Forensic DNA Typing: The Application of Nucleic Acid Based Technology to Human Identity Testing

Guest Lecture Series
Science and Engineering Alliance
Southern University and A&M College, Baton Rouge, LA

Dr. Peter M. Vallone, Biochemical Science Division, National Institute of Standards and Technology
Gaithersburg, Maryland

The National Institute of Standards and Technology (NIST)

Advanced Chemical Sciences Laboratory Building

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.

Outline

- NIST Campus
- DNA, Short Tandem Repeats (STRs) and the polymerase chain reaction (PCR)
- Uses of DNA testing
- Steps involved in sample typing
- NIST Standard Reference Materials (SRMs)
- Applications

Location of NIST (Gaithersburg, MD)

NIST

~30 miles

Washington D.C.

Reagan National Airport

Baltimore, MD

BWI Airport

Richmond, VA

Outline

- NIST Campus
- DNA, Short Tandem Repeats (STRs) and the polymerase chain reaction (PCR)
- Uses of DNA testing
- Steps involved in sample typing
- NIST Standard Reference Materials (SRMs)
- Applications

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Working with DNA is relatively easy

Lengths of DNA up to 100 base pairs can be commercially synthesized and purified.

DNA is relatively stable: can be stored in water or low salt buffer at 4°C for 6 months to a year.

Single strands can be functionalized and attached to beads or a glass/silicon surface.

Fluorescent dyes can be covalently attached on the 5' end of the molecule for detection purposes.

Single strands of DNA have a strong affinity for their complement

<table>
<thead>
<tr>
<th>Length</th>
<th>Sequence</th>
<th>$K_a$ (M$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5mer</td>
<td>gctca</td>
<td>8.0E+01</td>
</tr>
<tr>
<td>10mer</td>
<td>acgtatgctgatc</td>
<td>8.3E+05</td>
</tr>
<tr>
<td>15mer</td>
<td>gctatcgctgatc</td>
<td>1.4E+09</td>
</tr>
<tr>
<td>20mer</td>
<td>acgtatcgctgatc</td>
<td>5.5E+13</td>
</tr>
<tr>
<td>25mer</td>
<td>acgtatcgctgatc</td>
<td>3.1E+16</td>
</tr>
</tbody>
</table>

Other binding constants: Strepavidin-biotin[10$^{15}$], Drugs[~10$^6$], Antibodies[10$^7$ to 10$^{11}$]

DNA is a biopolymer that consists of only 4 monomers units

Relationship: $4^N$ where $N$ is the length of the sequence

<table>
<thead>
<tr>
<th>Length (nt)</th>
<th>Possible unique seqs</th>
<th>Base pairs in Human Genome</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>1024</td>
<td>3.2E+09</td>
</tr>
<tr>
<td>10</td>
<td>1.0E+06</td>
<td>3.2E+09</td>
</tr>
<tr>
<td>15</td>
<td>1.1E+09</td>
<td>3.2E+09</td>
</tr>
<tr>
<td>20</td>
<td>1.1E+12</td>
<td>3.2E+09</td>
</tr>
<tr>
<td>25</td>
<td>1.1E+15</td>
<td>3.2E+09</td>
</tr>
</tbody>
</table>

It is reasonable to assume that DNA molecules of ~15 units or greater are unique in the human genome (exceptions, repeats, duplicated regions etc).

DNA in the Cell

DNA in the Cell

Our DNA Comes from our Parents

Father's Sperm

Mother's Egg

Child's Cell

What Type of Genetic Variation?

- **Length Variation**
  - short tandem repeats (STRs)
  - CTAGTCGT(GATA)(GATA)(GATA)GCCATCGT

- **Sequence Variation**
  - single nucleotide polymorphisms (SNPs)
  - insertions/deletions
  - GCTAGTCGATGCT(G/A)GGATGCTGAGC

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Short Tandem Repeat (STR) Markers
An accordion-like DNA sequence that occurs between genes

\[\text{TCCCAAGCTCTTCC} \]
\[\text{TCTTCCCTAGATCAATACAGACAGAAGACA} \]
\[\text{GGTG} \]
\[\text{GATAGATAGATAGATAGATAGATAGATAGATAGATAGA} \]
\[\text{TAGATAGATA} \]
\[\text{TCATTGAAAGACAAAACAGAGATGGATGATAGAT} \]
\[\text{ACATGCTTACAGATGCACAC} \]
\[= 12 \text{ GATA repeats ("12" is all that is reported)} \]

The number of consecutive repeat units can vary between people

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

13 CODIS Core STR Loci with Chromosomal Positions

<table>
<thead>
<tr>
<th>Locus</th>
<th>Chromosomal Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>TPX</td>
<td>D3S1358</td>
</tr>
<tr>
<td>D5S818</td>
<td>D8S1179</td>
</tr>
<tr>
<td>CSF1PO</td>
<td>D21S11</td>
</tr>
<tr>
<td>FGA</td>
<td>D13S317</td>
</tr>
<tr>
<td>TH01</td>
<td>D18S51</td>
</tr>
<tr>
<td>VWA</td>
<td>D16S539</td>
</tr>
<tr>
<td>D7S820</td>
<td>D8S1179</td>
</tr>
<tr>
<td>D18S51</td>
<td>D21S111</td>
</tr>
<tr>
<td>AMEL</td>
<td>D16S539</td>
</tr>
<tr>
<td>AMEL</td>
<td>D18S51</td>
</tr>
</tbody>
</table>

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)

Chromosomes are paired so…

Homozygous – Alleles are identical on each chromosome

Heterozygous - Alleles differ on each on each chromosome

Outline

• NIST Campus
• DNA, Short Tandem Repeats (STRs) and the polymerase chain reaction (PCR)
• Uses of DNA testing
• Steps involved in sample typing
• NIST Standard Reference Materials (SRMs)
• Applications

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Fingerprints have been used since 1901

DNA since 1986

Methods for Human Identification

Characteristics of Genomic DNA

- Each person has a unique DNA profile (except identical twins)
- Each person’s DNA is the same in every cell (DNA from skin cells will match DNA from blood cells)
- An individual’s DNA profile remains the same throughout life
- Half of your DNA comes from your mother and half from your father

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, predisposition 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 person 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...

A Brief History of Forensic DNA Typing

- 1985: PCR developed
- 1990: mtDNA
- 1994: First STRs described
- 1996: First commercial fluorescent STR multiplex kits
- 1998: STR typing with CE is routinely used worldwide
- 2000: US core loci defined
- 2002: STR = short tandem repeat
- 2006: DNA is an important part of the criminal justice system

Unfortunately, current DNA testing cannot be performed as quickly as a commercial break...

The instruments on CSI are real – they just do not collect data as quickly as shown on TV.

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

• NIST Campus
• DNA, Short Tandem Repeats (STRs) and the polymerase chain reaction (PCR)
• Uses of DNA testing
• Steps involved in sample typing
• NIST Standard Reference Materials (SRMs)
• Applications

Steps in DNA Analysis

Collection
Extraction
Quantitation
Genotyping
Interpretation of Results
Database
Storage & Searching

Steps in DNA Analysis

Collection
Extraction
Quantitation
Genotyping
Interpretation of Results
Database
Storage & Searching

Sources of Biological Evidence

• Blood
• Semen
• Saliva
• Urine
• Hair
• Teeth
• Bone
• Tissue

Blood sample

Only a very small amount of blood is needed to obtain a DNA profile

Steps in DNA Analysis

A small punch taken from the buccal device and is incubated in an extraction buffer
It is important to remove the heme from blood sample as it will inhibit the PCR
Commercial chromatography columns are available for the extraction process

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Steps in DNA Analysis

Collection
Extraction
Quantitation
Genotyping
Interpretation of Results
Database

It is important that the optimal amount of DNA is added to the PCR. Typically, 0.5 ng to 2 ng works well. Too much or too little DNA will result in artifacts obscuring data interpretation. DNA can be quantitated using UV spectroscopy, fluorescence (after the addition of an intercalation dye), hybridization with a probe, and qRT-PCR.

Impact of DNA Amount into PCR

- Too much DNA
  - Off-scale peaks
  - Split peaks (+/−A)
  - Locus-to-locus imbalance

- Too little DNA
  - Heterozygote peak imbalance
  - Allele drop-out
  - Locus-to-locus imbalance

D3S1358

10 ng template (overloaded)
2 ng template (suggested level)

DNA Size (bp)

Relative Fluorescence (RFUs)

100 pg template
5 pg template

Stochastic effect when amplifying low levels of DNA produces allele drop out.

Steps in DNA Analysis

Collection
Extraction
Quantitation
Genotyping
Interpretation of Results
Database

The quantitated extract is then amplified (PCR) using a variety of commercial kits containing fluorescently labeled PCR primers that exist. Up to 16 loci are simultaneously amplified. Amplification takes approximately 2 hours. The PCR products are diluted and separated/detected on a gel or capillary platform.

Multiplex PCR
(Parallel Sample Processing)

Multiple primer sets target multiple sites in the human genome. 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)

Capillary Electrophoresis System

Capillary Electrophoresis Instrumentation

ABI 310
single capillary

ABI 3100
16-capillary array

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
**Principles of Sample Separation and Detection**

**Sample Detection**

- Labeled DNA fragments (PCR products)
- Capillary or Gel Lanes
- Size Separation
- Detection region

**Color Separation**

- Ar+ LASER (488 nm)
- Fluorescence
- ABI Prism spectrograph

**Butler, J.M. (2001) Forensic DNA Typing, Figure 10.8, ©Academic Press**

---

**Steps in DNA Analysis**

- **Collection**
- **Extraction**
- **Quantitation**
- **Genotyping**
- **Interpretation of Results**
- **Database Storage & Searching**

**The peak(s) for each locus are compared to an allelic ladder which contains all the possible alleles.**

**A repeat number is assigned for each observed peak.**

**The repeat values are tabulated for each locus.**

**Visual comparisons between data can be helpful.**

---

**Result from the Identifiler STR Kit**

![Image of the Identifiler STR Kit result](image_url)

**Identifiler STR Kit**

Information is tied together with multiplex PCR and data analysis.

---

**Companies Supply Allelic Ladders in STR Kits to Aid Interlaboratory Consistency**

![Image of Allelic Ladders](image_url)

---

**http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm**
Crime Scene - Two Suspects

Suspect 1

Suspect 2

Evidence

DNA Profiles from Multiple Regions

“Crime Scene” Evidence

“Suspects”

Evidence

Data Format

<table>
<thead>
<tr>
<th>AMEL</th>
<th>CSF1PO</th>
<th>FGA</th>
<th>TH01</th>
<th>VWA</th>
<th>D3S1358</th>
<th>D5S818</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>Y</td>
<td>11,12</td>
<td>10,21</td>
<td>8,1</td>
<td>8,8</td>
<td>15,18</td>
</tr>
</tbody>
</table>

The number of repeats observed for each locus is tabulated.

This data format is stored in databases and used for comparisons/matches.

Paternity Testing

Family Inheritance of STR Alleles (D13S317)

PCR product size (bp)

<table>
<thead>
<tr>
<th>100</th>
<th>150</th>
<th>200</th>
<th>250</th>
<th>300</th>
<th>350</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>14</td>
<td>12</td>
<td>14</td>
<td>11</td>
<td>12</td>
<td>8</td>
</tr>
</tbody>
</table>

Paternity Testing

Family Inheritance of STR Alleles (D13S317)

PCR product size (bp)

<table>
<thead>
<tr>
<th>100</th>
<th>150</th>
<th>200</th>
<th>250</th>
<th>300</th>
<th>350</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>14</td>
<td>12</td>
<td>14</td>
<td>11</td>
<td>12</td>
<td>8</td>
</tr>
</tbody>
</table>

Same DNA sample run with Applied Biosystems STR Kits

Random Match Probability

1.0 x 10^-3

7.8 x 10^-4

9.0 x 10^-11

2.4 x 10^-11

2.0 x 10^-7

4.5 x 10^-13
Steps in DNA Analysis

Collection
Extraction
Quantitation
Genotyping
Interpretation of Results
Database

Combined DNA Index System (CODIS)
- Used for linking serial crimes and unsolved cases with repeat offenders
- Convicted offender and forensic case samples
- Launched October 1998
- Requires 13 core STR markers
- Annual Results with NIST SRM required for submission of data to CODIS

Coordinated DNA Index System (CODIS)
- Used for linking serial crimes and unsolved cases with repeat offenders
- Convicted offender and forensic case samples
- Launched October 1998
- Requires 13 core STR markers
- Annual Results with NIST SRM required for submission of data to CODIS

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

>27,079 investigations Aided through August 2005

As of August 2005 the total profile composition of the National DNA Index System (NDIS) is as follows:
Total number of profiles: 2,695,885
Total Forensic profiles: 117,255
Total Convicted Offender Profiles: 2,578,630

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

STRBase
Short Tandem Repeat DNA Internet Database

Recent Additions
- Forensic SNP information (will be official site for ISFG SNP information)
- NIST publications and presentations as pdf files
- NIST Publications and Presentations as PDF files

We Regularly Update
- Reference List
- Variant Alleles
- Addresses for Scientists
- Links to Other Web Sites
- Y-STR Information

We will continue to add downloadable PowerPoint files that can be used for training purposes

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

NIST Standard Reference Materials (SRMs)

SRM 2390 - DNA Profiling Standard Meets RFLP Needs
NIST SRMs are used to help calibrate forensic genotyping laboratories

SRM 2392 - Mitochondrial DNA Standard Cell Lines and Cloned HV1 Plasmid
SRM 2393 – mtDNA heteroplasmy
SRM 2395 – Y chromosome DNA standards

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
NIST Y Chromosome 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

Outline

- NIST Campus
- DNA, Short Tandem Repeats (STRs) and the polymerase chain reaction (PCR)
- Uses of DNA testing
- Steps involved in sample typing
- NIST Standard Reference Materials (SRMs)
- Applications

INNOCENCE PROJECT
http://www.innocenceproject.org

August 17, 1998
FBI Report on Analysis of Stain on Monica Lewinsky’s Blue Dress

http://www.law.umkc.edu/faculty/projects/ftrials/clinton/lewinskydress.html


“We got him!”

Saddam was known to have many “stunt doubles” that acted as decoys for his own safety.

Saddam Hussein’s capture was verified with DNA testing conducted in Rockville, MD at the Armed Forces DNA Identification Laboratory

Source: www.cnn.com; The Scientist Dec 19, 2003

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Tomb of the Unknown Soldier

- Armed Forces DNA Identification Laboratory (AFDIL) (Rockville, MD)
- In June 1998 AFDIL identified Michael J. Blassie as the Vietnam Unknown in the Tomb of the Unknown Soldier (located in Arlington National Cemetery)
- There will be no more “unknown” soldiers.


World Trade Center Towers (Sept 11, 2001)

- Wreckage at Ground Zero
- Highly degraded DNA was recovered

Using Personal Effects to Identify Remains

DNA Testing on Bones from WTC Site

- Bone specimens received
- Bone powder extracted
- STR profile generated
- Typing data uploaded to WTC CODIS database to search for match

New DNA Tests Pioneered at NIST

- New test developed to aid in identification of World Trade Center victims of 9/11/01 terrorist attacks
- Has resulted in an increase in the number of WTC victims identified

World Trade Center – Phase I Summary

- 12,392 Bone samples processed
- Over 6800 profiles
- 3,405 Full profiles (13 STR loci)
- 2,143 High partial profiles (>7 STR loci)
- 2,670 Low partial profiles (<7 STR loci)
- 4,174 No loci

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Tsunami Survivor “Baby 81” Connected to His Parents with DNA

NEW YORK (AP) -- The parents of the infant tsunami survivor nicknamed “Baby 81” say they found it difficult to feel overjoyed about their reunion in the midst of so much tragedy. The 4-month-old Sri Lankan baby and his parents, who were reunited after court-ordered DNA tests proved their relationship, appeared on ABC’s “Good Morning America” Wednesday, a day after their 20-hour-long flight landed in New York.


Hurricane Katrina Victims Will Be Identified with Forensic DNA Testing Methods

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