Aspects of Forensic DNA Typing using John M Butler’s slides

http://www.cstl.nist.gov/biotech/strbase/FDT2e.htm

Copyright symbols have been placed at the bottom of figures used directly from the book. The publisher requests that these copyright symbols are retained by slide users in their presentations.

This has been retained, but material has been added or removed to suit the purposes of the class. Please check the originals to see exactly what changes, has been made. Most other slides are from John Butler’s collection too.
Some generalities
• Report published in Nov 2000

• Asked to estimate where DNA testing would be 2, 5, and 10 years into the future

Conclusions

STR typing is here to stay for a few years because of DNA databases that have grown to contain millions of profiles
Chapter 1  Overview & History of DNA Typing
Chapter 2  DNA Biology Review
Chapter 3  Sample Collection, Extraction, Quantitation
Chapter 4  PCR Amplification
Chapter 5  Common STRs and Commercial Kits
Chapter 6  Biology of STRs
Chapter 7  Forensic Issues
Chapter 8  Single Nucleotide Polymorphisms
Chapter 9  Y-Chromosome DNA Tests
Chapter 10  Mitochondrial DNA Analysis
Chapter 11  Non-Human DNA and Microbial Forensics
Chapter 12  DNA Separation Methods
Chapter 13  DNA Detection Methods
Chapter 14  Instrumentation for STR Typing: ABI 310, ABI 3100, FMBIO
Chapter 15  STR Genotyping Issues
Chapter 16  Lab Validation
Chapter 17  New Technologies, Automation, and Expert Systems
Chapter 18  CODIS and DNA Databases
Chapter 19  Basic Genetic Principles and Statistics
Chapter 20  STR Database Analyses
Chapter 21  Profile Frequency Estimates
Chapter 22  Statistical Analysis of Mixtures and Degraded DNA
Chapter 23  Kinship and Paternity Testing
Chapter 24  Mass Disaster DNA Victim Identification
Appendix I  Reported STR Alleles
Appendix II  U.S. Population Data-STR Allele Frequencies
Appendix III  Suppliers of DNA Analysis Equipment
Appendix IV  DAB QA Standards
Appendix V  DAB Recommendations on Statistics
Appendix VI  Application of NRC II to STR Typing
Appendix VII  Example DNA Cases
Human Identity Testing

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

Involves generation of DNA profiles usually with the same core STR (short tandem repeat) markers
Basis of DNA Profiling

The genome of each individual is unique (with the exception of identical twins) and is inherited from parents.

Probe subsets of genetic variation in order to differentiate between individuals (statistical probabilities of a random match are used).

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

Current standard DNA tests DO NOT look at genes – little/no information about race, predisposition to disease, or phenotypical information (eye color, height, hair color) is obtained.
Some DNA marker types & technologies

- **Markers Used**
  - (Biology)
  - RFLP
  - Single Locus Probes
  - Multi-Locus Probes
  - Multiplex STRs
  - PolyMarker
  - mtDNA
  - ABO blood groups

- **Power of Discrimination** (Genetics)
  - High
  - Low

- **Speed of Analysis** (Technology)
  - Slow
  - Fast

Figure 1.1, J.M. Butler (2005) *Forensic DNA Typing, 2nd Edition* © 2005 Elsevier Academic Press
Steps in DNA sample analysis and interpretation

**Biology**

- DNA Extraction
- DNA Quantitation
- PCR Amplification of Multiple STR markers

**Technology**

- Separation and Detection of PCR Products (STR Alleles)
- Sample Genotype Determination

**Genetics**

- Comparison of Sample Genotype to Other Sample Results
- Generation of Case Report with Probability of Random Match
- If match occurs, comparison of DNA profile to population databases

Figure 1.2, J.M. Butler (2005) *Forensic DNA Typing, 2nd Edition* © 2005 Elsevier Academic Press
Comparison of two electropherograms

Quick review of some biology
Human Genome
23 Pairs of Chromosomes + mtDNA

Located in cell nucleus

Autosomes

2 copies per cell

Nuclear DNA
3.2 billion bp

Located in mitochondria (multiple copies in cell cytoplasm)

mtDNA
16,569 bp

Mitochondrial DNA

Sex-chromosomes

100s of copies per cell

Figure 2.3, J.M. Butler (2005) Forensic DNA Typing, 2nd Edition © 2005 Elsevier Academic Press
Figure 2.4, J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press
Two important polymorphic marker types

(A) Single nucleotide polymorphism (SNP)

--------AGAC\text{TAG}\text{ACATT}\--------
--------AGAT\text{TAGG}\text{CATT}\--------

(B) Short tandem repeat (STR) polymorphism

--------(AATG)(AATG)(AATG)--------
\text{3 repeats}

--------(AATG)(AATG)--------
\text{2 repeats}

Based on Figure 2.5, J.M. Butler (2005) *Forensic DNA Typing*, 2\textsuperscript{nd} Edition © 2005 Elsevier Academic Press
Figure 2.6, J.M. Butler (2005) *Forensic DNA Typing, 2nd Edition* © 2005 Elsevier Academic Press
Overview of the technology
Sources of Biological Evidence

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

Blood stain

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

**ORGANIC**

1. **Blood stain**
2. **SDS, DTT, EDTA and proteinase K**
3. **INCUBATE (56 °C)**
4. **Centrifuge**
5. **Phenol, chloroform, isoamyl alcohol**
6. **VORTEX**
7. **Centrifuge**
8. **TRANSFER aqueous (upper) phase to new tube**
9. **INCUBATE (ambient)**
10. **Centrifuge**
11. **REMOVE supernatant**
12. **5% Chelex**
13. **INCUBATE (56 °C)**
14. **INCUBATE (100 °C)**
15. **Centrifuge**
16. **CONCENTRATE sample**
17. **(Centricon/Microcon-100 or ethanol precipitation)**
18. **Centrifuge**
19. **QUANTITATE DNA**
20. **PERFORM PCR**

**CHELEX**

1. **Blood stain**
2. **Water**
3. **INCUBATE (ambient)**
4. **Centrifuge**
5. **REMOVE supernatant**
6. **Chelex**
7. **INCUBATE (56 °C)**
8. **INCUBATE (100 °C)**
9. **Centrifuge**
10. **QUANTITATE DNA**
11. **PERFORM PCR**

**FTA Paper**

1. **Apply blood to paper and allow stain to dry**
2. **PUNCH**
3. **WASH Multiple Times with extraction buffer**
4. **REMOVE supernatant**
5. **PCR Reagents**
6. **(NO DNA QUANTITATION TYPICALLY PERFORMED WITH UNIFORM SAMPLES)**
7. **PERFORM PCR**

**Figures:****Figure 3.1, J.M. Butler (2005) *Forensic DNA Typing, 2nd Edition © 2005 Elsevier Academic Press***
Slot blot for human DNA quantitation

Uses a 40 bp primate-specific DNA probe (for details see book)

Importance of DNA Quantitation  
(prior to multiplex PCR)

**DNA amount**  
(log scale)

- 100 ng
- 10 ng
- 1 ng
- 0.1 ng
- 0.01 ng

**High levels of DNA create interpretation challenges (more artifacts to review)**

- Off-scale peaks
- Split peaks (+/-A)
- Locus-to-locus imbalance

**STR Kits Work Best in This Range**

- 2.0 ng
- 0.5 ng

**Too much DNA**

- Stochastic effect when amplifying low levels of DNA produces allele dropout

**Too little DNA**

- Heterozygote peak imbalance
- Allele drop-out
- Locus-to-locus imbalance
Calculation of the quantity of DNA in a cell

1. Molecular Weight of a DNA Basepair = 618g/mol
   = 313 g/mol; T: 304 g/mol;  A-T base pairs = 617 g/mol
   G = 329 g/mol;  C: 289 g/mol;  G-C base pairs = 618 g/mol

2. Molecular weight of DNA = $1.85 \times 10^{12}$ g/mol
   There are 3 billion base pairs in a haploid cell ~3 x 10^9 bp
   (~3 x 10^9 bp) x (618 g/mol/bp) = $1.85 \times 10^{12}$ g/mol

3. Quantity of DNA in a haploid cell = 3 picograms
   1 mole = 6.02 x 1023 molecules
   (1.85 x 1012 g/mol) x (1 mole/6.02 x 1023 molecules)
   = 3.08 x 10^{-12} g = 3.08 picograms (pg)
   A diploid human cell contains ~6 pg genomic DNA

4. One ng of DNA contains the DNA from 167 diploid cells
   1 ng genomic DNA (1000 pg)/6pg/cell = ~333 copies of each locus
   (2 per 167 diploid genomes)
Brief review of the Polymerase Chain Reaction
Separate strands (denature)

Add primers (anneal)

Make copies (extend primers)

Repeat Cycle, Copying DNA Exponentially

Schematic: the precise details are different
The denaturation time in the first cycle is lengthened to ~10 minutes when using AmpliTaq Gold to perform a “hot-start” PCR.

Schematic for multiplex PCR

(A) Simultaneous amplification of three locations on a DNA template

(B) Resolution of PCR products with a single size-based separation method

Commonly used Short Tandem Repeat (STR) markers
Minisatellite Marker (D1S80)

- Flanking regions
- Repeat region
- GAGGACCACCAGGAAG
- 16 bp repeat unit

STR Marker (TH01)

- Flanking regions
- Repeat region
- TCAT
- 4 bp repeat unit

Figure 5.1, J.M. Butler (2005) *Forensic DNA Typing, 2nd Edition* © 2005 Elsevier Academic Press
Example of DNA sequence in an STR repeat region

Note different repeat motifs and starting positions on different strands. By convention, this is a TCAT repeat, read 5’ to 3’.
Commercially available STR kits for the 13 CODIS loci

AmpFISTR® Identifiler™ kit (Applied Biosystems)

6-FAM (Blue)  D8S1179  D21S11  D7S820  CSF1PO
VIC (Green)   D3S1358  TH01  D13S317  D16S539  D2S1338
NED (Yellow)  D19S433  VWA  TPOX  D18S51
PET (Red)     AMEL  D5S818  FGA
LIZ (Orange)   GS500 LIZ size standard

PowerPlex® 16 kit (Promega Corporation)

FL (Blue)      D3S1358  TH01  D21S11  D18S51
J5E (Green)    D5S818  D13S317  D7S820  D16S539  CSF1PO
TMR (Yellow)   AMEL  VWA  D8S1179  TPOX  FGA
CXR (Red)      ILS600 CXR size standard

Figure 5.4, J.M. Butler (2005) Forensic DNA Typing, 2nd Edition © 2005 Elsevier Academic Press
Figure 5.5, J.M. Butler (2005) *Forensic DNA Typing, 2nd Edition* © 2005 Elsevier Academic Press
Figure 5.6, J.M. Butler (2005) Forensic DNA Typing, 2nd Edition © 2005 Elsevier Academic Press
Amelogenin sex-typing assay

Female: X, X

Male: X, Y

1:1 Mixture: 3X + 1Y

Figure 5.11, J.M. Butler (2005) Forensic DNA Typing, 2nd Edition © 2005 Elsevier Academic Press
Short Tandem Repeat DNA Internet Data Base

These data are intended to benefit research and application of short tandem repeat DNA markers to human identity testing. The authors are solely responsible for the information herein.

Created by John M. Butler and Dennis J. Reeder (NIST Biotechnology Division), with invaluable help from Jan Redman, Christian Ruitberg and Michael Tung

*Partial support for the design and maintenance of this website is being provided by The National Institute of Justice through the NIST Office of Law Enforcement Standards.*

Publications and Presentations from NIST Human Identity Project Team

- STRs101: Brief Introduction to STRs
- STR Fact Sheets (observed alleles and PCR product sizes)
- Sequence Information (annotated)
- Multiplex STR sets
- STR Training Materials
- Non-published Variant Allele Reports
- Three-Banded Patterns
- FBI CODIS Core STR Loci
- DNA Advisory Board Quality Assurance Standards
- NIST Standard Reference Material for PCR-Based Testing
- Chromosomal Locations
- Mutation Rates for Common Loci
- Published PCR primers
- Validation studies
- Population data
- Y-chromosome STRs
- Sex-typing markers
- Technology for resolving STR alleles
- Reference List Now 2059 references
- Addresses for scientists working with STRs
- Links to other web sites
- Glossary of commonly used terms

Figure 5.12, J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press
STR alleles with stutter products

DNA Size (bp)

(A) Normal replication

(B) Insertion caused by backward slippage

(C) Deletion caused by forward slippage

Non-template nucleotide addition

(A)

Forward Primer → Polymerase extension → Reverse Primer

5' → 3' → 5'

Polymerase extension

OR

5' 5'

(-A form) (+A form)

5' A 5'

(B)

Measurement Result with dye labeled DNA strand

Incomplete adenylation

+ A + A - A - A

Full-length allele (n)

allele + 1 base (n+1)

Shoulder peak

Split peak

D8S1179

Incomplete non-template addition with high levels of DNA template

DNA Size (bp)

Relative Fluorescence (RFUs)

- A
+ A

off-scale

10 ng template (overloaded)

D3S1358  VWA  FGA

2 ng template (suggested level)

Detection of a microvariant allele at the locus FGA

\[ \delta_1 = S_{25} - L_{25} = 244.34 - 244.46 = -0.12 \text{ bp} \]

\[ \delta_2 = S_{OL} - L_{28} = 257.51 - 256.64 = +0.87 \text{ bp} \]

\[ c = |\delta_1 - \delta_2| = |-0.12 - 0.87| = 0.99 \text{ bp} \]

Tri-allelic patterns at TPOX and D18S51
Possible sequence variation (*) and impact on PCR amplification

A)

Example:
TH01 9.3 allele
(-A in 7th repeat)

Amplicon size may be nonstandard

B)

Example:
D18S51 13.2 allele
(+AG in 3'-flanking region)

Amplicon size may be nonstandard

C)

Example:
Rare VWA allele amplified with AmpFISTR primers
(A-to-T in 2nd base from 3'end of forward primer)

PCR may fail

Impact of sequence polymorphism in the primer binding site

Heterozygous alleles are well balanced

Imbalance in allele peak heights

Allele 6 amplicon has "dropped out"

Mutational events observed in family trees

(a) 14,18  →  15,17  ↓  15,18
(b) 14,18  →  15,17  ↓  13,17

## Mutation Rates for Common STR Loci

http://www.aabb.org/About_the_AABB/Std
ds_and_Accred/ptannrpt03.pdf, Appendix 2

<table>
<thead>
<tr>
<th>STR System</th>
<th>Maternal Meioses (%)</th>
<th>Paternal Meioses (%)</th>
<th>Number from either</th>
<th>Total Number of Mutations</th>
<th>Mutation Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSF1PO</td>
<td>95/304,307 (0.03)</td>
<td>982/643,118 (0.15)</td>
<td>410</td>
<td>1,487/947,425</td>
<td>0.16%</td>
</tr>
<tr>
<td>FGA</td>
<td>205/408,230 (0.05)</td>
<td>2,210/692,776 (0.32)</td>
<td>710</td>
<td>3,125/1,101,006</td>
<td>0.28%</td>
</tr>
<tr>
<td>TH01</td>
<td>31/327,172 (0.009)</td>
<td>41/452,382 (0.009)</td>
<td>28</td>
<td>100/779,554</td>
<td>0.01%</td>
</tr>
<tr>
<td>TPOX</td>
<td>18/400,061 (0.004)</td>
<td>54/457,420 (0.012)</td>
<td>28</td>
<td>100/857,481</td>
<td>0.01%</td>
</tr>
<tr>
<td>VWA</td>
<td>184/564,398 (0.03)</td>
<td>1,482/873,547 (0.17)</td>
<td>814</td>
<td>2,480/1,437,945</td>
<td>0.17%</td>
</tr>
<tr>
<td>D3S1358</td>
<td>60/405,452 (0.015)</td>
<td>713/558,836 (0.13)</td>
<td>379</td>
<td>1,152/964,288</td>
<td>0.12%</td>
</tr>
<tr>
<td>D5S818</td>
<td>111/451,736 (0.025)</td>
<td>763/655,603 (0.12)</td>
<td>385</td>
<td>1,259/1,107,339</td>
<td>0.11%</td>
</tr>
<tr>
<td>D7S820</td>
<td>59/440,562 (0.013)</td>
<td>745/644,743 (0.12)</td>
<td>285</td>
<td>1,089/1,085,305</td>
<td>0.10%</td>
</tr>
<tr>
<td>D8S1179</td>
<td>96/409,869 (0.02)</td>
<td>779/489,968 (0.16)</td>
<td>364</td>
<td>1,239/899,837</td>
<td>0.14%</td>
</tr>
<tr>
<td>D13S317</td>
<td>192/482,136 (0.04)</td>
<td>881/621,146 (0.14)</td>
<td>485</td>
<td>1,558/1,103,282</td>
<td>0.14%</td>
</tr>
<tr>
<td>D16S539</td>
<td>129/467,774 (0.03)</td>
<td>540/494,465 (0.11)</td>
<td>372</td>
<td>1,041/962,239</td>
<td>0.11%</td>
</tr>
<tr>
<td>D18S51</td>
<td>186/296,244 (0.06)</td>
<td>1,094/494,098 (0.22)</td>
<td>466</td>
<td>1,746/790,342</td>
<td>0.22%</td>
</tr>
<tr>
<td>D21S11</td>
<td>464/435,388 (0.11)</td>
<td>772/526,708 (0.15)</td>
<td>580</td>
<td>1,816/962,096</td>
<td>0.19%</td>
</tr>
<tr>
<td>Penta D</td>
<td>12/18,701 (0.06)</td>
<td>21/22,501 (0.09)</td>
<td>24</td>
<td>57/41,202</td>
<td>0.14%</td>
</tr>
<tr>
<td>Penta E</td>
<td>29/44,311 (0.065)</td>
<td>75/55,719 (0.135)</td>
<td>59</td>
<td>163/100,030</td>
<td>0.16%</td>
</tr>
<tr>
<td>D2S1338</td>
<td>15/72,830 (0.021)</td>
<td>157/152,310 (0.10)</td>
<td>90</td>
<td>262/225,140</td>
<td>0.12%</td>
</tr>
<tr>
<td>D19S433</td>
<td>38/70,001 (0.05)</td>
<td>78/103,489 (0.075)</td>
<td>71</td>
<td>187/173,490</td>
<td>0.11%</td>
</tr>
<tr>
<td>SE33 (ACTBP2)</td>
<td>0/330 (&lt;0.30)</td>
<td>330/51,610 (0.64)</td>
<td>None reported</td>
<td>330/51,940</td>
<td>0.94%</td>
</tr>
</tbody>
</table>
Some special issues arising with forensic DNA typing
Illustration of typical single source vs mixed sample

(a) Heterozygous peak region

MIXTURE REGION

Stutter region

(b) Higher than typical stutter product (>15%)

Smaller peak area than normally seen with heterozygote partner alleles (<70%)

Wrong side of allele to be typical stutter product

Some of the DNA technology
Schematic of gel electrophoresis

*Side view*

*Top view*

Sample tray moves automatically beneath the cathode end of the capillary to deliver each sample in succession.

Larger DNA molecules interact more frequently with the gel and are thus retarded in their migration through the gel.
Illustration of fluorescence & excitation/emission spectra

Dyes used by ABI in 4-color detection

**FAM**
(blue)

**JOE**
(green)

**TAMRA**
(yellow)

**ROX**
(red)

Emission spectra of ABI dyes and wavelength bands used for detection

5-FAM  JOE  NED  ROX

Normalized Fluorescent Intensity

WAVELENGTH (nm)

Emission spectra of ABI dyes and wavelength bands used for detection

Laser excitation
(488, 514.5 nm)

310 Filter Set F
with color contributions

Example of a matrix used for color separation

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>G</th>
<th>Y</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>1.0000</td>
<td>0.8502</td>
<td>0.1380</td>
<td>0.0009</td>
</tr>
<tr>
<td>G</td>
<td>0.8300</td>
<td>1.0000</td>
<td>0.7622</td>
<td>0.0051</td>
</tr>
<tr>
<td>Y</td>
<td>0.6416</td>
<td>0.8324</td>
<td>1.0000</td>
<td>0.1102</td>
</tr>
<tr>
<td>R</td>
<td>0.4493</td>
<td>0.6484</td>
<td>0.7851</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Figure 13.6, J.M. Butler (2005) *Forensic DNA Typing, 2nd Edition* © 2005 Elsevier Academic Press
Traces before and after color separation (and perhaps other processing)

Mixture of dye-labeled PCR products from multiplex PCR reaction

Sample Injection

Size Separation

Sample Separation

Argon ion LASER (488 nm)

Fluorescence

Capillary

ABI Prism spectrograph

Color Separation

CCD Panel (with virtual filters)

Sample Detection

Processing with GeneScan/Genotyper software

Sample Interpretation

Schematic of the separation and detection process

Traces from ABI 310 capillary electrophoresis system

Profiler Plus™ multiplex STR result

DNA size (bp)

COfiler™ multiplex STR result

DNA size (bp)

Image of data from Fig 14.9 below
Deciphering Artifacts from the True Alleles

**Biological (PCR) artifacts**

- **Stutter products**
  - D3S1358
  - Incomplete adenylation
    - D8S1179

**STR alleles**

- Dye blob
- Stutter
- Pull-up (bleed-through)
- Spike

**Channels**

- Blue channel
- Green channel
- Yellow channel
- Red channel

STR population databases and statistical calculations

This material which follows is presented without comment as illustrative of the approach taken by law enforcement agencies.
Decide on Number of Samples and Ethnic/Racial Grouping

Gather Samples

Get IRB approval
Often anonymous samples from a blood bank

Analyze Samples at Desired Genetic Loci

Summarize DNA types

See Chapter 5 (STR kits available) and Chapter 15 (STR typing/interpretation)

Determine Allele Frequencies for Each Locus

See Table 20.2 and Appendix II

Perform Statistical Tests on Data

Hardy-Weinberg equilibrium for allele independence
Linkage equilibrium for locus independence

Ethnic/Racial Group 1

Ethnic/Racial Group 2

Examination of genetic distance between populations

Use Database(s) to Estimate an Observed DNA Profile Frequency

See Chapter 21

Usually >100 per group (see Table 20.1)
U.S. Population Samples
(Appendix II)

How Statistical Calculations are Made

• **Generate data** with set(s) of samples from desired population group(s)
  – Generally only 100-150 samples are needed to obtain reliable allele frequency estimates

• **Determine allele frequencies** at each locus
  – Count number of each allele seen

• **Allele frequency information** is used to **estimate the rarity of a particular DNA profile**
  – Homozygotes \( p^2 \), Heterozygotes \( 2pq \)
  – Product rule used (multiply locus frequency estimates)

For more information, see Chapters 20 and 21 in *Forensic DNA Typing, 2nd Edition*
Assumptions with Hardy-Weinberg Equilibrium

<table>
<thead>
<tr>
<th>The Assumption</th>
<th>The Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large population</td>
<td>Lots of possible allele combinations</td>
</tr>
<tr>
<td>No natural selection</td>
<td>No restriction on mating so all alleles have equal chance of becoming part of next generation</td>
</tr>
<tr>
<td>No mutation</td>
<td>No new alleles being introduced</td>
</tr>
<tr>
<td>No immigration/emigration</td>
<td>No new alleles being introduced or leaving</td>
</tr>
<tr>
<td>Random mating</td>
<td>Any allele combination is possible</td>
</tr>
</tbody>
</table>

None of these assumptions are really true…

Individual Genotypes Are Summarized and Converted into Allele Frequencies

<table>
<thead>
<tr>
<th>Genotype Array</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>Allele Count</th>
<th>Observed Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>8,8</td>
<td>8,9</td>
<td>8,10</td>
<td>8,11</td>
<td>8,12</td>
<td>8,13</td>
<td>8,14</td>
<td>8,15</td>
<td>68</td>
<td>0.11258</td>
</tr>
<tr>
<td>9</td>
<td>9,9</td>
<td>9,10</td>
<td>9,11</td>
<td>9,12</td>
<td>9,13</td>
<td>9,14</td>
<td>9,15</td>
<td></td>
<td>45</td>
<td>0.07450</td>
</tr>
<tr>
<td>10</td>
<td>11</td>
<td>10,11</td>
<td>10,12</td>
<td>10,13</td>
<td>10,14</td>
<td>10,15</td>
<td></td>
<td></td>
<td>31</td>
<td>0.05132</td>
</tr>
<tr>
<td>11</td>
<td>37</td>
<td>54</td>
<td>21</td>
<td>12</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td>205</td>
<td>0.33940</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td>12,12</td>
<td>12,13</td>
<td>12,14</td>
<td>12,15</td>
<td></td>
<td>150</td>
<td>0.24834</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13,13</td>
<td>13,14</td>
<td>13,15</td>
<td>7</td>
<td>75</td>
<td>0.12417</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>29</td>
<td>0.04801</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>0.00166</td>
</tr>
</tbody>
</table>

The **11,12** genotype was seen **54 times** in 302 samples (604 examined chromosomes)

## Allele Frequency Tables

**Allele frequencies denoted with an asterisk (*) are below the 5/2N minimum allele threshold recommended by the National Research Council report (NRCII) *The Evaluation of Forensic DNA Evidence* published in 1996.**

### D3S1358

<table>
<thead>
<tr>
<th>Allele</th>
<th>Caucasian N= 302</th>
<th>Caucasian N= 7,636</th>
<th>African American N=258</th>
<th>African American N= 7,602</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>0.0017*</td>
<td>0.0009</td>
<td>11</td>
<td>0.0003*</td>
</tr>
<tr>
<td>12</td>
<td>0.0017*</td>
<td>0.0007</td>
<td>12</td>
<td>0.0045</td>
</tr>
<tr>
<td>13</td>
<td>--</td>
<td>0.0031</td>
<td>13</td>
<td>0.0077</td>
</tr>
<tr>
<td>14</td>
<td>0.1027</td>
<td>0.1240</td>
<td>14</td>
<td>0.0905</td>
</tr>
<tr>
<td>15</td>
<td>0.2616</td>
<td>0.2690</td>
<td>15</td>
<td>0.2920</td>
</tr>
<tr>
<td>15.2</td>
<td>--</td>
<td>--</td>
<td>15.2</td>
<td>0.0010</td>
</tr>
<tr>
<td>16</td>
<td>0.2533</td>
<td>0.2430</td>
<td>16</td>
<td>0.3353</td>
</tr>
<tr>
<td>17</td>
<td>0.2152</td>
<td>0.2000</td>
<td>17</td>
<td>0.3300</td>
</tr>
<tr>
<td>18</td>
<td>0.15232</td>
<td>0.1460</td>
<td>18</td>
<td>0.2070</td>
</tr>
<tr>
<td>19</td>
<td>0.01160</td>
<td>0.0125</td>
<td>19</td>
<td>0.0630</td>
</tr>
<tr>
<td>20</td>
<td>0.0017*</td>
<td>0.0001*</td>
<td>20</td>
<td>0.0048</td>
</tr>
</tbody>
</table>

*Most common allele*
Illustrative calculations
## DNA Profile Frequency with all 13 CODIS STR loci

<table>
<thead>
<tr>
<th>Locus</th>
<th>allele</th>
<th>value</th>
<th>allele</th>
<th>value</th>
<th>1 in</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>D3S1358</td>
<td>16</td>
<td>0.2533</td>
<td>17</td>
<td>0.2152</td>
<td>9.17</td>
<td>9.17</td>
</tr>
<tr>
<td>VWA</td>
<td>17</td>
<td>0.2815</td>
<td>18</td>
<td>0.2003</td>
<td>8.87</td>
<td>81</td>
</tr>
<tr>
<td>FGA</td>
<td>21</td>
<td>0.1854</td>
<td>22</td>
<td>0.2185</td>
<td>12.35</td>
<td>1005</td>
</tr>
<tr>
<td>D8S1179</td>
<td>12</td>
<td>0.1854</td>
<td>14</td>
<td>0.1656</td>
<td>16.29</td>
<td>16,364</td>
</tr>
<tr>
<td>D21S11</td>
<td>28</td>
<td>0.1589</td>
<td>30</td>
<td>0.2782</td>
<td>11.31</td>
<td>185,073</td>
</tr>
<tr>
<td>D18S51</td>
<td>14</td>
<td>0.1374</td>
<td>16</td>
<td>0.1391</td>
<td>26.18</td>
<td>4,845,217</td>
</tr>
<tr>
<td>D5S818</td>
<td>12</td>
<td>0.3841</td>
<td>13</td>
<td>0.1407</td>
<td>9.25</td>
<td>44,818,259</td>
</tr>
<tr>
<td>D13S317</td>
<td>11</td>
<td>0.3394</td>
<td>14</td>
<td>0.0480</td>
<td>30.69</td>
<td>1.38 x 10^9</td>
</tr>
<tr>
<td>D7S820</td>
<td>9</td>
<td>0.1772</td>
<td></td>
<td></td>
<td>31.85</td>
<td>4.38 x 10^{10}</td>
</tr>
<tr>
<td>D16S539</td>
<td>9</td>
<td>0.1126</td>
<td>11</td>
<td>0.3212</td>
<td>13.8</td>
<td>6.05 x 10^{11}</td>
</tr>
<tr>
<td>THO1</td>
<td>6</td>
<td>0.2318</td>
<td></td>
<td></td>
<td>18.62</td>
<td>1.13 x 10^{13}</td>
</tr>
<tr>
<td>TPOX</td>
<td>8</td>
<td>0.5348</td>
<td></td>
<td></td>
<td>3.50</td>
<td>3.94 x 10^{13}</td>
</tr>
<tr>
<td>CSF1PO</td>
<td>10</td>
<td>0.2169</td>
<td></td>
<td></td>
<td>21.28</td>
<td>8.37 x 10^{14}</td>
</tr>
</tbody>
</table>

The Random Match Probability for this profile in the U.S. Caucasian population is **1 in 837 trillion (10^{12})**
The Same 13 Locus STR Profile in Different Populations

1 in 837 trillion

1 in 0.84 quadrillion ($10^{15}$) in U.S. Caucasian population (NIST)
1 in 2.46 quadrillion ($10^{15}$) in U.S. Caucasian population (FBI)*
1 in 1.86 quadrillion ($10^{15}$) in Canadian Caucasian population*

1 in 16.6 quadrillion ($10^{15}$) in African American population (NIST)
1 in 17.6 quadrillion ($10^{15}$) in African American population (FBI)*

1 in 18.0 quadrillion ($10^{15}$) in U.S. Hispanic population (NIST)

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


*http://www.csfs.ca/pplus/profiler.htm
### STR Cumulative Profile Frequency with Multiple Population Databases

<table>
<thead>
<tr>
<th>STR Locus</th>
<th>Profile Computed</th>
<th>Number of Populations Used</th>
<th>Cumulative Profile Frequency Range (1 in ...)</th>
<th>Cumulative Profile Frequency against U.S. Caucasians (Appendix II)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D3S1358</td>
<td>16,17</td>
<td>166</td>
<td>5.24 to 62.6</td>
<td>9.19</td>
</tr>
<tr>
<td>VWA</td>
<td>17,18</td>
<td>166</td>
<td>37.6 to 1080</td>
<td>81.8</td>
</tr>
<tr>
<td>FGA</td>
<td>21,22</td>
<td>166</td>
<td>737 to 119 000</td>
<td>1010</td>
</tr>
<tr>
<td>D8S1179</td>
<td>12,14</td>
<td>166</td>
<td>8980 to 5 430 000</td>
<td>16 400</td>
</tr>
<tr>
<td>D21S11</td>
<td>28,30</td>
<td>166</td>
<td>165 000 to 248 000 000</td>
<td>186 000</td>
</tr>
<tr>
<td>D18S51</td>
<td>14,16</td>
<td>166</td>
<td>$3.85 \times 10^6$ to $2.68 \times 10^{10}$</td>
<td>$4.88 \times 10^6$</td>
</tr>
<tr>
<td>D5S818</td>
<td>12,13</td>
<td>166</td>
<td>$2.28 \times 10^7$ to $4.22 \times 10^{11}$</td>
<td>$4.51 \times 10^7$</td>
</tr>
<tr>
<td>D13S317</td>
<td>11,14</td>
<td>166</td>
<td>$4.32 \times 10^8$ to $1.69 \times 10^{13}$</td>
<td>$1.38 \times 10^9$</td>
</tr>
<tr>
<td>D7S820</td>
<td>9,9</td>
<td>166</td>
<td>$1.17 \times 10^{10}$ to $2.98 \times 10^{16}$</td>
<td>$4.22 \times 10^{10}$</td>
</tr>
<tr>
<td>D16S539</td>
<td>9,11</td>
<td>97</td>
<td>$4.06 \times 10^{11}$ to $1.11 \times 10^{18}$</td>
<td>$5.82 \times 10^{11}$</td>
</tr>
<tr>
<td>TH01</td>
<td>6,6</td>
<td>97</td>
<td>$9.30 \times 10^{12}$ to $1.45 \times 10^{19}$</td>
<td>$1.05 \times 10^{13}$</td>
</tr>
<tr>
<td>TPOX</td>
<td>8,8</td>
<td>97</td>
<td>$3.33 \times 10^{13}$ to $1.54 \times 10^{20}$</td>
<td>$3.63 \times 10^{13}$</td>
</tr>
<tr>
<td>CSF1PO</td>
<td>10,10</td>
<td>97</td>
<td>$3.43 \times 10^{14}$ to $2.65 \times 10^{21}$</td>
<td>$7.43 \times 10^{14}$</td>
</tr>
</tbody>
</table>

$10^{14}$ to $10^{21}$
Example Calculations with Population Substructure Adjustments

Table 21.5
Example calculations with NRC II recommendations for population substructure adjustments (see Appendix VI). Scenarios with theta equal to 0.01 and 0.03 are examined.

<table>
<thead>
<tr>
<th>From U.S. Caucasian (N=302): Appendix II - sample in database</th>
<th>Under HWE</th>
<th>NRCII Recommendation 4.1</th>
<th>NRCII Recommendation 4.10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A1</td>
<td>A2</td>
<td>Allele 1 freq (p)</td>
</tr>
<tr>
<td>D135317</td>
<td>11</td>
<td>14</td>
<td>0.33940</td>
</tr>
<tr>
<td>TH01</td>
<td>6</td>
<td>6</td>
<td>0.23179</td>
</tr>
<tr>
<td>D18551</td>
<td>14</td>
<td>16</td>
<td>0.13742</td>
</tr>
<tr>
<td>D21511</td>
<td>28</td>
<td>30</td>
<td>0.15894</td>
</tr>
<tr>
<td>D351358</td>
<td>16</td>
<td>17</td>
<td>0.25331</td>
</tr>
<tr>
<td>D55818</td>
<td>12</td>
<td>13</td>
<td>0.38411</td>
</tr>
<tr>
<td>D75820</td>
<td>9</td>
<td>9</td>
<td>0.17715</td>
</tr>
<tr>
<td>D851179</td>
<td>12</td>
<td>14</td>
<td>0.18543</td>
</tr>
<tr>
<td>CSF1PO</td>
<td>10</td>
<td>10</td>
<td>0.21689</td>
</tr>
<tr>
<td>FGA</td>
<td>21</td>
<td>22</td>
<td>0.18543</td>
</tr>
<tr>
<td>D165539</td>
<td>9</td>
<td>11</td>
<td>0.11258</td>
</tr>
<tr>
<td>TPOX</td>
<td>8</td>
<td>8</td>
<td>0.53477</td>
</tr>
<tr>
<td>VWA</td>
<td>17</td>
<td>18</td>
<td>0.28146</td>
</tr>
<tr>
<td>AMEL</td>
<td>X</td>
<td>Y</td>
<td></td>
</tr>
</tbody>
</table>

1.20E-15  1.35E-15  1.70E-15  3.92E-15
### Table 21.6

Example calculations with corrections for relatives using the NRC II recommended formula.

<table>
<thead>
<tr>
<th>From U.S. Caucasian (N = 302); Appendix II – sample In database</th>
<th>Under HWE</th>
<th>NRCII Recommendation 4.4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A1</td>
<td>A2</td>
</tr>
<tr>
<td>D135317</td>
<td>11</td>
<td>14</td>
</tr>
<tr>
<td>TH01</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>D165539</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>D18551</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>D21511</td>
<td>28</td>
<td>30</td>
</tr>
<tr>
<td>D351358</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>D55818</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>D75820</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>D851179</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>CSG1PO</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>FGA</td>
<td>21</td>
<td>22</td>
</tr>
<tr>
<td>TPOX</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>VWA</td>
<td>17</td>
<td>18</td>
</tr>
</tbody>
</table>

1.20E–15 | 3.17E–09 | 1.68E–11 | 3.74E–13 |