Topics and Techniques for Forensic DNA Analysis

Mixture Interpretation

Houston DNA
Training Workshop

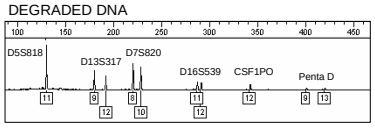
Houston, TX
April 3-4, 2007

Dr. John M. Butler
National Institute of Standards and Technology

john.butler@nist.gov

Common Casework Challenges

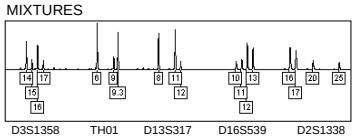
DEGRADED DNA



Loss of signal at larger size loci

miniSTRs can help

MIXTURES

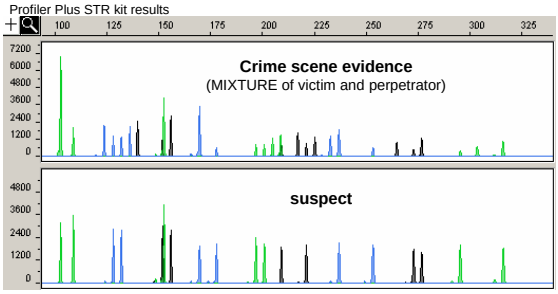


More than two alleles at multiple loci

From Butler, J.M. (2004) Short tandem repeat analysis for human identity testing. *Current Protocols in Human Genetics*, John Wiley & Sons, Hoboken, NJ, Unit 14.8, *(Supplement 41)*, pp. 14.8.1-14.8.22

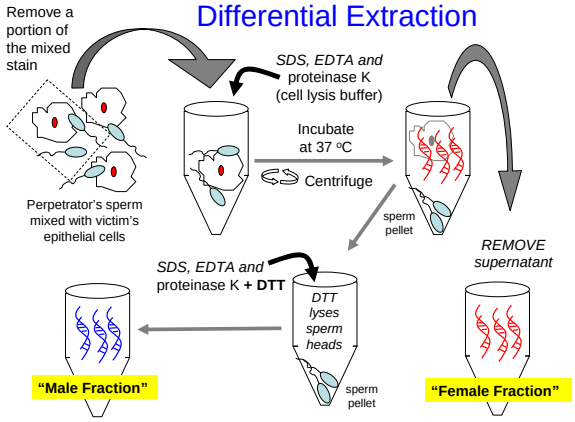
Forensic Casework DNA Profiles

Profiler Plus STR kit results

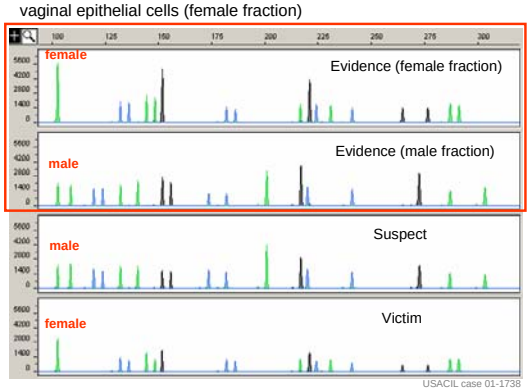


All of the suspect's alleles are included in the evidentiary DNA profile

Differential Extraction



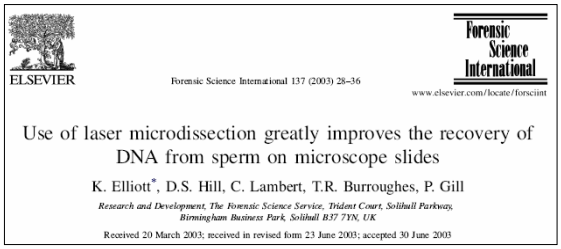
Differential extraction used to separate sperm (male fraction) from vaginal epithelial cells (female fraction)



USACIL case 01-1738

The four samples typically associated with a forensic DNA case…

Evidence Enrichment
Through Physical Capture of Sperm Cells



ELSEVIER

Forensic Science International 137 (2003) 28–36

Forensic Science International

www.elsevier.com/locate/forsciint

Use of laser microdissection greatly improves the recovery of DNA from sperm on microscope slides

K. Elliott*, D.S. Hill, C. Lambert, T.R. Burroughes, P. Gill

Research and Development, The Forensic Science Service, Trident Court, Solihull Parkway, Birmingham Business Park, Solihull B37 7YN, UK

Received 20 March 2003; received in revised form 23 June 2003; accepted 30 June 2003

Laser capture microdissection permits collection of individual sperm. Typically >50 sperm heads enable collection of full diploid DNA profile…

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…

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)

Victim Reference and Spouse or Boyfriend Reference

- **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

Reference elimination samples are useful in deciphering both situations

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

J Forensic Sci, May 2004, Vol. 49, No. 3
Paper ID JFS2003366
Available online at: www.astm.org

Peter Gill,[1] Ph.D. and Amanda Kirkham,[1] B.Sc.

Development of a Simulation Model to Assess the Impact of Contamination in Casework Using STRs

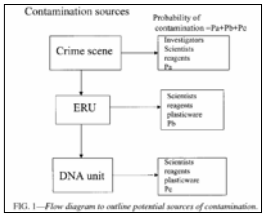


FIG. 1—*Flow diagram to outline potential sources of contamination.*

- **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

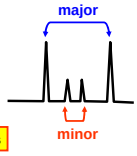
Potential Impact of Contamination on Cold Cases or Post-Conviction Testing

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

- 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.*

Principles of Mixture Interpretation

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

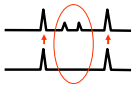


Torres *et al.* (2003) *Forensic Sci. Int.* 134:180-186 examined 1,547 cases from 1997-2000 containing **2,424 typed samples** of which **163 (6.7%) contained a mixed profile** with **only 8 (0.3%) coming from more than two contributors**

95.1% (155/163) were 2-component mixtures

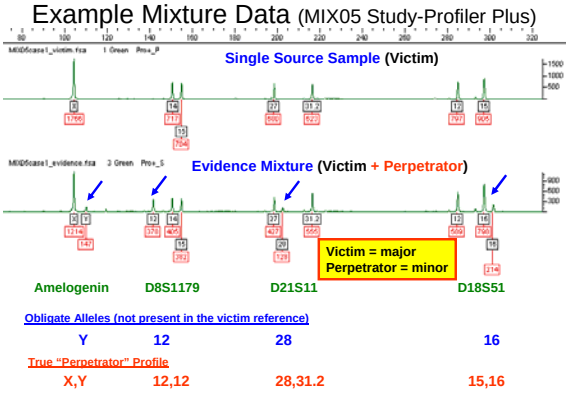
Ratios of the various mixture components stay fairly constant between multiple loci enabling deduction of the profiles for the major and minor components

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)



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

Setting thresholds for the ABI 310/3100

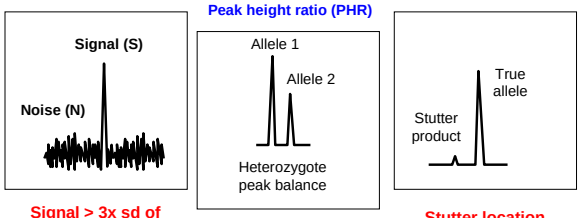
- Where do current ideas on instrument thresholds for the ABI 310/3100 come from?
- How do I set these values in my laboratory?
- Why might they vary from one instrument to the next?
- How do these thresholds affect data interpretation?

Future defense attacks will likely focus on detection thresholds – can you defend your current threshold (e.g., 50 RFU or 150 RFU)?

What is a true peak (allele)?

GeneScan function *Genotyper function*

Peak detection threshold **Peak height ratio (PHR)** **Stutter percentage**



Signal > 3x sd of noise

PHR consistent with single source Typically above 60%

Stutter location above 15%

Threshold Settings for the ABI 310/3100

Detection Limit: 3x the standard deviation of the noise.
Estimated using **2x peak to peak noise**. (approximately 35 - 50 RFUs)

Limit of Quantitation: 10x the standard deviation of the noise
Estimated using **7x peak to peak noise** (150-200 RFUs)
Below this point estimates of peak area or height are unreliable.

Dynamic Range: The range of sample quantities that can be analyzed from the lowest to the highest **(linear range is also important)**

Stochastic Threshold: Level of quantifiable DNA below which peaks can show severe imbalance (peak height ratios below 60%) Approximately 150 -200 RFUs. Enhanced stutter also occurs at these signal levels.

Will be covered more in the low copy number section of this workshop…

The Scientific Reasoning behind the Concept of an Analytical Threshold (limit of detection)

- This is fundamentally an issue of reliability
- **For a peak intensity three times the standard deviation of the noise there is a limited chance that such a signal is the result of a random fluctuation**
- This is because 99.7 percent of all noise signals fall below this value (from the definition of a Gaussian curve)
- **Below this point the very real possibility exists that what you think is a peak is simply a statistical fluctuation in the baseline noise**.

How Does Your Laboratory Derive Its Interpretation Rules?

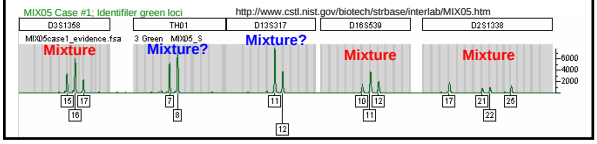
From your Validation Studies or Others?

- **Peak detection threshold** – set to 50 RFU or 150 RFU based on your lab data or what FBI or manufacturer has done? Do you use S/N >3 for determining if something is a true peak?
- **Peak height ratio threshold** – Set at 70% due to suggestion by manufacturer? Or 50-70% based on other data?
- **Stutter product threshold** – are Genotyper macros set to 15%, manufacturer values, or adjusted based on your validation? Does it matter? How do these values play into your mixture interpretation guidelines?
- **Sample Cleanup** - Post PCR concentration a sample may also remove salts artificially enhancing injection. Will this move results into stochastic range?

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.



Is a DNA Profile Consistent with Being a Mixture?

From J.M. Butler (2005) *Forensic DNA Typing, 2nd Edition*, pp. 156-157

If the answer to any one of the following three questions is yes, then the DNA profile may very well have resulted from a mixed sample:

- Do any of the loci show more than two peaks in the expected allele size range?
- Is there a severe peak height imbalance between heterozygous alleles at a locus?
- Does the stutter product appear abnormally high (e.g., >15-20%)?

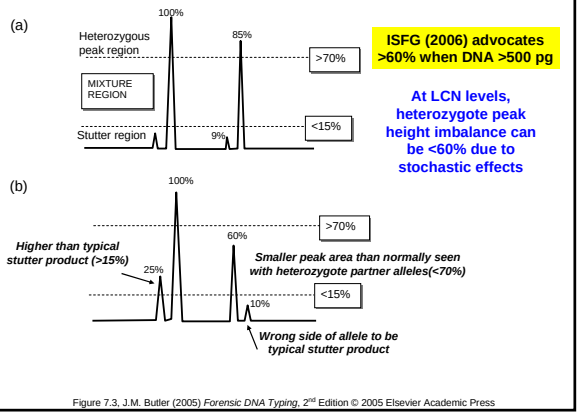


Figure 7.3, J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press

Steps in the Interpretation of Mixtures

(Clayton *et al.* 1998)

Step #1 Identify the Presence of a Mixture

Step #2 Designate Allele Peaks

Step #3 Identify the Number of Potential Contributors

Step #4 Estimate the Relative Ratio of the Individuals Contributing to the Mixture

Step #5 Consider All Possible Genotype Combinations

Step #6 Compare Reference Samples


Analysis and interpretation of mixed forensic stains using DNA STR profiling

T.M. Clayton, J.P. Whitaker, R. Sparkes, P. Gill

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

Gill *et al.* (2006) DNA Commission of the International Society of Forensic Genetics: Recommendations on the interpretation of mixtures. *Forensic Sci. Int.* 160: 90-101

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

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 510-339-1911, USA

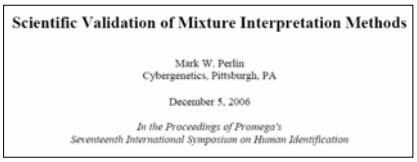
Abstract

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

Some of Mark Perlin's Recent Statements

http://www.promega.com/geneticidproc/ussymp17proc/oralpresentations/Perlin.pdf



Scientific Validation of Mixture Interpretation Methods

Mark W. Perlin
Cybergenetics, Pittsburgh, PA

December 5, 2006

In the Proceedings of Promega's Seventeenth International Symposium on Human Identification

- **Different laboratories follow different mixture interpretation guidelines.** Moreover, different examiners within the same laboratory who are following the same guidelines often infer different STR profiles.
- **Therefore, there is no concordance in current forensic practice on what constitutes a "correct" mixture solution.** Thus, it is not possible to conduct a mixture interpretation concordance study in order to validate a mixture interpretation method.
- **DNA mixture evidence currently fails the general acceptance test of both Frye and Daubert, since there are no generally accepted methods for interpreting mixed stains.**

A High Degree of Variability Currently Exists with Mixture Interpretation

- **"If you show 10 colleagues a mixture, you will probably end up with 10 different answers"**
  - Peter Gill, Human Identification E-Symposium, April 14, 2005
- **Interlaboratory studies help to better understand why variability may exist between laboratories**
- Most analysts are only concerned about their own lab protocols and do not get an opportunity to see the big picture from the entire community that can be provided by a well-run interlaboratory study

NIST Initiated Interlaboratory Studies

Studies involving STRs	# Labs	Publications
Evaluation of CSF1PO, TPOX, and TH01	34	Kline MC, Duewer DL, Newall P, Redman JW, Reeder DJ, Richard M. (1997) Interlaboratory evaluation of STR triplex CTT. *J. Forensic Sci.* 42: 897-906
Mixed Stain Studies #1 and #2 (Apr–Nov 1997 and Jan–May 1999)	45	Duewer DL, Kline MC, Redman JW, Newall PJ, Reeder DJ. (2001) NIST Mixed Stain Studies #1 and #2: interlaboratory comparison of DNA quantification practice and short tandem repeat multiplex performance with multiple-source samples. *J. Forensic Sci.* 46: 1199-1210
MSS3 **Mixed Stain Study** #3 (Oct 2000-May 2001)	74	Kline, M.C., Duewer, D.L., Redman, J.W., Butler, J.M. (2003) NIST mixed stain study 3: DNA quantitation accuracy and its influence on short tandem repeat multiplex signal intensity. *Anal. Chem.* 75: 2463-2469. Duewer, D.L., Kline, M.C., Redman, J.W., Butler, J.M. (2004) NIST Mixed Stain Study #3: signal intensity balance in commercial short tandem repeat multiplexes, *Anal. Chem.* 76: 6928-6934.
DNA Quantitation Study (Jan-Mar 2004) QS04	80	Kline, M.C., Duewer, D.L., Redman, J.W., Butler, J.M. (2005) Results from the NIST 2004 DNA Quantitation Study, *J. Forensic Sci.* 50(3):571-578
Mixture Interpretation Study (Jan - Aug 2005) MIX05	**69**	***Data analysis currently on-going ...*** **Poster at 2005 Promega meeting (Sept 2005); available on STRBase**

Overall Lessons Learned from NIST MSS 1,2,&3

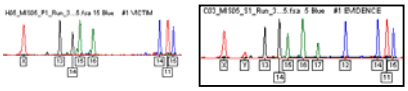
- Laboratories have instruments with different sensitivities
- **Different levels of experience and training plays a part in effective mixture interpretation**
- Amount of input DNA makes a difference in the ability to detect the minor component (labs that put in "too much" DNA actually detected minor components more frequently)

Purpose of MIX05 Study

- **Goal is to understand the "lay of the land" regarding mixture analysis across the DNA typing community**
- One of the primary benefits we hope to gain from this study is **recommendations for a more uniform approach to mixture interpretation** and training tools to help educate the community

Mixture Interpretation Interlab Study (MIX05)

- **Only involves interpretation of data – to remove instrument detection variability and quantitation accuracy issues**
- **94 labs enrolled** for participation
- **69 labs have returned results** (17 from outside U.S.)
- Four mock cases supplied with "victim" and "evidence" electropherograms (GeneScan .fsa files – that can be converted for Mac or GeneMapper; gel files made available to FMBIO labs)
- Data available with Profiler Plus, COfiler, SGM Plus, PowerPlex 16, Identifiler, PowerPlex 16 BIO (FMBIO) kits
- Summary of results will involve training materials to illustrate various approaches to solving mixtures



Perpetrator Profile(s) ??

Along with reasons for making calls and any stats that would be reported

MIX05 Study Design and Purpose

Interlab studies provide a "big picture" view of the community

- Permit a large number of forensic practioners to **evaluate the same mixture data**
- Provide multiple cases representing a range of mixture scenarios
- Generate data from multiple STR kits on the same mixture samples to compare performance for detecting minor components
- The primary variable should be the laboratory's interpretation guidelines rather than the DNA extraction, PCR amplification, and STR typing instrument sensitivity
- **Are there best practices in the field that can be advocated to others?**

Requests for Participants in MIX05

Mixtures representing four different case scenarios have been generated at NIST with multiple STR kits and provided to laboratories as electropherograms.

We would like to receive the following information:

1) **Report the results as though they were from a real case** including whether a statistical value would be attached to the results. *Please summarize the perpetrator(s) alleles in each "case" as they might be presented in court—along with an appropriate statistic (if warranted by your laboratory standard operating procedure) and the source of the allele frequencies used to make the calculation.* Please indicate which kit(s) were used to solve each case.

2) **Estimate the ratio for samples present in the evidence mixture** and how this estimate was determined.

3) **Provide a copy of your laboratory mixture interpretation guidelines** and a brief explanation as to why conclusions were reached in each scenario

A MIX05 Participant Noted…

"Things we do not do:

- **Calculate mixture ratios for casework**
  - **Calculation used for this study:** Find loci with 4 alleles (2 sets of sister alleles). Make sure sister alleles fall within 70%, then take the ratio of one allele from one sister set to one allele of the second sister set, figure ratios for all combinations and average. Use peak heights to calculate ratios.
- **Provide allele calls in reports**
- **Provide perpetrator(s) alleles or statistics in court without a reference sample to compare to the DNA profile obtained from the evidence. We will try to determine the perpetrator(s) profile for entry into CODIS."**

We recognize that some of the information requested in this interlab study may not be part of a lab's standard operating procedure

MIX05 Case Scenarios

Genomic DNA samples with specific allele combinations ("evidence") were mixed in the following ratios:

Based on Identifiler 15 STR loci

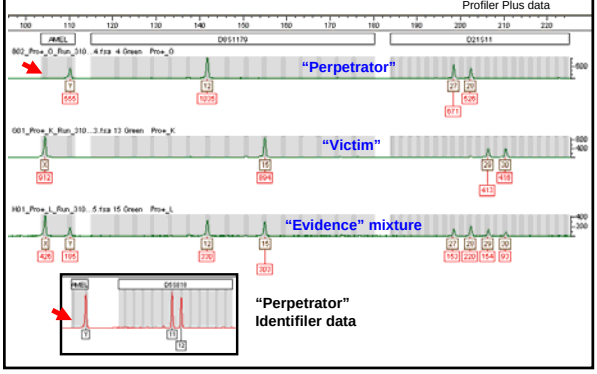
	#alleles		#loci with #alleles				
	N all	N unq	N 1	N 2	N 3	N 4	N 5
Case #1 – victim is major contributor (3F:1M)	39	26	2	6	5	2	0
Case #2 – perpetrator is major contributor (1F:3M)	55	52	0	1	4	10	0
Case #3 – balanced mixture (1F:1M) • Male lacked amelogenin X	48	37	0	3	8	4	0
Case #4 – more extreme mixture (7F:1M) • Male contained tri-allelic pattern at TPOX	50	42	0	3	7	4	1

Female victim DNA profile was supplied for each case

Labs asked to deduce the perpetrator DNA profile – suspect(s) not provided

Amelogenin X allele is missing in male perpetrator DNA sample for MIX05 Case #3

Profiler Plus data

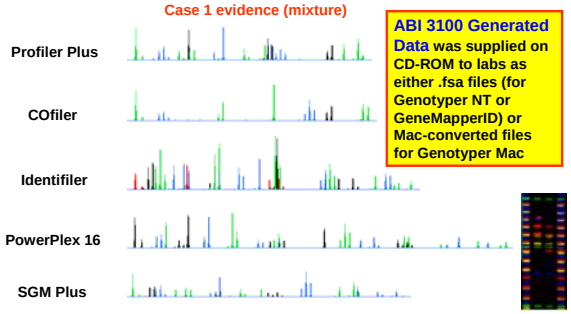


"Perpetrator" Identifiler data

MIX05 Results on Multiple Kits

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

Case 1 evidence (mixture)



ABI 3100 Generated Data was supplied on CD-ROM to labs as either .fsa files (for Genotyper NT or GeneMapperID) or Mac-converted files for Genotyper Mac

FMBIO data was also made available upon request

Summary of MIX05 Responses

94 labs enrolled for participation

69 labs returned results (17 from outside U.S.)

50 labs made allele calls
39 labs estimated ratios
29 labs provided stats

All participants were supplied with all data and could choose what kits to examine based on their experience and lab protocols

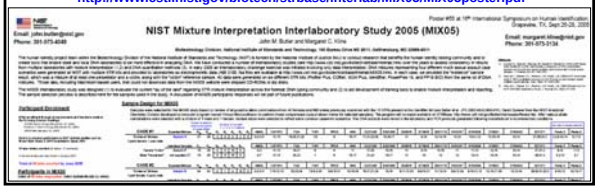
STR kit results used
- 34 ProfilerPlus/COfiler
- 10 PowerPlex 16
- 7 PP16 BIO
- 5 Identifiler
- 2 SGM Plus
- 1 All ABI kit data
- 9 Various combinations

Generally Identifiler data was of poorer quality in the electropherograms we provided...which caused some labs to not return results (they indicated a desire for higher quality data through sample re-injection to reduce pull-up prior to data interpretation)

What MIX05 Participants Have Received Back from NIST…

- Certificate of participation in the interlab study
- Copy of the poster presented at the Promega Sept 2005 meeting displaying "correct" results for the perpetrator in each case scenario as well as an explanation of study design and preliminary results

http://www.cstl.nist.gov/biotech/strbase/interlab/MIX05/MIX05poster.pdf


NIST Mixture Interpretation Interlaboratory Study 2005 (MIX05)

When is a Sample a Potential Mixture?

According to several MIX05 participant interpretation guidelines

- Number of Observed Peaks
  - Greater than two peaks at a locus
  - More than two alleles are present at two or more loci, although three banded patterns can occur
  - Presence of 3 alleles at a single locus within a profile
  - 4 peaked patterns (if observed at any locus), 3 peaked patterns (if observed at two or more loci), significant imbalances (peak height ratios <60%) of alleles for a heterozygous genotype at two or more loci with the exception of low template amplifications, which should be interpreted with caution
- Imbalance of heterozygote alleles
  - thresholds range from 50-70%
- Stutter above expected levels
  - generally 15-20%

Detection thresholds also varied in the range of 50-200 RFUs

These protocol differences can lead to variation in reported alleles and therefore the deduced profile and resulting statistics

Summary of Some MIX05 Reported Results

Case #2 has perpetrator as major component and thus is the easiest to solve...

CASE #2		D3S1358	VWA	FGA	AMEL	D8S1179	D21S11	D18S51	D5S818	D13S317	D7S820	D16S539	TH01	TPOX	CSF1PO
True Perp	*Z779619*	15,15	15,15	20,24	X,Y	11,13	28,32.2	17,18	8,13	12,14	8,10	10,11	7,9.3	9,10	7,10
LabID	Kit Used														
16	ProPlus/Cofiler	--	--	--	--	--	--	--	--	--	--	--	--	--	--
6	ProPlus/Cofiler	15	15	20,24	X,Y	11,13	28,32.2	17,18	8,13	12,14	8,10	10,11	7,9.3	9,10	7,10
91	SGM Plus	15	15	20,24	X,Y	11,13	28,32.2	17,18				10,11	7,9.3		
46	PP16	--	--	--	--	--	--	--	--	--	--	--	--	--	--
37	ProPlus/Cofiler	--	15	20	X,Y	13	28,32.2	17,18	8,13	12,14	8,10	10,11	7,9.3	9,10	7,10
2	PP16	15	15,15	20,24	X,Y	11,13	28,32.2	17,18	8,13	INC	8,10	10,11	7,9.3	9,10	7,10
13	PP16 & Identifiler	15	15	20,24	--	11,13	28,32.2	17,18	8,13	12,14	8,10	10,11	7,9.3	9,10	7,10
34	ProPlus/Cofiler	15	15	20,24	--	11,13	28,32.2	17,18	8,13	12,14	8,10	10,11	7,9.3	9,10	7,10
70	Identifiler	15	15	20,24	X,Y	11,13	28,32.2	17,18	8,13	12,14	8,10	10,11	7,9.3	9,10	7,10
55	ProPlus/Cofiler	15	15	20,24	--	11,13	28,32.2	17,18	8,13	12,14	8,10	10,11	7,9.3	9,10	7,10
21	ProPlus/Cofiler	15,15	15,15	20,24	X,Y	11,13	28,32.2	17,18	8,13	12,14	8,10	10,11	7,9.3	9,10	7,10
73	ProPlus/Cofiler	15,15	15,15	20,24	X,Y	11,13	28,32.2	17,18	8,13	12,14	8,10	10,11	7,9.3	9,10	7,10
29	Identifiler	15	15	20,24	X,Y	11,13	28,32.2	17,18	8,13	12,14	8,10	10,11	7,9.3	9,10	7,10
54	All Kits	15,15	15,15	20,24	X,Y	11,13	28,32.2	17,18	8,13	12,14	8,10	10,11	7,9.3	9,10	7,10
90	ProPlus/Cofiler	15	15	20,24	X,Y	11,13	28,32.2	17,18	8,13	12,14	8,10	10,11	7,9.3	9,10	7,10
9	ProPlus/Cofiler	15	15	20,24	X,Y	11,13	28,32.2	17,18	8,13	12,14	8,10	10,11	7,9.3	9,10	7,10
4	ProPlus/Cofiler	15	15	20,24	X,Y	11,13	28,32.2	17,18	8,13	12,14	8,10	10,11	7,9.3	9,10	7,10
33	ProPlus/Cofiler	--	--	--	--	--	--	--	--	--	--	--	--	--	--
12	ProPlus/Cofiler	15	15	20,24	X,Y	11,13	28,32.2	17,18	8,13	12,14	8,10	10,11	7,9.3	9,10	7,10
67	PP16	15	15,16	20,24	X,Y	11,13	28,32.2	17,18	8,13	12,14	8,10	10,11	7,9.3	9,10	7,10
86	ProPlus/Cofiler	15,15	15,15	20,24	--	11,13	28,32.2	17,18	8,13	12,14	8,10	10,11	7,9.3	9,10	7,10
79	ProPlus/Cofiler	15,15	15,15	20,24	--	11,13	28,32.2	17,18	8,13	12,14	8,10	10,11	7,9.3	9,10	7,10
77	Identifiler	--	--	--	--	--	--	--	--	--	--	--	--	--	--
60	PP16	15	15	20,24	X,Y	11,13	28,32.2	17,18	8,13	12,14	8,10	10,11	7,9.3	9,10	7,10
61	Identifiler	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Most calls were correct (when they were made)

Some Mixture Ratios Reported in MIX05

Many labs do not routinely report the estimated ratio of mixture components

LabID	Case1 (F:M)	Case2 (M:F)	Case3 (M:F)	Case4 (F:M)
13	2	5	<2	10
34	1.8--3.6	3.9--6.7	1.6--1.8	6.2--7.6
70				
55	68%:32%	85%:15%	64%:36%	
21				
73	2:1	6:1	2:1	not determined
29				
54	2:1	6:1	2:1	6:1
90	male23-39%	not determined	male64-71%	
9	3 or 4:1	4 or 5:1	1.4:1	~10:1
4	10:1	6:1	1:1	not determined
33	male60-78%	male80-90%	male58-71%	victim86%
12	male25%	male85%	male40-45%	unknown10%
67	1:2.3	6.4:1	2:1	1:6.8
86	2:1	6-6.5:1	1.6-2:1	4-4.5:1
79	~3:1 to ~2:1	~6:1 to ~4:1	~2:1*	a lot of victim
77				
60	2:1	5:1	2:1	10:1
61				

Some Reported Stats for MIX05 Case #1

Many of the 29 labs providing statistics used PopStats 5.7

		Case1		
LabID	Kits Used	Caucasians	African Americans	Hispanics
77	Identifiler	PE calculated	PE calculated	PE calculated
73	ProPlus/Cofiler	none provided	none provided	none provided
4	ProPlus/Cofiler	none provided	none provided	none provided
12	ProPlus/Cofiler	none provided	none provided	none provided
29	Identifiler	none provided	none provided	none provided
90	ProPlus/Cofiler	1.18E+15	2.13E+14	3.09E+15
34	ProPlus/Cofiler	2.40E+11	7.00E+09	9.80E+10
46	PP16	5.60E+09	3.80E+11	none provided
33	ProPlus/Cofiler	2.94E+08	1.12E+08	1.74E+09
6	ProPlus/Cofiler	40,000,000	3,500,000	280,000,000
9	ProPlus/Cofiler	1.14E+07	1.97E+07	1.54E+08
61	Identifiler	1.50E+06	260,000	2.40E+07
79	ProPlus/Cofiler	930,000	47,900	1,350,000
16	ProPlus/Cofiler	434,600	31,710	399,100

Which loci are included in each calculation?

Some Differences in Reporting Statistics

		Case1		
LabID	Kits Used	Caucasians	African Americans	Hispanics
90	ProPlus/Cofiler	1.18E+15	2.13E+14	3.09E+15
34	ProPlus/Cofiler	2.40E+11	7.00E+09	9.80E+10
33	ProPlus/Cofiler	2.94E+08	1.12E+08	1.74E+09
6	ProPlus/Cofiler	40,000,000	3,500,000	280,000,000
9	ProPlus/Cofiler	4.14E+07	1.97E+07	1.54E+08
79	ProPlus/Cofiler	930,000	47,900	1,350,000
16	ProPlus/Cofiler	434,600	31,710	399,100

~10 orders of magnitude difference (10^5 to 10^{15}) based on which alleles were deduced and reported

Remember that these labs are interpreting the same MIX05 electropherograms

Further Examination of These 7 Labs

		Case 1	ASCLD-LAB accredited?	Solved loci listed?
LabID	Kits Used	Caucasians		
90	ProPlus/Cofiler	1.18E+15	Yes	Yes
34	ProPlus/Cofiler	2.40E+11	Yes	Yes
33	ProPlus/Cofiler	2.94E+08	Yes	No
6	ProPlus/Cofiler	40,000,000	Yes	Yes
9	ProPlus/Cofiler	4.14E+07	No	No (CPE)
79	ProPlus/Cofiler	930,000	Yes	Yes
16	ProPlus/Cofiler	434,600	Yes	No

Possible Reasons for Variability in Reported Statistics:

- Different types of calculations (CPE vs RMP)
- Different loci included in calculations (due to different thresholds used)
- Different allele frequency population databases (most use PopStats)
- Use of victim (e.g., major component in Case 1) profile stats

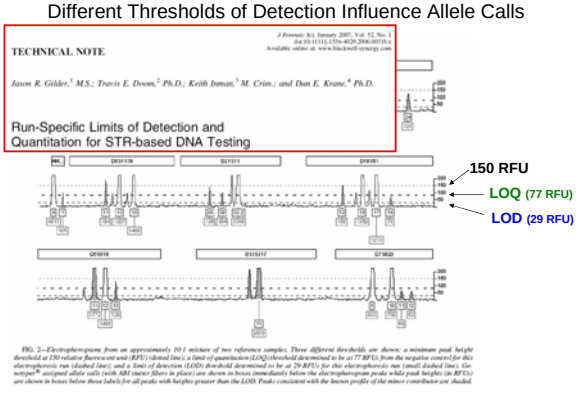
Different Stats Used

Combined Probability of Exclusion

Random Match Probability on Deduced Profiles

- **Lab 9** (4.14 x 10^7) used 1/CPI
- **Lab 6** (4.0 x 10^7) used selected loci and summed all possible genotypes for loci not completely deduced
- **Lab 90** (1.18 x 10^{15}) used theta value of 0.03 and deduced alleles at all 13 loci (***correctly deduced all perpetrator alleles***)

Different Thresholds of Detection Influence Allele Calls



Gilder, J.R., Doom, T.E., Inman, K., Krane, D.E. (2007) Run-specific limits of detection and quantitation for STR-based DNA testing. *J. Forensic Sci.* 52(1): 97-101.

Different Detection Thresholds Used

		Case 1	
LabID	Kits Used	Caucasians	
90	ProPlus/Cofiler	1.18E+15	**75 RFUs**; all 13 STRs; **all results correct**
34	ProPlus/Cofiler	2.40E+11	**Not stated**; 8 STRs, 2 partial, 3 INC
33	ProPlus/Cofiler	2.94E+08	**75 RFUs**; no deduced alleles reported
6	ProPlus/Cofiler	40,000,000	**Not provided**; 3 STRs, 6 partial, 4 INC
9	ProPlus/Cofiler	4.14E+07	**100 RFUs**; no deduced alleles reported
79	ProPlus/Cofiler	930,000	**150 RFUs**; 2 STR, 5 partial, 6 INC
16	ProPlus/Cofiler	434,600	**Not stated**; no deduced alleles reported

- **Lab 90** has **specific, detailed mixture interpretation guidelines** with worked examples and a fabulous flowchart
- **Lab 16** has **vague guidelines** that begin with "mixture interpretation is not always straightforward. Analysts must depend on their knowledge and experience…"

Questions for Consideration

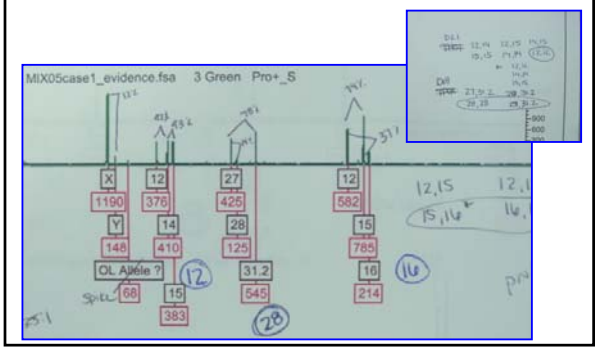
- Do you look at the evidence data first without considering the suspect's profile?
- Without a suspect, does your lab proceed with mixture interpretation?
- Do you have a decision point whereby you consider a mixture too complicated and do not try to solve it? If so, is the case declared inconclusive?
- What kind of training materials would benefit your lab in improving consistency in mixture interpretation?

Examples of MIX05 Report Formats

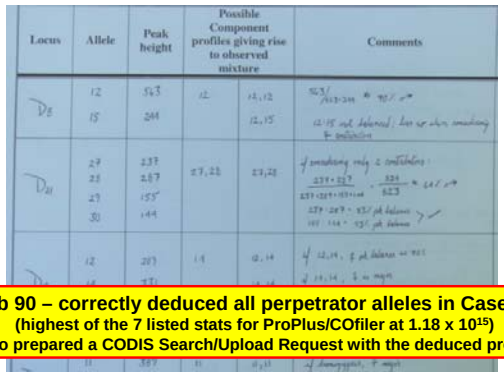
All examples with Case #1

(~3:1 mixture with female victim as the major component – and victim profile is provided)

Manual Solving of MIX05 Peak Ratios and Possible Mixture Combinations



Manually Solving Mixture Component Profiles



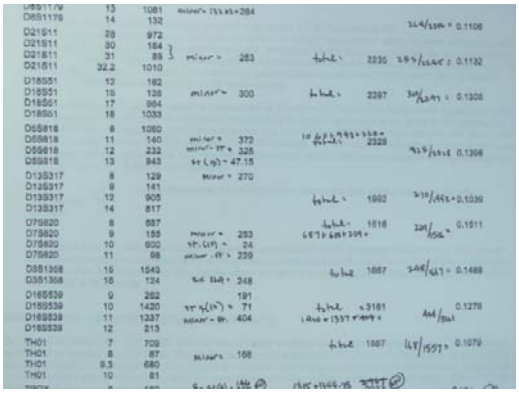
Lab 90 – correctly deduced all perpetrator alleles in Case #1
(highest of the 7 listed stats for ProPlus/COfiler at 1.18 x 10^{15})
Also prepared a CODIS Search/Upload Request with the deduced profile

A Model Report of Analysis…

- "The Profiler Plus and COfiler sample files were evaluated by **four different analysts**, using both NT and MAC analysis platforms. **The analysts checked for concordance, and a single conclusion for each mock case has been issued**."
- They detailed all assumptions made outside the course of routine casework:
  - Assumed intimate samples
  - That a comparison of deduced "foreign" alleles had been made with the perpetrator's known standard in order to calculate the significance of the inclusion with the evidentiary profile
- For Case #4: "A **Combined Probability of Inclusion was calculated** and reported for only those loci where all the alleles were above threshold [75 RFUs]. However, a minor profile(s) could not be deduced from this sample. **Please note that our laboratory may employ strategies to gain more information from the sample, such as a 10 second injection of the CE and Y-STR analysis**.

Lab 90

Another MIX05 Participant Manually Solving a Mixture



Semi-Automated Locus-by-Locus Interpretation Performed by One MIX05 Participant

D21S11

Peak	A	B	C	D
Allele	28	30	31	32.2
RFU	988	167	92	1025

Known type: K1 = 30, K2 = 31

Possible combinations

					Thresholds 70%	Thresholds 60%	Htrzyg. Pair 1	Htrzyg. Pair 2	Mixture ratio(1/2)	
28	30	AND	31	32.2	N	N	16.90%	8.98%		
28	31	AND	30	32.2	N	N	9.31%	16.29%		
28	32.2	AND	30	31	N	N	96.39%	55.09%		Best fit

D18S51

Peak	A	B	C	D
Allele	12	15	17	18
RFU	163	139	972	1047

Known type: K1 = 12, K2 = 15

Possible combinations

					Thresholds 70%	Thresholds 60%	Htrzyg. Pair 1	Htrzyg. Pair 2	Mixture ratio(1/2)	
12	15	AND	17	18	Y	Y	85.28%	92.84%	0.15	Known present
12	17	AND	15	18	N	N	16.77%	13.28%		
12	18	AND	15	17	N	N	15.57%	14.30%		

Excel spreadsheet used to examine possible component combinations

Different Reporting Formats for MIX05 Data

Different Reporting Formats for MIX05 Data

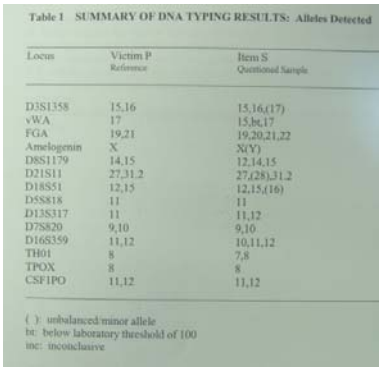


Table 1 SUMMARY OF DNA TYPING RESULTS: Alleles Detected

Locus	Victim P Reference	Item S Questioned Sample
D3S1358	15,16	15,16,(17)
vWA	17	15,bc,17
FGA	19,21	19,20,21,22
Amelogenin	X	X(Y)
D8S1179	14,15	12,14,15
D21S11	27,31.2	27,(28),31.2
D18S51	12,15	12,15,(16)
D5S818	11	11
D13S317	11	11,12
D7S820	9,10	9,10
D16S359	11,12	10,11,12
TH01	8	7,8
TPOX	8	8
CSF1PO	11,12	11,12

( ): unbalanced/minor allele
bt: below laboratory threshold of 100
inc: inconclusive

No attempt to deduce perpetrator alleles (foreign profile)

Different Reporting Formats for MIX05 Data

Profile that would be put into CODIS

LOCI	CODIS ENTRY * obligate allele	OTHER ALLELE'S IN SUSPECT'S POSSIBLE PROFILE
D3S1358	17	16,17
VWA	15*	15,17
FGA	20,22	20,22
D8S1179	12	12,12
D21S11	28*	28,31.2
D18S51	15*	15,16
D5S818	--	
D13S317	12	12,12
D7S820	--	10
D16S539	10,11*	10,11
THO1	7*	7,8 maybe
TPOX	8	8 maybe
CSF1PO	--	11,12 maybe

Kits – Profiler Plus and Cofiler
Ratio - 1:2 (perpetrator:victim)

Different Reporting Formats for MIX05 Data

	Items	
Locus	"S" Case 1 Evid.	"P" Case 1 Victim
D3S1358	15, 16, *	15, 16
D16S539	(10), 11, (12)	11, 12
AMEL	X, *	X
THO1	(7), 8	8
TPOX	8	8
CSF1PO	11, 12	11, 12
D7S820	9, 10	9, 10
vWA	(15), 17	17
FGA	19, 20, 21, 22	19, 21
D8S1179	12, 14, 15	14, 15
D21S11	27, 31.2, *	27, 31.2
D18S51	12, 15, (16)	12, 15,
D5S818	11	11
D13S317	11, 12	11

() indicates apparent minor peaks in a mixture.
"*" indicates peaks below the VFL threshold (150 rfus) for reporting.

Different Reporting Formats for MIX05 Data

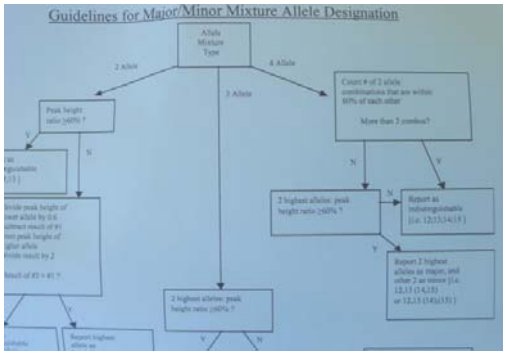
Case 1:

Item: description	D3S1358	VWA	FGA	AMEL	D8S1179	D21S11	D18S51	D5S818	D13S317	D7S820	D16S539	TH01	TPOX	CSF1PO
Pro+/CO_S: evid 1	15,16 (17)	15|17	19|20| 21|22	X,X (Y)	12|14|15	27,31.2 (28)	12,15 (16)	11,11	11|12	9|10	10|11 |12	7|8	8,8	11,12
Pro+/CO_P: victim 1 reference	15,16	17,17	19,21	X,X	14,15	27,31.2	12,15	11,11	11,11	9,10	11,12	8,8	8,8	11,12
Male interpreted from evidence 1	17	15,15 15,17	20,22	X,Y	12,12	28	16	11,11	12,12	Nd	10,11	7,7 7,8	Nd	Nd

Two allele values separated by a comma represent a genotype. Genotype calls assume biallelic donors with no null alleles.
() Indicates minor allele detected.
Single numbers and numbers separated by "|" represent an allele only designation rather than a genotype.
Interpreted profile assumes that the victim is present in the evidence mixture of two people. More than one genotype may be listed where a single genotype could not be conclusively determined. Nd=not determined due to level of results.

The community would benefit from more uniform reporting formats and mixture solving strategies…

Some Protocols Have Flow Charts to Help Make Decisions in Mixture Resolution



Quotes from One Lab's MIX05 Report

- Case 1: STR typing results from the Evidence sample indicate a DNA mixture profile. The victim cannot be excluded as a possible donor of the genetic material in the Evidence sample. No statistics will be generated at this time.
- The Evidence samples would have to be rerun in order to verify any alleles called in the final profiles. This is true for any mixed sample profiles as per our laboratory guidelines.
- **Our laboratory does not "pull out" any profile from a mixture for interpretation or statistical purposes.** The exception to this is for CODIS profiles where the alleles that can be unambiguously attributed to the victim are removed.
- **We currently do not calculate and report statistics on mixture samples.**

Lab 88

The Same Lab's "Mixture Interpretation Grid"

The Mixture Interpretation Grid provides an objective summary of how many alleles the two profiles have in common. The results will fall into one of the following categories:

"Can not be excluded"
-If the majority of alleles from the exemplar specimen are not present and/or a number of alleles foreign to the exemplar specimen are present ?

"Excluded"
-If the majority of alleles from the exemplar specimen are not included in the mixture profile

" No conclusion can be made"
-Cases where the mixture profile is limited

See laboratory mixture interpretation guidelines for further explanation.
All the cases in the study fell into the "can not be excluded" category.

Lab 88

Value of the MIX05 Study

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

- Data sets exist with multiple mixture scenarios and a variety of STR kits that **can be used for training purposes**
- A wide variety of approaches to mixture interpretation have been applied on the **same data sets evaluated as part of a single study**
- **Interpretation guidelines from many laboratories are being compared to one another for the first time in an effort to determine challenges facing future efforts to develop "expert systems" for automated mixture interpretation**
- **We are exploring the challenges of supplying a common data set to a number of forensic laboratories** (e.g., if a standard reference data set was ever desired for evaluating expert systems)

Conclusions
(Opportunities for Improvement)

- It is worth taking a closer look at protocol differences between labs to see the impact on recovering information from mixture data
- Training should help bring greater consistency
- Expert systems (when they become available and are used) should help aid consistency in evaluating mixtures and help produce more uniform reporting formats

Software Programs (Expert Systems) for Mixture Deconvolution

These programs do not supply stats (only attempt to deduce mixture components)

- Linear Mixture Analysis (LMA)
  - **Part of TrueAllele system** developed by Mark Perlin (Cybergenetics)
  - Perlin, M. W. and Szabady, B. (2001) Linear mixture analysis: a mathematical approach to resolving mixed DNA samples. *J.Forensic Sci.* 46(6): 1372-1378
- Least Squares Deconvolution (LSD)
  - Described by T. Wang (University of Tennessee) at Oct 2002 Promega meeting
  - Available for use at **https://lsd.lit.net/**
- PENDULUM
  - **Part of FSS i-3 software suite (i-STReam)**
  - Bill, M., Gill, P., Curran, J., Clayton, T., Pinchin, R., Healy, M., and Buckleton, J. (2005) PENDULUM-a guideline-based approach to the interpretation of STR mixtures. *Forensic Sci.Int.* 148(2-3): 181-189

USACIL program developed by Tom Overson

NIST Software Programs to Aid Mixture Work

Excel-based programs developed by David Duewer (NIST)

- **mixSTR** *(developed at request of Palm Beach Sheriff's Office)*
  - Does not interpret data (relies on user inputted alleles following STR data review)
  - Aids in the organization of STR mixture information
  - Considers only the presence/absence of alleles (no peak heights used)
- **Virtual MixtureMaker** *(developed to aid MIX05 sample selection)*
  - Creates mixture combinations through pairwise comparisons of input STR profiles
  - Returns information on the number of loci possessing 0,1,2,3,4,5, or 6 alleles in each 2-person mixture (also reports number of loci in each sample with 0,1,2, or 3 alleles)
  - Useful for selection of samples in mixture or validation studies with various degrees of overlapping alleles in combined STR profiles
  - Useful in checking for potentially related individuals in a population database

Programs can be downloaded from NIST STRBase web site:
http://www.cstl.nist.gov/div831/strbase/software.htm

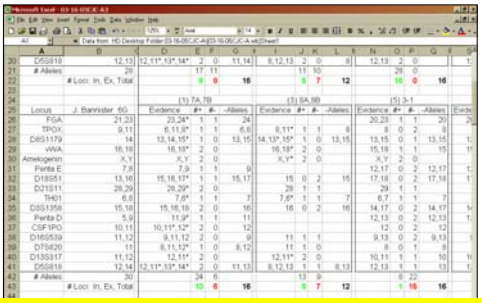
mixSTR Program

Comparisons are made between

- suspect and evidence (S/E) alleles,
- suspect and suspect (S/S) alleles **(to look for potential close relatives),**
- evidence and other evidence (E/E) sample(s) alleles **(to see how various evidentiary samples compare to one another),** and
- controls to evidence (C/E) and controls to suspect (C/S) alleles **(as a quality control contamination check).**

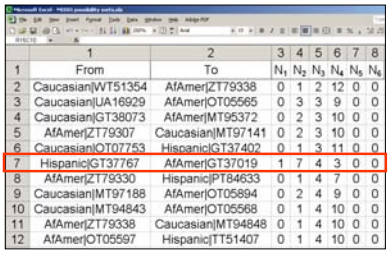
mixSTR S/E output

Data from Palm Beach County Sheriff's Office Case
Supplied by Catherine Cothran



Example of suspect to evidence (S/E) comparisons made in this case. Note that the suspect is 21,23 at FGA while the evidence contains 23,24* (* indicates that allele 24 is a minor component). Thus this suspect has allele 23 in common and is missing allele 24 in the evidence.

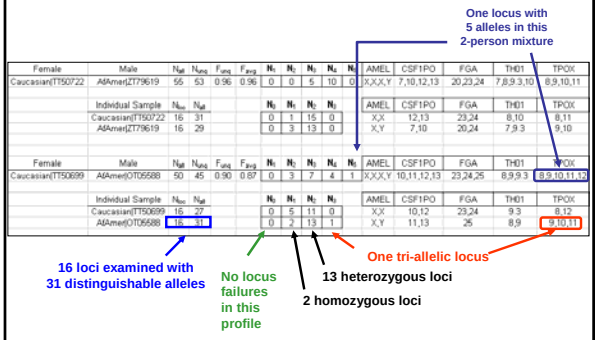
Virtual MixtureMaker Output



	1	2	3	4	5	6	7	8
1	From	To	N_1	N_2	N_3	N_4	N_5	N_6
2	Caucasian|WT51354	AfAmer|ZT79338	0	1	2	12	0	0
3	Caucasian|UA16929	AfAmer|OT05565	0	3	3	9	0	0
4	Caucasian|GT38073	AfAmer|MT95372	0	2	3	10	0	0
5	AfAmer|ZT79307	Caucasian|MT97141	0	2	3	10	0	0
6	Caucasian|OT07753	Hispanic|GT37402	0	1	3	11	0	0
7	Hispanic|GT37767	AfAmer|GT37019	1	7	4	3	0	0
8	AfAmer|ZT79330	Hispanic|PT84633	0	1	4	7	0	0
9	Caucasian|MT97188	AfAmer|OT05894	0	2	4	9	0	0
10	Caucasian|MT94843	AfAmer|OT05568	0	1	4	10	0	0
11	AfAmer|ZT79338	Caucasian|MT94848	0	1	4	10	0	0
12	AfAmer|OT05597	Hispanic|TT51407	0	1	4	10	0	0

When the STR profiles for these two individuals are combined to create a 2-person mixture, the mixture profile will contain 1 locus with a single allele, 7 loci with two alleles, 4 loci with three alleles, and 3 loci with four alleles (and no loci with 5 or 6 alleles, which is only possible if one or both samples possess tri-allelic patterns at the same STR locus).

Virtual MixtureMaker Output



Future Plans

- Develop training information based on lessons learned from the MIX05 study
- Create other useful software tools like ***mixSTR*** and ***Virtual MixtureMaker*** to increase mixture interpretation capabilities of the forensic DNA typing community
- Conduct mixture interpretation training
  - Develop worked specific examples

Some Final Thoughts…

- It is of the highest importance in the art of detection to be able to recognize out of a number of facts, which are incidental and which vital. Otherwise your energy and attention must be dissipated instead of being concentrated (Sherlock Holmes, *The Reigate Puzzle*).
- "**Don't do mixture interpretation unless you have to**" (Peter Gill, Forensic Science Service, 1998).
- Mixture interpretation consumes a large part of DNA analysts' time – software tools that improve consistency in analysis will speed casework reporting and hopefully cases solved

Conclusion


ELSEVIER

Forensic Science International

DNA mixtures in forensic casework: a 4-year retrospective study

Yolanda Torres[a,*], Inmaculada Flores[a], Victoria Prieto[a], Manuel López-Soto[a], María José Farfán[a], Angel Carracedo[b], Pilar Sanz[a]

[a]Instituto Nacional de Toxicología, A. Pestal 983, E-41080 Sevilla, Spain
[b]Instituto de Medicina Legal, Universidad de Santiago de Compostela, Sevilla, Spain

Received 27 March 2003; accepted 1 April 2003

"Mixture interpretation theory is well established and used in forensic laboratories. Most mixtures detected in casework are satisfactorily solved. But from this revision we can conclude that the behaviour of each mixed sample can be different and multifactorial and occasionally its interpretation turns out to be complicated—sometimes paralleling the importance of the evidence in the resolution of the case. In some casework mixtures our experience has proved that theoretical assumptions from studies with laboratory samples, albeit very useful, can turn out to be impracticable. **We consider that more sharing of day to day forensic laboratory problems is needed to refine our technical procedures in the resolution of specially difficult evidence."**

Acknowledgments

Funding from interagency agreement 2003-IJ-R-029 between NIJ and the NIST Office of Law Enforcement Standards

NIST Human Identity Project Team – Leading the Way in Forensic DNA…


John Butler, Margaret Kline, Pete Vallone, Jan Redman, Amy Decker, Becky Hill, Dave Duewer

Role in MIX05

- Margaret Kline (running study, sample prep, data review)
- John Butler (study design and data review)
- Becky Hill (GeneMapper*ID* data review)
- Jan Redman (Access database entry, shipping)
- Dave Duewer (*Virtual MixtureMaker* to aid sample selection; **mixSTR program**)
- Chris Tomsey & Frank Krist (FMBIO Mac data)
- Kermit Channel & Mary Robnett (FMBIO NT data)

Mandy Sozer for early discussions on study design

The many forensic scientists and their supervisors who took time out of their busy schedules to examine the MIX05 data provided as part of this interlaboratory study