a very small amount of blood is needed to obtain a DNA profile.
DNA in the Cell

- Double stranded DNA molecule
- ~3 billion total base pairs
- 22 pairs + XX or XY

Target Region for PCR

What Type of Genetic Variation?

- Length Variation
  - short tandem repeats (STRs)
  - CTAGTCGT\((GATA)(GATA)(GATA)\)GCGATCGT

- Sequence Variation
  - single nucleotide polymorphisms (SNPs)
  - insertions/deletions
  - GCTAGTCGATGCTC\((G/A)\)GCGTATGCTGTAGC
Basic Concepts

**PCR polymerase chain reaction** – method of amplifying a specific region of the genome – go from 1 to over a billion copies in about 2 hours

**Locus** region of the genome being examined

**Allele** the state of the genetic variation being examined

- (STRs = number of repeat units)
- (SNPs = base sequence at the site)

Chromosomes are paired so...

- **Homozygous** – Alleles are identical on each chromosome
- **Heterozygous** - Alleles differ on each on each chromosome
PCR Copies DNA Exponentially through Multiple Thermal Cycles

In 30 cycles at 100% efficiency, 1.07 billion copies of targeted DNA region are created

Short Tandem Repeats (STRs)

the repeat region is variable between samples while the flanking regions where PCR primers bind are constant

Homozygote = both alleles are the same length

Heterozygote = alleles differ and can be resolved from one another
Capillary Electrophoresis System

Butler, J.M. (2001) Forensic DNA Typing, Figure 9.3, ©Academic Press

Principles of Sample Separation and Detection

Butler, J.M. (2001) Forensic DNA Typing, Figure 10.8, ©Academic Press
Dr. Peter M. Vallone
Allelic Ladders

<table>
<thead>
<tr>
<th>Sample #1</th>
<th>Sample #2</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMEL</td>
<td>AMEL</td>
</tr>
<tr>
<td>D3S1358</td>
<td>D3S1358</td>
</tr>
<tr>
<td>TH01</td>
<td>TH01</td>
</tr>
<tr>
<td>D21S11</td>
<td>D21S11</td>
</tr>
<tr>
<td>D18S51</td>
<td>D18S51</td>
</tr>
<tr>
<td>Penta D</td>
<td>Penta D</td>
</tr>
<tr>
<td>Penta E</td>
<td>Penta E</td>
</tr>
</tbody>
</table>

PCR Product Size (bp)

- Allelic Ladders
- Sample #1
- Sample #2

All heterozygous
Multiplex PCR
(Parallel Sample Processing)

- Multiple primers target more than one site on the DNA strand
- Commercial kits are available for targeting and simultaneously amplifying 15 STR markers
- Spectrally distinguishable fluorescent dyes are used as labels

Advantages of Multiplex PCR
- Increases information obtained per unit time (increases power of discrimination)
- Reduces labor to obtain results
- Reduces template required (smaller sample consumed)

Methods for Parallel Sample Processing

- Multiplex by Size
- Multiplex by Dye Color
- Multiplex by Number of Capillaries
High-Throughput STR Typing on the ABI 3100 (16-capillary array)

256 data points in 45 minutes with STR 16plex and 16 capillaries

Human Identity Testing with Multiplex STRs
AmpFISTR® SGM Plus™ kit

Amelogenin D19 D3 D8 TH01 VWA D21
FGA D16 D18 D2

Two different individuals

Probability of a random match: ~1 in 3 trillion

Results obtained in less than 5 hours with a spot of blood the size of a pinhead

Simultaneous Analysis of 10 STRs and Gender ID

Dr. Peter M. Vallone
### Product Rule

For heterozygous loci

\[ P = 2pq \]

\( P \) = probability; \( p \) and \( q \) are frequencies of allele in a given population

Example: For the locus D3S1358 and individual is 15,18 with frequencies of 0.2825 and 0.1450 respectively

\[ P = 2(0.2825)(0.1450) = 0.0819 \text{ or 1 in 12} \]

For 5 loci the Profile Probability = \((P_1)(P_2)\ldots(P_n)\)

\[ = (0.0819)(0.0875)(0.0687)(0.0245)(0.0984) \]

\[ = 0.000001187 \text{ or 1 in 842,539} \]

---

**DNA Profiles from a Single Region (Locus)**

```
DNA Lineup of the “Suspects”
```

**“Crime Scene” Evidence**

A
B
C
D
E
DNA Profiles from Multiple Regions

“Crime Scene” Evidence

“Suspects”

B

E

C

D

A

Paternity Testing
Family Inheritance of STR Alleles (D13S317)

PCR product size (bp)

Child #1

Child #2

Child #3

Mother
**Paternity Testing**

Family Inheritance of STR Alleles (D13S317)

<table>
<thead>
<tr>
<th>PCR product size (bp)</th>
<th>110</th>
<th>120</th>
<th>130</th>
<th>140</th>
<th>150</th>
<th>160</th>
<th>170</th>
<th>180</th>
<th>190</th>
<th>200</th>
<th>210</th>
<th>220</th>
<th>230</th>
<th>240</th>
<th>250</th>
</tr>
</thead>
<tbody>
<tr>
<td>Father</td>
<td>11</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child #1</td>
<td>12</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child #2</td>
<td>8</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child #3</td>
<td>11</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother</td>
<td>8</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*FBI database*
Combined DNA Index System

Used for linking serial crimes and unsolved cases with repeat offenders
Convicted offender and forensic case samples
Launched October 1998 and links all 50 states
Requires 13 core STR markers
Current backlog of >750,000 samples ($15M in FY2002 to reduce backlog)
As of April 2003 the total profile composition of the National DNA Index System (NDIS) is as follows:

- Total number of profiles: 1,376,749
- Total Forensic profiles: 54,895
- Total Convicted Offender Profiles: 1,321,854

All 50 states now require convicted offenders to submit a sample for DNA testing purposes.

7,788 Investigations Aided through April 2003

http://www.fbi.gov/hq/lab/codis/clickmap.htm

Automation and Robotics

Automated PCR Setup and CE Sample Preparation
The Role of NIST Scientists

- **Develop DNA standards** so that laboratories around the world may compare their results.
- **Conduct tests of laboratories** around the world to insure accurate results in DNA testing.
- **Develop new DNA tests** which are more rapid and efficient than those currently used.
- **Create useful information databases** (STRBase) [http://www.cstl.nist.gov/biotech/strbase](http://www.cstl.nist.gov/biotech/strbase).
- **Evaluation and development** of new technologies.

NIST Standard Reference Materials (SRMs)

- SRM 2390 - DNA Profiling Standard
  Meets RFLP Needs
- SRM 2391 - PCR-Based DNA Standard
  Cell Lines and Genomics
- SRM 2392 - Mitochondrial DNA Standard
  Cell Lines and Cloned HV1 Plasmid
- **SRM 2393 – mtDNA heteroplasmy**
- **SRM 2395 – Y chromosome DNA standards**
DAB Quality Assurance Standards for Forensic DNA Testing Laboratories

STANDARD 9.5

The laboratory shall check its DNA procedures annually or whenever substantial changes are made to the protocol(s) against an appropriate and available NIST standard reference material or standard traceable to a NIST standard.
NIST Y Chromosome Standard

6 genomic DNA samples
5 male and 1 female
Typing Information on 27 Y STRs and 50 Y SNP markers

Available as of 07/2003
Sequence Determination of Y STR Repeat Region for Each Component

**DYS392 (forward) A**
13 TAT repeats

**DYS392 (forward) B**
11 TAT repeats

**DYS392 (forward) C**
11 TAT repeats

**DYS392 (forward) D**
11 TAT repeats

**DYS392 (forward) E**
12 TAT repeats

**Sequencing Performed**
- DYS19
- DYS385 a/b
- DYS389 I/II
- DYS390
- DYS391
- DYS392
- DYS393
- DYS388
- DYS426
- DYS435
- DYS436
- DYS437
- DYS438
- DYS439
- DYS447
- DYS460 (A7.1)
- DYS461 (A7.2)
- Y-GATA-H4
- DYS462

WTC DNA Identifications
World Trade Center Towers  
(Sept 11, 2001)

DNA typing being used as only possible method to identify over 2,000 victims of this tragedy

Highly degraded DNA; ~20,000 samples recovered; 3 years to complete...

Wreckage at Ground Zero

Identifying Mass Disaster Victims

Medical Specimen (Biopsy sample/bloodspot)
Personal Effect (toothbrush/hairbrush)
Kin (close relative)
Remains (bone/ tissue)

DNA profiles obtained are stored in a database and attempts at identification are made

Challenge lies in typing degraded samples
Aged Blood Stain Studies
(results from two different DNA samples)

“Decay curve” of degraded DNA

6 years old (-20 °C storage)

Results with SGM Plus STR kit (Applied Biosystems)

Typing Result on Aged Blood Stain
(15 years at room temperature storage)

When working with degraded samples it is difficult to generate longer PCR products
STR Size Reduction
Through Moving Primer Positions Closer to Repeat

**Advantages of Approach:**
Size reduction enhances success rate with degraded DNA
Retains same marker information (database compatibility)
Uses highly polymorphic STR loci (high discriminatory power)

Limitation: Lower levels of multiplexing 5-6 plex vs 10-15 plex

Development of miniSTR Assays

- Primers designed to come as close as possible to the repeat region to generate the smallest possible PCR products
- Equivalent genotypes are obtained when compared with commercial STR kits
- Available as singleplexes or miniplexes (usually one locus per dye color)
- Smaller amplicons offer improved chances of success with degraded DNA samples
- Project begun in November 2001 at the request of Bob Shaler to aid WTC DNA identifications
Development of miniSTRs to Aid Testing of Degraded DNA

Identical Typing Results in ALL 100+ Samples Tested

COfiler™ kit

Over 150 bp size reduction!

Comparison between miniSTR and Commercial Kit Results (Aged Blood Stain Known to Contain Degraded DNA)
Current Status of WTC Samples

28,251 Total Profiles
Personal Effects 4,903
Kin 6,876
Remains 16,472

2795 reported missing – 1511 Identified

Use of miniSTRs has resulted in ~20 additional identifications

WTC Kinship and Data Analysis Panel
Cat STR 12plex under development at NIST

4 ng male cat DNA; Sample #5; 3/12/02
Acknowledgments

Funding:
Interagency Agreement between NIJ and NIST Office of Law Enforcement Standards

NIST Personnel:
(Project Leader) John Butler

Margaret Kline
Jan Redman
Rich Schoske (AU grad student)

peter.vallone@nist.gov (x4872) 227 B242