Mixtures and Low-Level DNA:

Addressing Questions Regarding Forensic DNA Typing

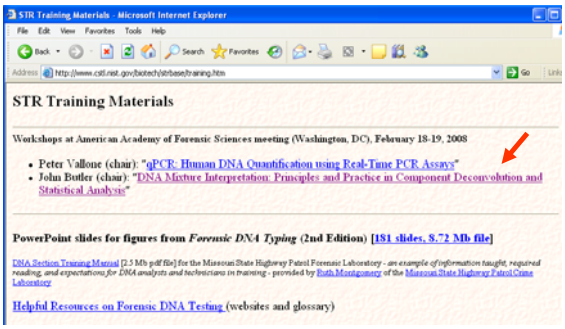
John M. Butler, Ph.D.
National Institute of Standards and Technology

NDAA Training at the National Advocacy Center
Columbia, SC
May 20, 2009

True Identity: DNA
DNA Fingerprinting on the Witness Stand

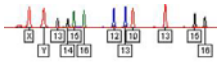
Training Materials Available on STRBase

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

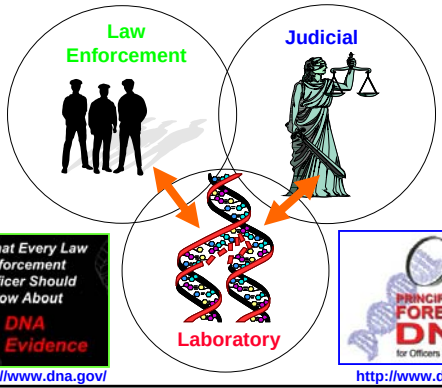


My Desire to Aid the Legal Community

- Growing up, I wanted to be a lawyer…until I took some science classes in high school (and my uncle who is a lawyer encouraged me not to since there were already too many lawyers)
- I served on the Grand Jury in Montgomery County, Maryland from June to August 2003 – but we only had three DNA cases…
- I have been contacted in the past by prosecutors, defense attorneys, and judges…I am happy to help where I can but the NIST lawyers have told me that I cannot testify – but in a sense my book and website testify for me
- I am committed to developing better training tools for lawyers to help with the proper use of DNA evidence



Interfaces Between Disciplines Are Crucial



http://www.dna.gov/

http://www.dna.gov/

Information Resources for Defense Attorneys

http://www.nlada.org/Defender/forensics/for_lib/Index/DNA/exhibits/index_html

NATIONAL LEGAL AID & DEFENDER ASSOCIATION

Defense Lawyers and Experts are becoming more united and informed

Common Defense Attacks

Compiled from Forensic Bioinformatics website

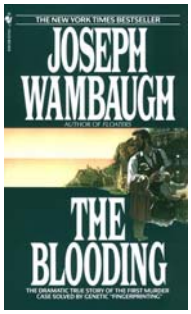
forensic bioinformatics
http://www.bioforensics.com

- Contamination
- Statistical Weight of a Match
- Degradation/PCR Inhibition of "True" Perp
- Artifacts and "Complications" to DNA Profiles
- **Thresholds Set Too High (missing peaks)**
- Examiner Bias
- **Improper Mixture Interpretation**
- Meaning of a Database Hit
- Protocol Violations

Forensic Bioinformatics
7th Annual Conference
The Science of DNA Profiling: A National Expert Forum
August 15 - 17, 2008
Dayton, OH

See http://www.bioforensics.com/conference08/index.html

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.

NIST Background

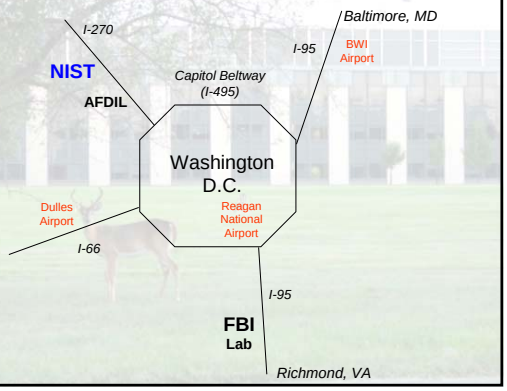
NIST History and Mission

- 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.**

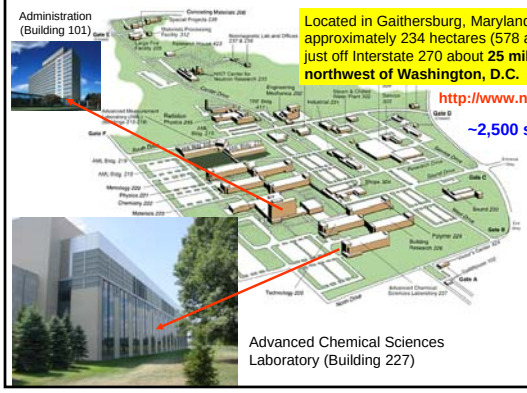
$603 for 3 jars

DNA typing standard

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.nist.gov

~2,500 staff

NIST Human Identity Project Team

…Bringing traceability and technology to the scales of justice…

John Butler
Group Leader

Amy Decker

Becky Hill

Margaret Kline

Jan Redman

Pete Vallone

Dave Duewer
(data analysis)

Angie Dolph
(summer 2007)

Michelle Burns
(summer 2008+)

And many wonderful collaborators…

Since 2000:
>100 publications
>250 presentations
>30 training workshops

Funding from the **National Institute of Justice (NIJ)** through NIST Office of Law Enforcement Standards

Our team publications and presentations are available at:
http://www.cstl.nist.gov/biotech/strbase/NISTpub.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.

National Institute of Justice

The Research, Development, and Evaluation Agency of the U.S. Department of Justice

Current Areas of NIST Effort with Forensic DNA

- **Standards** http://www.cstl.nist.gov/biotech/strbase/
  - 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

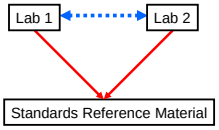
Standard Reference Materials

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

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

Helps meet DAB Std. 9.5 and ISO 17025

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

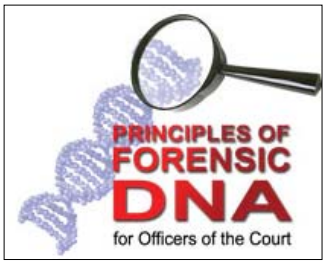
http://www.cstl.nist.gov/biotech/strbase

Includes information on:
- Core STR loci
- Validation
- STR reference list
- NIST publications
- miniSTRs
- Forensic SNPs
- Variant STR alleles
- Population data resources
- Addresses of scientists

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

DNA Training for Officers of the Court

PRESIDENT'S DNA INITIATIVE *Advancing Justice Through DNA Technology*



- CD-ROM available from the U.S. National Institute of Justice (http://www.ncjrs.gov)
- On-line training available at http://www.DNA.gov

http://www.dna.gov/training/otc/

Principles of Forensic DNA for Officers of the Court

PRESIDENT'S DNA INITIATIVE

1. **Introduction**
2. **Biology of DNA**
3. **Practical Issues Specific to DNA Evidence**
4. **Forensic DNA Laboratory**
5. **Assuring Quality in DNA Testing**
6. **Understanding a Forensic DNA Lab Report**
7. **Statistics and Population Genetics**
8. **Mitochondrial DNA & Y-STR Analysis**
9. **Forensic DNA Databases**
10. **Collection of DNA Evidence**
11. **Pretrial DNA Evidence Issues**
12. **Victim Issues**
13. **Trial Presentation**
14. **Postconviction DNA Cases**
15. **Emerging Trends**

http://www.dna.gov/training/otc/

Content of Section 15 "Emerging Trends" from *Officers of the Court*

- Topic 1 :: **Single Nucleotide Polymorphisms (SNPs)**
- Topic 2 :: Automation
  - Microarrays (Chip Technology)
  - **Portable DNA Typing Laboratory**
  - Low Copy Number DNA Analysis
- Topic 3 :: Microbial Forensics and DNA Testing
- Topic 4 :: Other **Non-human Forensic DNA Analysis**
- Topic 5 :: **DNA Typing and Physical Appearance**
  - Biogeographical Ancestry
  - Approximate Age Determination

http://www.dna.gov/training/otc/

How Are DNA Results Obtained?

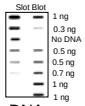
Steps in DNA Analysis

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

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


Blood Stain Buccal swab
Sample Collection & Storage


DNA Extraction



DNA Quantitation

Biology

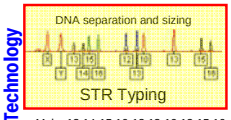

Multiplex PCR Amplification

Genetics

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


DNA Database Search

Technology



Male: 13,14-15,16-12,13-10,13-15,16

Interpretation of Results

Short Tandem Repeat (STR) Markers

An accordion-like DNA sequence that occurs between genes

TCCCAAGCTCTTCCTCTTCCCTAGATCAATACAGACAGAAGACA
GGTG**GATAGATAGATAGATAGATAGATAGATAGATAGATAGA**
TAGATAGATATCATTGAAAGACAAAACAGAGATGGATGATAGAT
ACATGCTTACAGATGCACAC

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

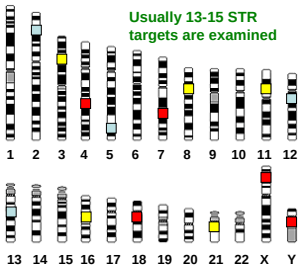
Target region
(short tandem repeat)

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

What is a DNA Profile?

Human Genome
23 Pairs of Chromosomes (~3 billion bp)

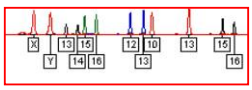


Unique regions of the human genome are targeted

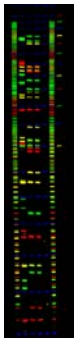
These regions consist of a few hundred base pairs

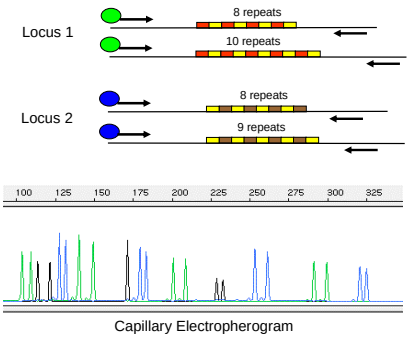
The regions are copied by the **polymerase chain reaction (PCR)** – billions of exact copies are made

The copied fragments are labeled with fluorescent dyes for detection purposes



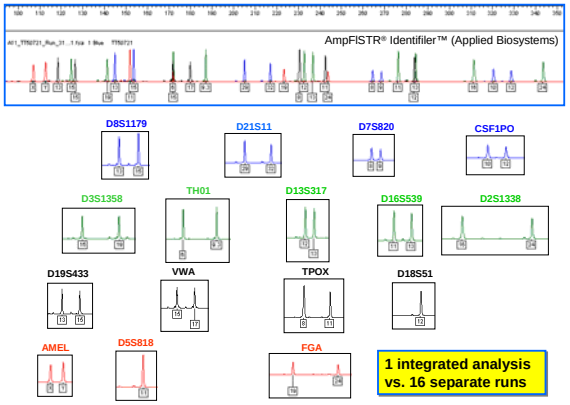
STR Allele Separation Can Be Performed by Gel or Capillary Electrophoresis with Detection of Fluorescent Dyes Labeling Each PCR Product


Scanned **Gel Image**



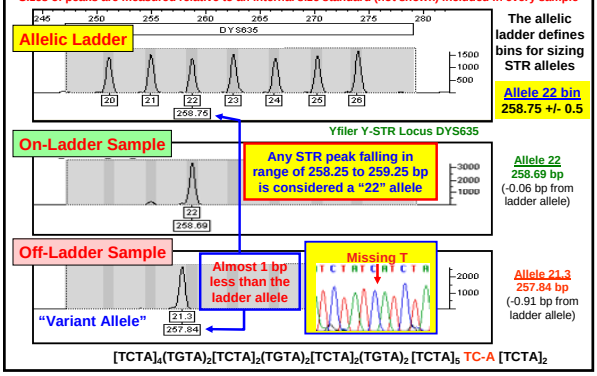
Capillary Electropherogram

Information is tied together with multiplex PCR and data analysis



A Locus-Specific Allelic Ladder Composed of Common Alleles is Used to Calibrate Size Measurements

Sizes of peaks are measured relative to an internal size standard (not shown) included in every sample



$[TCTA]_4(TGTA)_2[TCTA]_2(TGTA)_2[TCTA]_2(TGTA)_2\,[TCTA]_5$ TC-A $[TCTA]_2$

STR Data is Tabulated by Genotype Calls for Each Locus

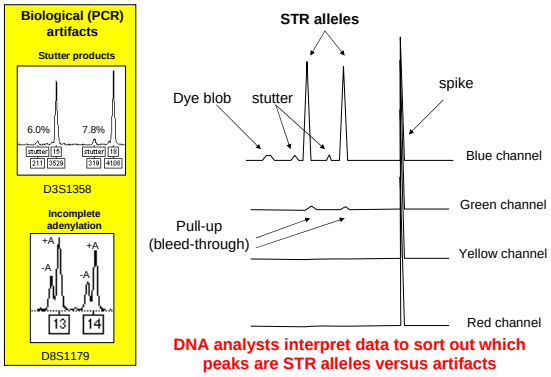
	AMEL	CSF1PO	FGA	TH01	TPOX	VWA	D3S1358	D5S818
Ind(1)	X,Y	11,12	19,21	6,7	8,8	15,18	14,18	10,13

The number of repeats observed for each locus is tabulated

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

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

Deciphering Artifacts from the True Alleles



DNA analysts interpret data to sort out which peaks are STR alleles versus artifacts

Butler, J.M. (2005) *Forensic DNA Typing, 2nd Edition*, Figure 15.4, © Elsevier Science/Academic Press

Mixtures: Issues and Challenges

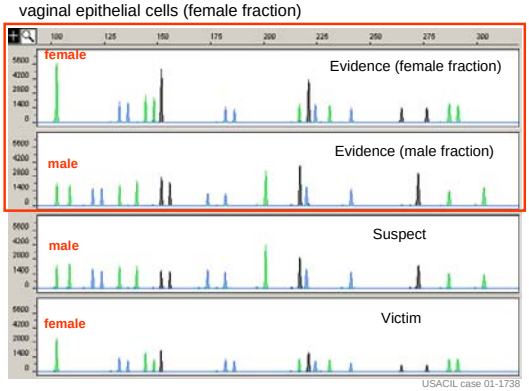
From J.M. Butler (2005) *Forensic DNA Typing, 2nd Edition*, p. 154

- 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…

Differential extraction used to separate sperm (male fraction) from vaginal epithelial cells (female fraction)

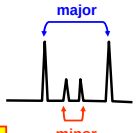


USACIL case 01-1738

The four samples typically associated with a forensic DNA case…

Principles of Mixture Interpretation

Most mixtures encountered in casework are 2-component mixtures arising from a combination of victim and perpetrator DNA profiles

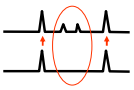


Torres *et al.* (2003) *Forensic Sci. Int.* 134:180-186 examined 1,547 cases from 1997-2000 containing **2,424 typed samples** of which **163 (6.7%) contained a mixed profile** with **only 8 (0.3%) coming from more than two contributors**

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

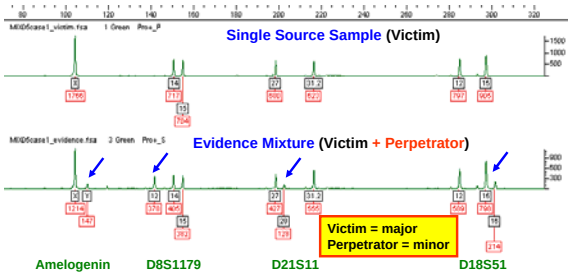
Some mixture interpretation strategies involve using victim (or other reference) alleles to help isolate obligate alleles coming from the unknown portion of the mixture



http://www.cstl.nist.gov/biotech/strbase/interlab/MIX05.htm

MIX05 Case #1; Profiler Plus green loci

Example Mixture Data (MIX05 Study-Profiler Plus)



Amelogenin	D8S1179	D21S11	D18S51
Obligate Alleles (not present in the victim reference)			
Y	12	28	16
True "Perpetrator" Profile			
X,Y	12,12	28,31.2	15,16

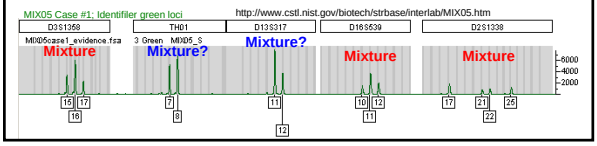
Mixtures: Issues and Challenges

- 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

From J.M. Butler (2005) *Forensic DNA Typing, 2nd Edition*, p. 155

- 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

- Determination of alleles present in the evidence and **deconvolution of mixture components** where possible
  - Many times through comparison to victim and suspect profiles
- **Providing some kind of statistical answer** regarding the weight of the evidence
  - There are multiple approaches and philosophies

Statistical Approaches with Mixtures

See Ladd *et al.* (2001) *Croat Med J.* 42:244-246

- **Inferring Genotypes of Contributors** - Separate major and minor components into individual profiles and compute the random match probability estimate as if a component was from a single source
- **Calculation of Exclusion Probabilities** - CPE/CPI (RMNE) – The probability that a random person (unrelated individual) would be excluded as a contributor to the observed DNA mixture
- **Calculation of Likelihood Ratio Estimates** – Comparing the probability of observing the mixture data under two (or more) alternative hypotheses; in its simplest form LR = 1/RMP

DNA Advisory Board (DAB) Recommendations on Statistics

February 23, 2000
Forensic Sci. Comm. 2(3); available on-line at
http://www.fbi.gov/hq/lab/fsc/backissu/july2000/dnastat.htm

"The DAB finds either one or both PE or LR calculations acceptable and strongly recommends that one or both calculations be carried out whenever feasible and a mixture is indicated"

- Probability of exclusion (PE)
  - Devlin, B. (1992) Forensic inference from genetic markers. *Statistical Methods in Medical Research*, 2, 241–262.
- Likelihood ratios (LR)
  - Evett, I. W. and Weir, B. S. (1998) *Interpreting DNA Evidence*. Sinauer, Sunderland, Massachusetts.

ISFG Recommendations on Mixture Interpretation

July 13, 2006 issue of *Forensic Science International*

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


ELSEVIER

Forensic Science International 160 (2006) 90–101

www.elsevier.com/locate/forsciint

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

P. Gill [a,*], C.H. Brenner [b], J.S. Buckleton [c], A. Carracedo [d], M. Krawczak [e], W.R. Mayr [f], N. Morling [g], M. Prinz [h], P.M. Schneider [i], B.S. Weir [j]

[a] Forensic Science Service, Trident Court, 2960 Solihull Parkway, Birmingham, UK
[b] Forensic Science Group, School of Public Health, University of California, Berkeley, CA 510-330-1911, USA

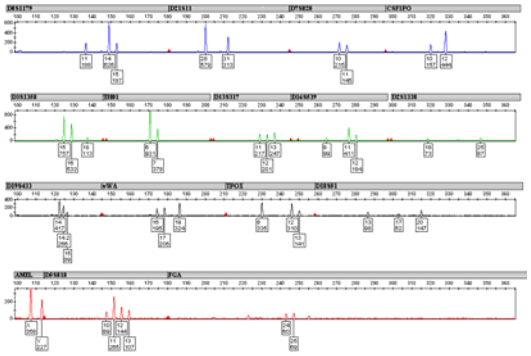
Abstract

The DNA commission of the International Society of Forensic Genetics (ISFG) was convened at the 21st congress of the International Society for Forensic Genetics held between 13 and 17 September in the Azores, Portugal. The purpose of the group was to agree on guidelines to encourage best practice that can be universally applied to assist with mixture interpretation. In addition the commission was tasked to provide guidance on low copy number (LCN) reporting. Our discussions have highlighted a significant need for continuing education and research into this area. We have attempted to present a consensus from experts but to be practical we do not claim to have conveyed a clear vision in every respect in this difficult subject. For this reason, we propose to allow a period of time for feedback and reflection by the scientific community. Then the DNA commission will meet again to consider further recommendations.

Keywords: STR typing; Biostatistical analysis; Likelihood ratio; Probability of exclusion; Mixtures; ISFG DNA commission

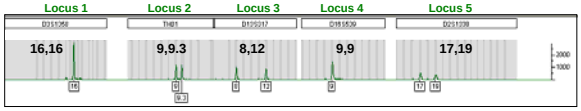
Mixture Interpretation – A Major Challenge…

Not a clear cut answer because DNA result is from multiple contributors



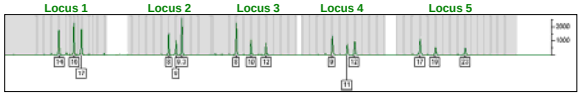
Single Source vs. Mixture Samples

Single Source Sample



One or two peaks observed at each locus (tested DNA region)

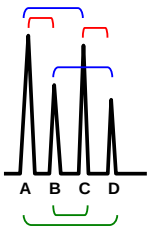
Mixture Sample



More than two peaks observed at more than two loci (tested DNA regions)

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

With Some Mixtures, Multiple Genotype Combinations Are Possible

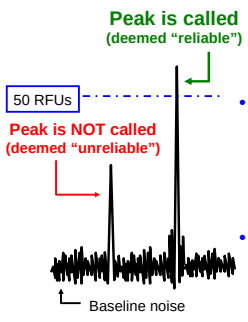


AC
BD
AB
CD
BC
AD

Depends on PHR and proportion of mixture components from the various contributors

Peak Height Ratios (PHR)
Minimum Peak Height (mPH)
Proportion (p) or mixture proportion (M_x)

Thresholds for Measuring DNA Data



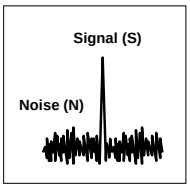
These thresholds for reliable data are determined through validation studies

- **Detection (analytical) threshold**
  - Dependent on instrument sensitivity ~50 RFU (relative fluorescence units)
  - Impacted by instrument baseline noise
- **Dropout (stochastic) threshold**
  - Dependent on biological sensitivity ~150-200 RFU
  - **Important in mixture interpretation**

STR Data Interpretation Involves Determining What is a True Allele (Peak)

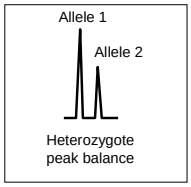
All of these issues impact mixture interpretation

Peak detection threshold



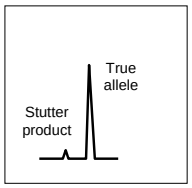
Signal >3x sd of noise (or S/N >3)

Peak height ratio (PHR)



PHRs consistent with single source are typically above 60%

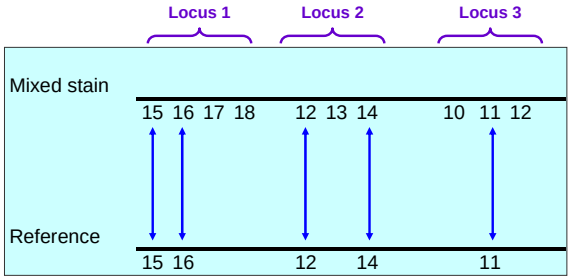
Stutter percentage



Stutter is usually one repeat position less and <15% than true allele

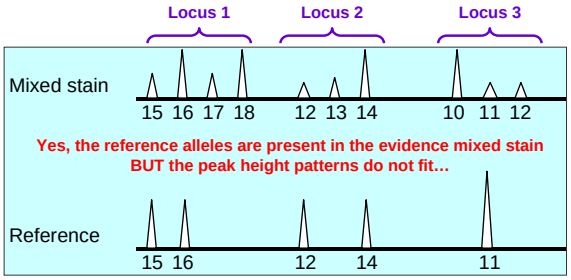
Mixture Example

Comparing Alleles Only



Mixture Example

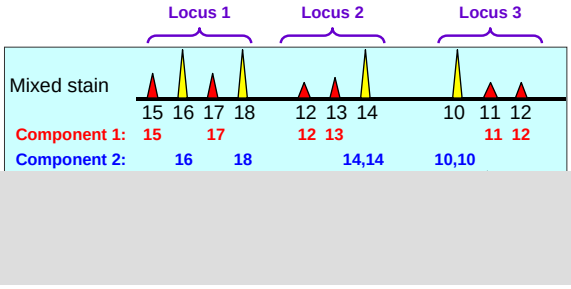
Showing Importance of Using Peak Height Information



Yes, the reference alleles are present in the evidence mixed stain BUT the peak height patterns do not fit…

Mixture Example

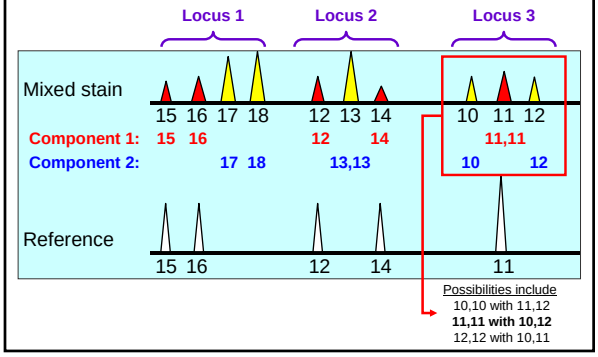
Solving Components Prior to Comparison to Suspect Reference



Reference (suspect) does not match either component of the mixed stain and therefore could not have contributed to the evidence sample

Mixture Example

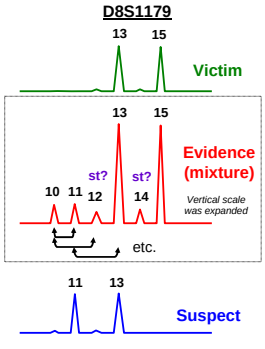
Different Evidence Sample…



Possibilities include
10,10 with 11,12
11,11 with 10,12
12,12 with 10,11

A Mixture Example

That Got the Duke Rape Case Prosecutor in Trouble



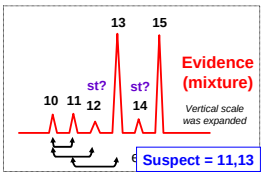
Conclusions from the evidence:

1. **Major contributor = 13,15 (victim) –** *to be expected with an intimate sample like a fingernail or vaginal swab*
2. Alleles 12 and 14 are **likely stutter products** of the major contributor's 13 and 15 alleles but could also be masking minor contributor alleles
3. A number of minor contributor combinations are possible (e.g., 10,11 or 10,12 or 10,13 or 11,13, etc.)
4. Could have more than two contributors present in this mixture

"Suspect cannot be excluded" BUT statistics needs to be qualified by statistics because a large percentage of the population might also not be able to be excluded…

Probability of Exclusion Calculation
for a Single STR Locus

The case may grow stronger against a suspect with information from additional STR loci…



From VA DFS STR Allele Frequencies
http://www.dfs.virginia.gov/manuals/manuals.cfm?id=5

D8S1179 alleles	AA (n=384)	C (n=346)	H (n=366)
10	0.0287	0.1069	0.0820
11	0.0495	0.0925	0.0465
12	0.1094	0.1416	0.1093
13	0.2422	0.3093	0.3224
14	0.2969	0.1965	0.2623
15	0.1849	0.0896	0.1202
SUM	0.9115	0.9364	0.9426
Sq SUM = PI	0.8308	0.8769	0.8886
PE = 1-PI	0.1692	0.1231	0.1114
PE (%)	**16.9%**	**12.3%**	**11.1%**
	African Am.	Caucasians	Hispanics

"Suspect cannot be excluded" BUT we would expect to see, for example, only 11.1% of Hispanics excluded (or 88.9% cannot be excluded) based on results at this one locus

The fact that in this case a suspect is included is not very informative because ~9 out of 10 people examined from any population could potentially be included in the evidence mixture…

Fox News Atlanta Story
http://tinyurl.com/MixedSampleDNA

- **I-Team: The GBI and DNA**
- **Monday, 27 Apr 2009, 7:08 PM EDT**
- **http://www.myfoxatlanta.com/dpp/news/ITeam_The_GBI_and_DNA_042709**

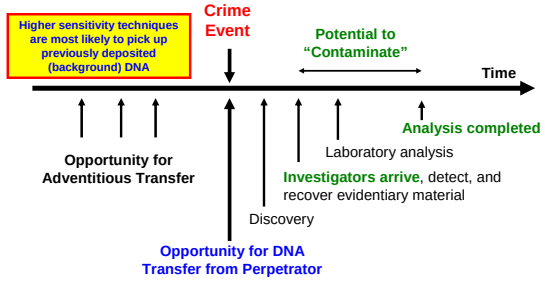
- **I-Team: The GBI and DNA Part 2**
- Tuesday, 28 Apr 2009, 10:50 PM EDT
- http://www.myfoxatlanta.com/dpp/news/ITeam_The_GBI_and_DNA_Pt_2_042809

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

From J.M. Butler (2005) *Forensic DNA Typing, 2nd Edition*, p. 154

…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.*

Time Line Showing the Potential for DNA Deposition/Transfer



Adapted from Gill, P. (2002) *BioTechniques* 32(2): 366-385, Figure 5

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

Conclusions

- This is an exciting time to be involved in forensic DNA testing
- However, it is a little scary because technology is advancing so rapidly on some fronts
- Thus, training for both the scientific and legal communities is vital to make the most effective use of the wonderful power of DNA technology

Data is Tabulated

	AMEL	CSF1PO	FGA	TH01	TPOX	VWA	D3S1358	D5S818
Ind(1)	X,Y	11,12	19,21	6,7	8,8	15,18	14,18	10,13

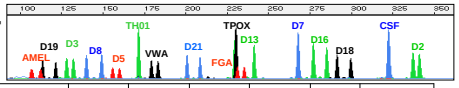
The number of repeats observed for each locus is tabulated

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

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

DNA Profile Frequency with all 13 CODIS STR loci

AmpFISTR® Identifiler™ (Applied Biosystems)



What would be entered into a DNA database for searching: 16,17-17,18-21,22-12,14-28,30-14,16-12,13-11,14-9,9-9,11-6,6-8,8-10,10

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

PRODUCT RULE

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 2 pq)

NIST study: Butler, J.M., *et al*. (2003) Allele frequencies for 15 autosomal STR loci on U.S. Caucasian, African American, and Hispanic populations. *J. Forensic Sci.* 48(4):908-911. (http://www.cstl.nist.gov/biotech/strbase/NISTpop.htm)

*http://www.csfs.ca/pplus/profiler.htm

DNA Degradation

- What causes DNA degradation?
  - Heat, humidity, long term exposure to the elements
  - DNA breaks down into small fragments; smaller than the targeted PCR product size
- Mass disasters (aviation, WTC)
- Aged samples (missing persons, remains of soldiers, ancient DNA)

DNA Degradation

Intact sample

Target region for PCR

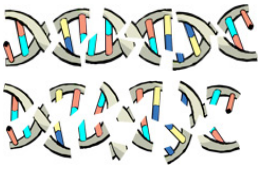
300 base pair PCR product can be produced

Degraded sample

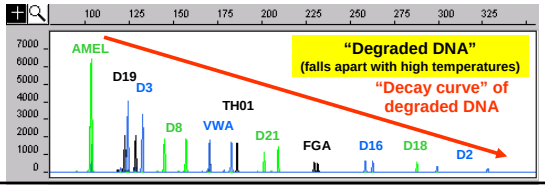
Target region for PCR is fragmented

300 base pair PCR product can not be produced or only in limited quantities

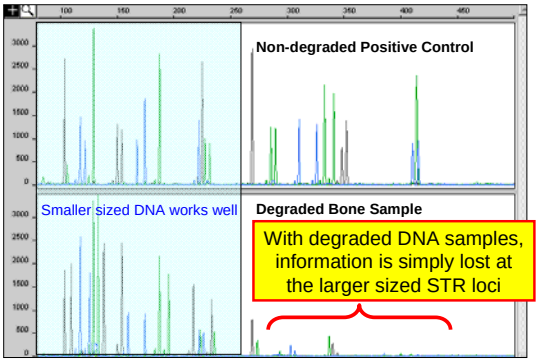
Degraded DNA



Larger segments of DNA cannot be recovered when DNA molecules have fragmented into small pieces (caused by heat, water, or bacteria)



DNA Degradation Means Less Loci Work



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…**

  ----------4049 or ----301-9-------