DNA and THE LAW

Brian Zubel, J.D.
and
John M. Butler, Ph.D.
National Institute of Standards and Technology
Frankenmuth, MI
May 4, 2007

NIST and NIJ Disclaimer

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

Points of view are mine and do not necessarily represent
the official position or policies of the US Department of Justice or the
National Institute of Standards and Technology.

 Certain commercial equipment, instruments and materials are identified
in order to specify experimental procedures as completely as
possible. In no case does such identification imply a
recommendation or endorsement by the National Institute of
Standards and Technology nor does it imply that any of the
materials, instruments or equipment identified are necessarily the
best available for the purpose.

Interfaces Between Disciplines Are Crucial

Law

Enforcement

Judicial

Laboratory

Lessons from the First Case Involving DNA Testing

Describes the first use of DNA (in 1986) to solve a double rape-homicide case in England; about 5,000 men asked to give blood or saliva to compare to crime stains

• Connection of two crimes (1983 and 1986)
• Use of DNA database to screen for perpetrator (DNA only done on 10% with same blood type as perpetrator)
• Exoneration of an innocent suspect
• DNA was an investigative tool – did not solve the case by itself (confession of accomplice)

A local baker, Colin Pitchfork, was arrested and his DNA profile matched with the semen from both murders. In 1988 he was sentenced to life for the two murders.

Impact of Forensic DNA Testing

Guilt

Innocence

Colin Pitchfork
Kirk Bloodsworth Josiah Sutton
Roger Coleman

Some DNA Resources for Lawyers

http://www.cstl.nist.gov/biotech/strbase/training.htm
Information Resources for Defense Attorneys

Defense Lawyers and Experts are becoming more united and informed

Common Defense Attacks
Compiled from Forensic Bioinformatics website
- Contamination
- Statistical Weight of a Match
- Degradation/PCR Inhibition of “True” Perp
- Artifacts (N+4 stutter, etc.)
- Thresholds Set Too High (missing peaks)
- Examiner Bias
- Improper Mixture Interpretation
- Meaning of a Database Hit

See http://www.bioforensics.com/conference07/index.html

NIST Background

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.

Location of NIST (near Washington, DC)

NIST Gaithersburg Campus
Located in Gaithersburg, Maryland, on approximately 234 hectares (578 acres) just off Interstate 270 about 25 miles northwest of Washington, D.C.

http://www.cstl.nist.gov/biotech/strbase/training.htm
Our Team Mission Statement

• The NIST Human Identity Project Team is trying to lead the way in forensic DNA... through research that helps bring traceability and technology to the scales of justice.

NIST Human Identity Project Team

• John Butler
• Margaret Kline
• Pete Valenza
• Jan Beeman
• Amy Decker
• Becky Hill
• Dave Duewer

All NIST publications and presentations available on STRBase:
http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm

• 26 publications from Jan-Dec 2006
• 45 presentations and 10 workshops to the community from Jan-Dec 2006

Current Areas of NIST 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
  – Assay and software development

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

Standard Reference Materials
http://www.cstl.nist.gov/biotech/strbase/srm_tab.htm

Traceable standards to ensure accurate measurements in our nation’s crime laboratories

- SRM 2391b – CODIS STRs
- SRM 2392-I – mtDNA
- SRM 2395 – Y-STRs
- SRM 2372 – DNA quantitation

Calibration with SRMs enables confidence in comparisons of results between laboratories

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

Provides up-to-date information and has been used in court cases to support application of DNA technology

Lab Procedures
Steps in DNA Analysis

Steps Involved
- Collection
- Sample Storage
- Extraction
- Quantitation
- Multiplex PCR
- STR Typing
- Interpretation of Results
- Database Storage & Searching
- Calculation of Match Probability

Steps in DNA Analysis

- DNA Extraction
- Multiplex PCR Amplification
- Male: 13,14-15,16-12,13-10,13-15,16
- Interpretation of Results

Sample Collection & Storage

Buccal swab
Blood Stain
DNA Quantitation
Slot Blot

1 ng
0.3 ng
1 ng
1 ng
0.7 ng
0.5 ng
0.5 ng
No DNA

Usually 1-2 day process (a minimum of ~5 hours)

If a match occurs, comparison of DNA profile to population allele frequencies to generate a case report with probability of a random match to an unrelated individual.

STR Typing
DNA separation and sizing

Technology
Biology
Genetics
DNA Database
Search

DNA Evidence Received in the Lab

- Evidentiary samples (commonly in the form of cotton swabs) are brought or shipped to the DNA laboratory after collection from the crime scene or victim
- Sexual assault evidence collection kits provide swabs and bags for clothing collections from the victim

DNA Collection

- Cotton swabs are commonly used to collect biological material from bloodstains or semen from sexual assault victims
- The amount of DNA needed has decreased dramatically in the past decade due to sensitivity of the PCR process (which makes millions of copies of targeted regions)

DNA Reference Sample from Suspect

- Blood samples may be collected but require a phlebotomist to draw blood
- Easier to collect a buccal swab from the inside of an individual’s mouth, which scrapes off some cheek cells

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

best results with >100 cells, but DNA profiles can be recovered from as little as a single cell

http://www.cstl.nist.gov/biotech/strbase/training.htm
Buccal Swab DNA Collection

- The inside of the cheek is scrubbed to collect cells
- Less invasive than drawing blood
- Swab must be dried before storing and shipping to lab to avoid mold and bacterial growth

DNA Extraction

- DNA is extracted from proteins that protect it in the nucleus of a cell
- Chemicals are added to digest the protecting proteins and produce “naked” DNA molecules
- The final solution looks like a tube of water

DNA Quantitation

- DNA quantitation is important to determine how much human DNA (as opposed to bacterial DNA) is present in a sample
- A commonly used DNA quantitation kit is called Quantifiler (sold by Applied Biosystems)

Impact of DNA Amount into PCR

Reason that DNA Quantitation is Important Prior to Multiplex Amplification

- 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

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

PCR Amplification

- PCR = polymerase chain reaction
- Process that copies a particular region of DNA using two “primers” (short pieces of DNA)
- Each strand of DNA is used as a template to create a replicate that permits a doubling of the number of target molecules with each cycle of heating and cooling

http://www.cstl.nist.gov/biotech/strbase/training.htm
PCR Amplification (Thermal Cycling)
- The polymerase chain reaction (PCR) copies sections of DNA through heating and cooling the sample
- Each DNA strand is copied with each temperature cycle
- A thermal cycler heats and cools DNA samples (usually 28 cycles)

DNA in the Cell
- Cell nucleus
- Double stranded DNA molecule
- Approximately 3 billion total base pairs

Identification of Information
<table>
<thead>
<tr>
<th>Printed Information</th>
<th>Genetic Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Library</td>
<td>Body</td>
</tr>
<tr>
<td>Book</td>
<td>Cell</td>
</tr>
<tr>
<td>Chapter</td>
<td>Nucleus</td>
</tr>
<tr>
<td>Page Number</td>
<td>Chromosome</td>
</tr>
<tr>
<td>Line on Page</td>
<td>Locus (part of chromosome)</td>
</tr>
<tr>
<td>Word</td>
<td>Short DNA sequence</td>
</tr>
<tr>
<td>Letter</td>
<td>DNA nucleotides</td>
</tr>
</tbody>
</table>

What is a DNA Profile?
- Unique regions of the human genome are targeted
- These regions consist of a few hundred base pairs
- The regions are copied by the polymerase chain reaction (PCR) – billions of exact copies
- The copied fragments now contain fluorescent dyes for detection

Short Tandem Repeat (STR) Markers
- An accordion-like DNA sequence that occurs between genes
- TCCCAAGCTCTCCCTCCCTCCTCCTCCTAGATCAAATACAGACAGAAGACA
- CGTTGATAAGATGATAAGATGATAAGATGATAAGATGATAAGATGATAAGAT
- TAGATAGATAAGATGATAAGATGATAAGATGATAAGATGATAAGATGATAAGAT
- ACATGCTTACAGATGACAC
- Target region (short tandem repeat)

13 CODIS Core STR Loci with Chromosomal Positions
- 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

http://www.cstl.nist.gov/biotech/strbase/training.htm
Short Tandem Repeat (STR) Markers

PCR primers anneal to unique sequences bracketing the variable STR repeat region. Fluorescent dye labels the amplified DNA template containing STR marker.

PCR product size generated

DNA template

PCR primer

Forward

Reverse

GATA

GATA

GATA

GATA

GATA

GATA

GATA

GATA

PCR primers anneal to unique sequences bracketing the variable STR repeat region.

= 12 GATA repeats ("12" is all that is reported)

DNA Reaction Setup

- DNA sample is added (about 1 ng based on DNA quantitation performed) – 10 μL
- PCR primers and other reaction chemicals from an STR typing kit are added – 15 μL

Strip of 8 tubes containing ~25 μL of solution

STR Typing Kit

- Kit Components:
  - Primer mix
  - PCR Reaction Buffer and Building Blocks
  - DNA Polymerase (Taq Gold)

- Most expensive reagent

- Common kits used:
  - Identifiler (Applied Biosystems)
  - Profiler Plus/COffer (Applied Biosystems)
  - PowerPlex 16 (Promega)

What is in an STR Typing Kit?

- Primer mix
  - containing fluorescently labeled oligonucleotides used to target specific regions of the human genome

- Applied Biosystems has not published their primer sequences

- PowerPlex 16, which amplifies 16 genomic sites, contains 32 PCR primers

PCR Primers in an STR Kit

<table>
<thead>
<tr>
<th>Locus</th>
<th>Dye</th>
<th>Primer Seq.</th>
<th>Primer Sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Locus 1</td>
<td>JOE</td>
<td>D3S1358-F</td>
<td>ACTGCAGTCCAATCTGGGT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D3S1358-R</td>
<td>FL ATGAAATCAACAGAGGCTTGC</td>
</tr>
<tr>
<td>Locus 2</td>
<td>JOE</td>
<td>TH01-F</td>
<td>FL GTGATTCCCATTGGCCTGTTC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TH01-R</td>
<td>ATTCCTGTGGGCTGAAAAGCTC</td>
</tr>
<tr>
<td>Locus 3</td>
<td>JOE</td>
<td>D21S11-F</td>
<td>ATATGTGAGTCAATTCCCCAAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D21S11-R</td>
<td>FL TGTATTAGTCAATGTTCTCCAGAGAC</td>
</tr>
<tr>
<td>Locus 4</td>
<td>JOE</td>
<td>D18S51-F</td>
<td>FL TTCTTGAGCCCAGAAGGTTA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D18S51-R</td>
<td>ATTCTACCAGCAACAACACAAATAAAC</td>
</tr>
<tr>
<td>Locus 5</td>
<td>JOE</td>
<td>PentaE-F</td>
<td>FL ATTACCAACATGAAAGGGTACCAATA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PentaE-R</td>
<td>FL TGGGTTATTAATTGAGAAAACTCCTTACAATTT</td>
</tr>
<tr>
<td>Locus 6</td>
<td>JOE</td>
<td>D5S818-F</td>
<td>GGTGATTTTCCTCTTTGGTATCC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D5S818-R</td>
<td>JOE AGCCACAGTTTACAACATTTGTATCT</td>
</tr>
<tr>
<td>Locus 7</td>
<td>JOE</td>
<td>D13S317-F</td>
<td>ATTACAGAAGTCTGGGATGTGGAGGA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D13S317-R</td>
<td>JOE GGCAGCCCAAAAAGACAGA</td>
</tr>
<tr>
<td>Locus 8</td>
<td>JOE</td>
<td>D7S820-F</td>
<td>JOE ATGTTGGTCAGGCTGACTATG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D7S820-R</td>
<td>GATTCCACATTTATCCTCATTGAC</td>
</tr>
<tr>
<td>Locus 9</td>
<td>JOE</td>
<td>D16S539-F</td>
<td>GGGGGTCTAAGAGCTTGTAAAAAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D16S539-R</td>
<td>JOE GTTTGTGTGTGCATCTGTAAGCATGTATC</td>
</tr>
<tr>
<td>Locus 10</td>
<td>JOE</td>
<td>CSF1PO-F</td>
<td>JOE CCGGAGGTAAAGGTGTCTTAAAGT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CSF1PO-R</td>
<td>JOE TGGGTTATTAATTGAGAAAACTCCTTACAATTT</td>
</tr>
<tr>
<td>Locus 11</td>
<td>JOE</td>
<td>PentaD-F</td>
<td>JOE GAAGGTCGAAGCTGAAGTG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PentaD-R</td>
<td>ATTAGAATTCTTTAATCTGGACACAAG</td>
</tr>
<tr>
<td>Locus 12</td>
<td>JOE</td>
<td>AMEL-F</td>
<td>TMR CCCTGGGCTCTGTAAAGAA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AMEL-R</td>
<td>ATCAGAGCTTAAACTGGGAAGCTG</td>
</tr>
<tr>
<td>Locus 13</td>
<td>JOE</td>
<td>vWA-F</td>
<td>GCCCTAGTGGATGATAAGAATAATCAGTATGTG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>vWA-R</td>
<td>TMR GGACAGATGATAAATACATAGGATGGATGG</td>
</tr>
<tr>
<td>Locus 14</td>
<td>JOE</td>
<td>D8S1179-F</td>
<td>TMR ATTGCAACTTATATGTATTTTTGTATTTCATG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D8S1179-R</td>
<td>ACCAAATTGTGTTCATGAGTATAGTTTC</td>
</tr>
<tr>
<td>Locus 15</td>
<td>JOE</td>
<td>TPOX-F</td>
<td>TMR GCACAGAACAGGCACTTAGG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TPOX-R</td>
<td>TMR CGCTCAAACGTGAGGTTG</td>
</tr>
<tr>
<td>Locus 16</td>
<td>JOE</td>
<td>FGA-F</td>
<td>TMR GGCTGCAGGGCATAACATTA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FGA-R</td>
<td>ATTCTATGACTTTGCGCTTCAGGA</td>
</tr>
</tbody>
</table>

The polymerase chain reaction (PCR) is used to amplify STR regions and label the amplicons with fluorescent dyes using locus-specific primers.
### Transfer of DNA Samples

- Following PCR, a small portion of the sample is transferred for analysis.
- This aliquot of the sample is mixed with a molecular size marker (termed an internal size standard) that permits calibration of sizing measurements.

### Sample Plates Spun Down via a Centrifuge

- Sample plates are spun to remove bubbles that would interfere with the injection (loading) process onto the capillary electrophoresis instrument.

### ABI 3130xl DNA Analysis Instrument

- Import sample names
- Determine run conditions (voltages and times to be used based on laboratory protocols)

### Data Collection on ABI 3130xl Instrument

- Data analysis is performed on an Applied Biosystems (ABI) 3130xl capillary electrophoresis instrument.

### Capillary Electrophoresis Instrumentation

- **ABI 310**
  - Single capillary
- **ABI 3100**
  - 16-capillary array

### Steps in STR Typing with ABI 310/3100

1. **Sample Preparation**
   - Mixture of dye-labeled PCR products from multiplex PCR reaction
2. **Sample Injection**
   - Sample is injected into the capillary
3. **Capillary**
   - Polymer solution
4. **Size Separation**
   - According to molecular weight
5. **Fluorescence**
6. **Color Separation**
7. **CCD Panel**
   - With virtual filters
8. **Sample Detection**
   - Fluorescence detected by a microplate reader
9. **Sample Interpretation**
   - Genetic profiles are analyzed by computer software

---

For more detailed information, please visit [http://www.cstl.nist.gov/biotech/strbase/training.htm](http://www.cstl.nist.gov/biotech/strbase/training.htm)

A DNA Profile is Produced by Separating DNA Molecules by Size and Dye Color

The labeled fragments are separated (based on size) and detected on a gel or capillary electrophoresis instrument.

~2 hours or less

Peaks represent labeled DNA fragments separated by electrophoresis. This ‘profile of peaks’ is unique for an individual—a DNA type.

Data Transfer

- Following data collection, the data (.fsa) files are typically transferred from the lab computer to one in an office where data analysis is performed.
- A USB thumb drive permits rapid and easy transfer of data files.

Data Analysis

- The analyst carefully reviews the DNA data (electropherogram) and checks software genotype calls and edits out artifacts.
- Software designates sample genotypes via comparison to an allelic ladder (mixture of common allele possibilities).

Comparison of Allelic Ladder to Samples to Convert Size into Allele Repeat Number

difference = -0.02 bp

difference = +0.05 bp

± 0.5 bp bin defined around each allele.

STR Results

- Individuals will differ from one another in terms of their STR profile.
- STR genotype can then be put into an alpha numeric form for search on a DNA database.

What would be entered into a DNA database for searching:

16,17-17,17-17,17-18,17-21,22-12,14-28,30-14,16-12,13-11,14-9,9-11,13-6,6-8,8-10,10

http://www.cstl.nist.gov/biotech/strbase/training.htm
Data is Tabulated

<table>
<thead>
<tr>
<th>Locus</th>
<th>Allele</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSF1PO</td>
<td>17</td>
<td>0.2819</td>
</tr>
<tr>
<td>FGA</td>
<td>21</td>
<td>0.0894</td>
</tr>
<tr>
<td>TH01</td>
<td>12</td>
<td>0.1805</td>
</tr>
<tr>
<td>TPOX</td>
<td>28</td>
<td>0.0786</td>
</tr>
<tr>
<td>VWA</td>
<td>14</td>
<td>0.1327</td>
</tr>
<tr>
<td>D3S1358</td>
<td>16</td>
<td>0.1278</td>
</tr>
<tr>
<td>D5S818</td>
<td>15</td>
<td>0.1374</td>
</tr>
<tr>
<td>D16S539</td>
<td>14</td>
<td>0.2782</td>
</tr>
<tr>
<td>D7S820</td>
<td>11</td>
<td>0.3394</td>
</tr>
<tr>
<td>D13S317</td>
<td>11</td>
<td>0.3841</td>
</tr>
<tr>
<td>D5S818</td>
<td>16</td>
<td>0.1391</td>
</tr>
<tr>
<td>D18S51</td>
<td>28</td>
<td>0.2185</td>
</tr>
<tr>
<td>D21S11</td>
<td>14</td>
<td>0.1854</td>
</tr>
<tr>
<td>D8S1179</td>
<td>10</td>
<td>0.2152</td>
</tr>
<tr>
<td>FGA</td>
<td>81</td>
<td>0.2003</td>
</tr>
<tr>
<td>D2D19</td>
<td>9.17</td>
<td>0.2148</td>
</tr>
<tr>
<td>D2D19</td>
<td>16</td>
<td>0.1374</td>
</tr>
</tbody>
</table>

The Random Match Probability for this profile in the U.S. Caucasian population is 1 in 837 trillion. (10^12)

DNA Profile Frequency with all 13 CODIS STR loci

<table>
<thead>
<tr>
<th>Locus</th>
<th>Allele</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSF1PO</td>
<td>17</td>
<td>0.2819</td>
</tr>
<tr>
<td>FGA</td>
<td>21</td>
<td>0.0894</td>
</tr>
<tr>
<td>TH01</td>
<td>12</td>
<td>0.1805</td>
</tr>
<tr>
<td>TPOX</td>
<td>28</td>
<td>0.0786</td>
</tr>
<tr>
<td>VWA</td>
<td>14</td>
<td>0.1327</td>
</tr>
<tr>
<td>D3S1358</td>
<td>16</td>
<td>0.1278</td>
</tr>
<tr>
<td>D5S818</td>
<td>15</td>
<td>0.1374</td>
</tr>
<tr>
<td>D16S539</td>
<td>14</td>
<td>0.2782</td>
</tr>
<tr>
<td>D7S820</td>
<td>11</td>
<td>0.3394</td>
</tr>
<tr>
<td>D13S317</td>
<td>11</td>
<td>0.3841</td>
</tr>
<tr>
<td>D5S818</td>
<td>16</td>
<td>0.1391</td>
</tr>
<tr>
<td>D18S51</td>
<td>28</td>
<td>0.2185</td>
</tr>
<tr>
<td>D21S11</td>
<td>14</td>
<td>0.1854</td>
</tr>
<tr>
<td>D8S1179</td>
<td>10</td>
<td>0.2152</td>
</tr>
<tr>
<td>FGA</td>
<td>81</td>
<td>0.2003</td>
</tr>
<tr>
<td>D2D19</td>
<td>9.17</td>
<td>0.2148</td>
</tr>
<tr>
<td>D2D19</td>
<td>16</td>
<td>0.1374</td>
</tr>
</tbody>
</table>

The Random Match Probability for this profile in the U.S. Caucasian population is 1 in 837 trillion. (10^12)

Finally a case report is written based on tabulated STR genotype calls.

The Same 13 Locus STR Profile in Different Populations

1 in 837 trillion
1 in 0.84 quadrillion in U.S. Caucasian population (NIST)
1 in 2.46 quadrillion in U.S. Caucasian population (FBI)*
1 in 16.6 quadrillion in African American population (NIST)
1 in 17.6 quadrillion in African American population (FBI)*
1 in 18.0 quadrillion in U.S. Hispanic population (NIST)

These values are for unrelated individuals assuming no population substructure (using only p^2 and 2pq)

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

Crime Scene Evidence compared to Suspect(s) (Forensic Case)
Child compared to Alleged Father (Paternity Case)
Victim’s Remains compared to Biological Relative (Mass Disaster ID)
Soldier’s Remains compared to Direct Reference Sample (Armed Forces ID)

Steps in DNA Analysis

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collection</td>
<td>DNA is collected from the sample</td>
</tr>
<tr>
<td>Extraction</td>
<td>DNA is extracted from the sample</td>
</tr>
<tr>
<td>Quantitation</td>
<td>DNA is quantified</td>
</tr>
<tr>
<td>Genotyping</td>
<td>DNA is genotyped</td>
</tr>
<tr>
<td>Interpretation of Results</td>
<td>DNA results are interpreted</td>
</tr>
<tr>
<td>Database Storage &amp; Searching</td>
<td>DNA results are stored and searched</td>
</tr>
</tbody>
</table>

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

No names are associated with DNA profiles uploaded to NDIS if my profile was entered for searching:
16,17-18,21-22-12,14-28,30-14,16-12,13-11-14-9,9-11-6,6-8,8-10,10

Samples in State DNA Database

<table>
<thead>
<tr>
<th>State</th>
<th>As of Feb 2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michigan</td>
<td></td>
</tr>
</tbody>
</table>

As of February 2007 the profile composition of the National DNA Index System (NDIS):
Total number of profiles: 4,398,639
Total Forensic profiles: 167,103
Total Convicted Offender profiles: 4,231,536

http://www.fbi.gov/hq/lab/codis/mi.htm
Companies Supply Allelic Ladders in STR Kits to Aid Interlaboratory Consistency

Profiler Plus kit allelic ladders (Applied Biosystems)

PCR Product Size (bp)

Sample #1

Sample #2

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

 lvl

Crime Scene - Two Suspects

Suspect 1

Suspect 2

Evidence

Mixture vs Single Source Samples

Many possible combinations could have given rise to the particular mixture observed

Checks and Controls on DNA Results

Community

FBI DNA Advisory Board’s Quality Assurance Standards (also interlaboratory studies)

Laboratory

ASCLDLAB Audits and Accreditation

Analyst

Proficiency Tests & Continuing Education

Method/Instrument

Validation of Performance (along with traceable standard sample)

Protocol

Standard Operating Procedure is followed

Data Sets

Allelic ladders, positive and negative amplification controls, and reagent blanks are used

Individual Sample

Internal size standard present in every sample

Interpretation of Result

Second review by qualified analyst/supervisor

Court Presentation of Evidence

Defense attorneys and experts with power of discovery requests

http://www.cstl.nist.gov/biotech/strbase/training.htm
**Additional Challenges**

- Multiplex STR amplification requires a fairly narrow amount of input DNA to produce high quality results.
- High-throughput needs for databanking labs
  - Automated software for data review
- An attitude of being (and needing to be) “error-free”
- Separating biological fluids – perpetrator’s sperm from victim’s vaginal epithelial cells
- Mixture components can be difficult to decipher

**Mixtures: Issues and Challenges**


  - Mixtures arise when two or more individuals contribute to the sample being tested.
  - Mixtures can be challenging to detect and interpret without extensive experience and careful training. Even more challenging with poor quality data when degraded DNA is present...
  - Differential extraction can help distinguish male and female components of many sexual assault mixtures. Y-chromosome markers can help here in some cases...

**Mixtures: Issues and Challenges**

- Artifacts of PCR amplification such as stutter products and heterozygote peak imbalance complicate mixture interpretation.
- Thus, only a limited range of mixture component ratios can be solved routinely.

---

**Principles of Mixture Interpretation**

Most mixtures encountered in casework are 2-component mixtures arising from a combination of victim and perpetrator DNA profiles. Torres et al. (2003) *Forensic Sci. Int.* 134:180-186 examined 1,547 cases from 1997-2000 containing 2,424 typed samples of which 163 (6.7%) contained a mixed profile with only 8 (0.3%) coming from more than two contributors. 95.1% (155/163) were 2-component mixtures.

Ratios of the various mixture components stay fairly constant between multiple loci enabling deduction of the profiles for the major and minor components.

Some mixture interpretation strategies involve using victim (or other reference) alleles to help isolate obligate alleles coming from the unknown portion of the mixture.

---

**Example Mixture Data**

- Single Source Sample (Victim)
- Evidence Mixture (Victim + Perpetrator)

**Obligate Alleles (not present in the victim reference)**

<table>
<thead>
<tr>
<th>Allele</th>
<th>Victim = major</th>
<th>Perpetrator = minor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>D8S1179</td>
<td>28</td>
<td>16</td>
</tr>
<tr>
<td>D21S11</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>D18S51</td>
<td>28</td>
<td>16</td>
</tr>
<tr>
<td>Amelogenin</td>
<td>12,12</td>
<td>12,12</td>
</tr>
</tbody>
</table>

- **True “Perpetrator” Profile**

| Allele   | 12,12 | 28,31,2 | 15,16 |

---

**Mixtures: Issues and Challenges**

- Artifacts of PCR amplification such as stutter products and heterozygote peak imbalance complicate mixture interpretation.
- Thus, only a limited range of mixture component ratios can be solved routinely.
Mixtures: Issues and Challenges


- The probability that a mixture will be detected improves with the use of more loci and genetic markers that have a high incidence of heterozygotes.

- The detectability of multiple DNA sources in a single sample relates to the ratio of DNA present from each source, the specific combinations of genotypes, and the total amount of DNA amplified.

- Some mixtures will not be as easily detectable as other mixtures.

Two Parts to Mixture Interpretation

- Deduction of alleles present in the evidence (compared to victim and suspect profiles)

- Providing some kind of statistical answer regarding the weight of the evidence

- An ISFG DNA Commission (Peter Gill, Bruce Weir, Charles Brenner, etc.) is evaluating the statistical approaches to mixture interpretation and has made recommendations


ISFG Recommendations on Mixture Interpretation

Our discussions have highlighted a significant need for continuing education and research into this area.

DNA commission of the International Society of Forensic Genetics: Recommendations on the interpretation of mixtures

P. Gill*, C.B. Weir†, J.S. Buckleton‡, A. Carcamo§, M. Hovanciklá, W.R. May‡, N. Marling†, M. Peru‡, P.M. Schmidt‡, B.S. West†

*Forensic Science Unit, Royal Netherlands Academy of Sciences, Netherlands; †Medical Research Council Forensic Science Service, U.K.; ‡CSIRO, Division of Forensic & Applied Science, Australia; §Instituto de Investigaciones Forenses, Mexico

Reference elimination samples are useful in deciphering both situations.

Impact of Degraded DNA Samples

- Comparison to a phone number (string of 13 numbers) 001-301-975-4049

- If you only had ‘4049’…this information would be of limited value since it is not as specific (and could match other phone numbers from different area codes)

- DNA profiles are essentially a string of numbers – if the DNA is damaged, then the string of numbers is shorter and less informative…

How Can DNA Mixtures Arise?

- Two (or more) individuals contribute to the biological evidence examined in a forensic case (e.g., sexual assault with victim and perpetrator or victim, consensual sexual partner, and perp)

- Contamination of a single source sample from – evidence collection staff – laboratory staff handling the sample – Low-level DNA in reagents or PCR tubes or pipet tips

Examine Staff Profiles (Elimination Database). Maintain Contamination Log

Victim Reference and Spouse or Boyfriend Reference

Reference elimination samples are useful in deciphering both situations.

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

14
Contamination

- **Systematic**
  - e.g., Contaminated water or PCR buffer
- **Sporadic**
  - e.g., individual PCR tube contamination

- To reduce risks of contamination:
  - Careful lab cleanliness
  - Constant monitoring of reagents and consumables

- Contaminants are more likely to show up in the low molecular weight STR loci because they amplify more efficiently (miniSTRs will have a greater chance of detecting contaminating DNA)

- A negative control can detect systematic contamination but may not detect sporadic contamination, such as could be found in a single PCR tube

Impact of Contamination on Casework

- Use negative controls to predict the level of overall contamination in a lab
- Conclude that most likely outcome of a contamination event is a false exclusion...if contaminating DNA is preferentially amplified over original LCN material

Potential Impact of Contamination on Cold Cases or Post-Conviction Testing


- While this contamination possibility might only rarely impact a careful forensic DNA laboratory, it can have potential significance on old cases under review including the Innocence Project. For example, if biological evidence from a 20-year old case was handled by ungloved police officers or evidence custodians (prior to knowledge regarding the sensitivity of modern DNA testing), then the true perpetrator’s DNA might be masked by contamination from the collecting officer. Thus, when a DNA test is performed, the police officer’s or evidence custodian’s DNA would be detected rather than the true perpetrator. In the absence of other evidence, the individual in prison might then be falsely declared “innocent” because his DNA profile was not found on the original crime scene evidence. This scenario emphasizes the importance of considering DNA evidence as an investigative tool within the context of a case rather than the sole absolute proof of guilt or innocence.

How Are Such Large Numbers Generated with Random Match Probabilities?

- Each allele is sampled multiple times to produce a statistically stable allele frequency
- Using theoretical model from genetics, multiple loci are multiplied together to produce an estimate of the rarity of a particular DNA profile (combination of STR alleles based on individual allele frequencies)
- Remember that relatives will share genetic characteristics and thus have STR profiles that are more similar to one another than unrelated individuals
- We are not looking at every person on the planet nor are we looking at every nucleotide in the suspect’s genome

DNA Testing Has Become Extremely Sensitive...

- What does it mean to obtain a DNA match between a suspect and material from a crime scene?
- Is the fact that a DNA profile obtained mean that this information is probative?
- More complicated samples (mixtures) and more items per case being submitted to labs

Time Line Showing the Potential for DNA Deposition/Transfer

Adapted from Gill, P. (2002) BioTechniques 32(2): 385-385, Figure 5

http://www.cstl.nist.gov/biotech/strbase/training.htm
Some Final Thoughts

- "DNA" + "Match" ➜ “Guilty” in the minds of many jurors
- Be careful to state assumptions going into the weight of the evidence particularly for mixtures
- General population (i.e., jury pool) is becoming more informed regarding DNA testing thanks to genetic genealogy and TV shows like CSI
- Low-level DNA recovered from a crime scene may not be relevant to the committed crime