



Why are mixtures hard to interpret?

The more you know the harder they
get!

Plan for Presentations

Presenter	Topic
Robin Cotton	Introduction
Charlotte Word	Complex mixture issues
<i>15 minute break</i>	
Mike Coble	Statistical issues
Robin Cotton	BU website preview
John Butler	What we have learned

***Thanks to [NIJ](#) for arranging this workshop.
And special thanks to Henry Maynard for his help
throughout the process.***

Grant to **Boston University**,

- Funded by: **National Institute of Justice** Forensic Science Training Development and Delivery Program
- **NIJ** Grant # 2008-DN-BX-K158, awarded to Biomedical Forensic Science Program at **Boston University** School of Medicine
- Thank you to  for the contributions of John Butler and Mike Coble to the workshops
- Funded by: Interagency Agreement 2008-DN-R-121 between the **National Institute of Justice** and NIST Office of Law Enforcement Standards

NIST and NIJ Disclaimer

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

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

Grant Supported Mixture Workshops

- Six workshops to practitioners
 - **Two** held at the International Symposium for Human Identification (2010 & 2011)
 - 8 hour workshop
 - ~ 150 NIJ sponsored attendees/workshop
 - **Four** regional locations
 - Miami, FL
 - Houston, TX
 - Lansing, MI
 - Phoenix, AZ
 - ~ 50 analysts per location
 - 4 hours afternoon of day 1, and 4 hours morning of day 2

Workshops Incorporated an “Audience Response System”

- We used the system available from Turning Technologies <http://www.turningtechnologies.com/>
- Allowed us to ask questions, see the participants’ answers and get their opinions **“live”**

Audience Response Components

- Each participant has a clicker
- Presenter's computer has receiver and software
- Responses to questions are received and displayed live



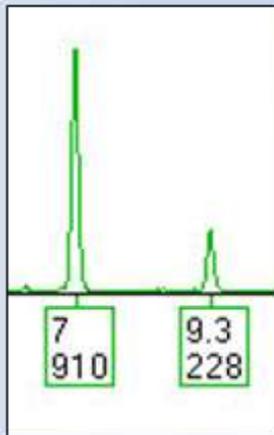
Clickers



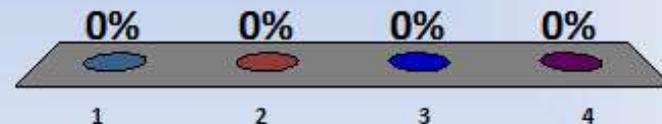
Receiver

- Use of the audience response system & *clickers*

Based on the information we have thus far,
which possibility do you like most for the
contributor genotypes?



1. 7,7 + 9.3,9.3
2. 7,9.3 + 7,7
3. 7,7 + 9.3, X
4. 2 or 3 equally likely



Grant Supported Website Development

- Website contains:
 - Lessons
 - Problems
 - Profiles



Building the Website: Lessons

Built by **BU** Information **S**ervices & **T**echnology

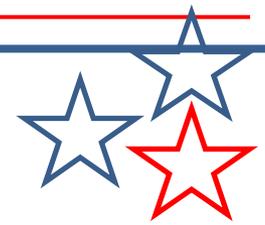
- Website is constructed on a Django platform which is an open source application for the Python language
- Development managed by Sarah Cortes
- Lessons are written off line and uploaded with associated profiles, figures and tables
 - Charlotte Word, Margaret Terrill, Catherine Grgicak, Robin Cotton
- Profiles by Catherine Grgicak, Margaret Terrill

Building the Website: Mixtures

A+B	1:19	1:9	1:4	1:2	1:1	2:1	4:1	9:1	19:1
C+D	1:19	1:9	1:4	1:2	1:1	2:1	4:1	9:1	19:1
A+B+C	2:1:1	1:2:1	1:1:2	4:1:1	1:4:1	1:1:4	1:1:1		
A+C+D	2:1:1	1:2:1	1:1:2	4:1:1	1:4:1	1:1:4	1:1:1		
A+B+C+D	2:1:1:1	1:2:1:1	1;1:2:1	1:1:1:2	2:2:1:1	1:2:2:1	1:1:2:2	1:1:1:1	

Building the Website: Profiles

- Profiles available
 - Amplify single source profiles A, B, C, and D
 - Using: ID, PP16, MI, Y
 - With 2, 1, 0.5, 0.25, 0.125, 0.0625 ng, (4 replicates of each)
 - Two person mixtures AB and CD
 - Using: ID, PP16, MI, Y
 - With 2, 1, 0.5, 0.25, 0.125, 0.0625 ng
 - Three person mixtures A,B,C and A,C,D
 - Using: ID, PP16, MI, Y
 - With 4, 2, 1, 0.5, 0.25, 0.125 ng
 - Four person mixture A,B,C,D
 - Using ID
 - With 4, 2, 1, 0.5, 0.25, 0.125 ng



Some perspective on mixtures:

- **Why are mixtures difficult?**
- It seems that the more you know the harder they get!
- **The answer is twofold:**
- We are working with **evidence**,
 - 1- We do not know the **number or ratio** of contributors before testing the sample

If a sample is a 1:1 mixture of source A and B:

Amount of DNA	Total Cells in sample	~ # of cells from each component
1ng	143	71
0.5ng	71	36
0.25ng	36	18
0.125ng	18	9
0.0625ng	9	4

If a sample is a 1:9 mixture of source A and B :

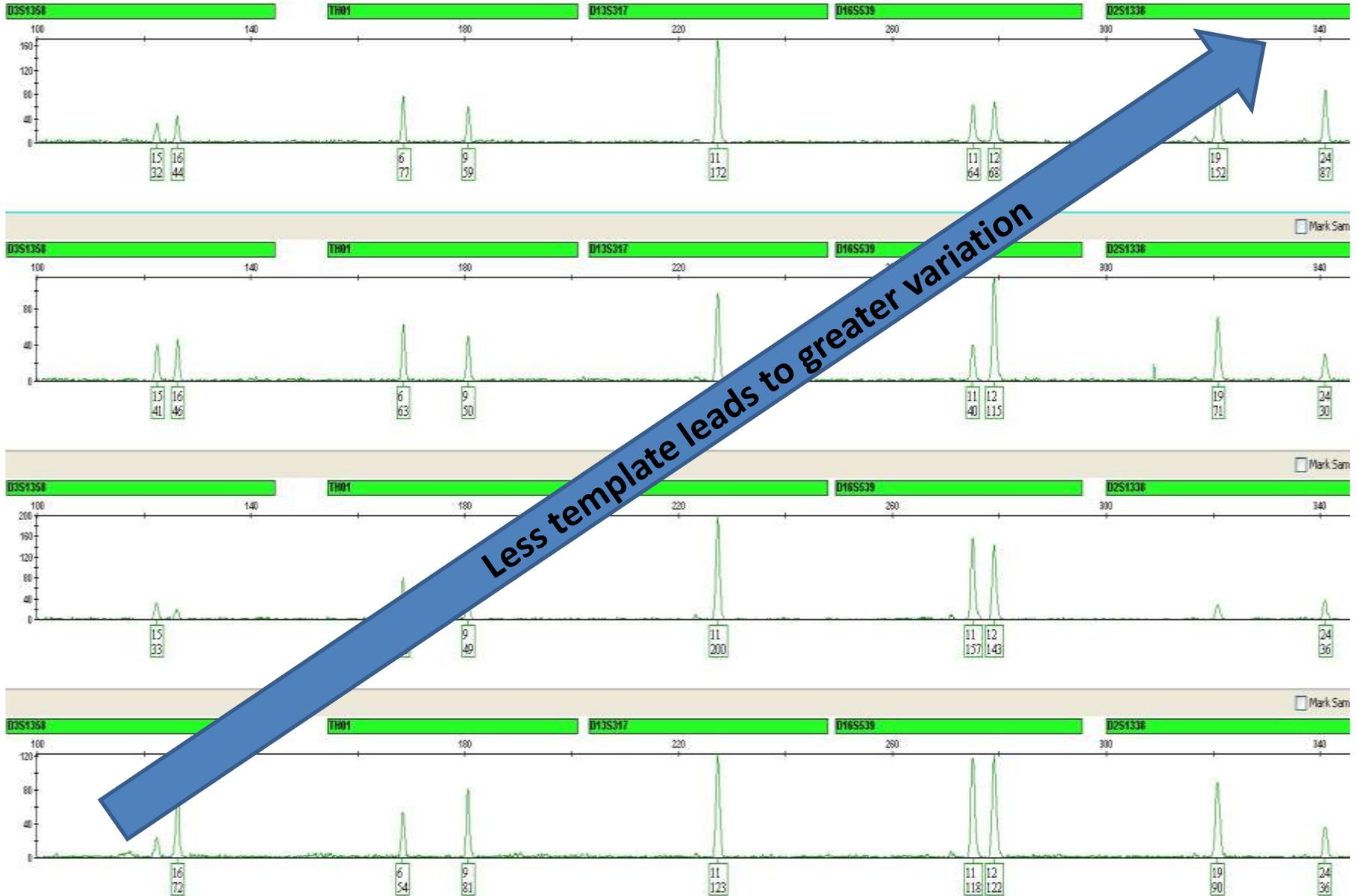
Amount of DNA	~ # of cells from major component	~ # of cells from minor component
1ng	129	14
0.5ng	64	7
0.25ng	32	4
0.125ng	16	2
0.0625ng	8	1

Some perspectives on mixtures:

- **The answer is:**

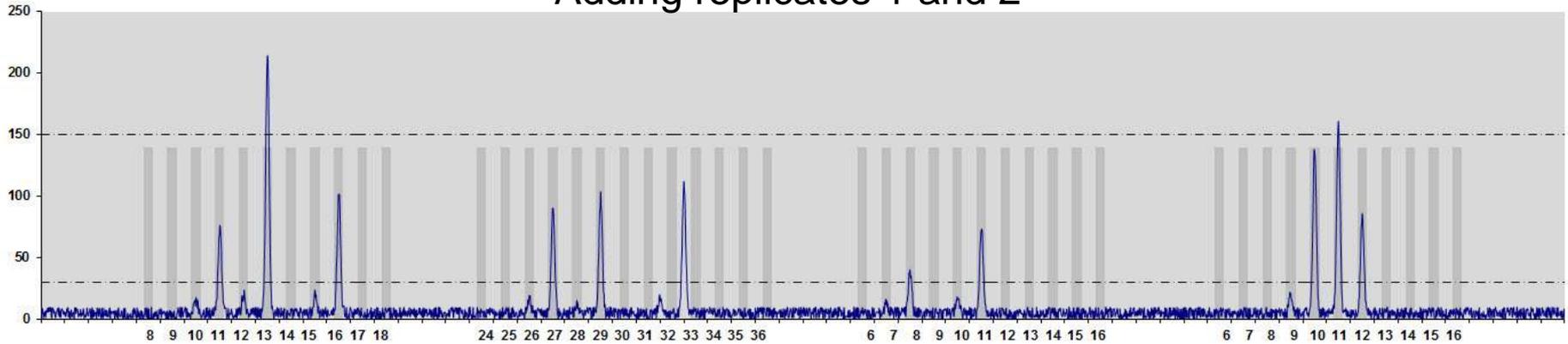
- 1- We do not know the **number or ratio** of contributors before testing the sample
- **and**
- 2- We cannot control the PCR chemistry sufficiently to prevent variation in the amount of product produced for two alleles at the same locus even in a single source sample.
- Therefore we have **peak height** and **peak height ratio variation**

Amplification replicates of 125 pg from single master mix

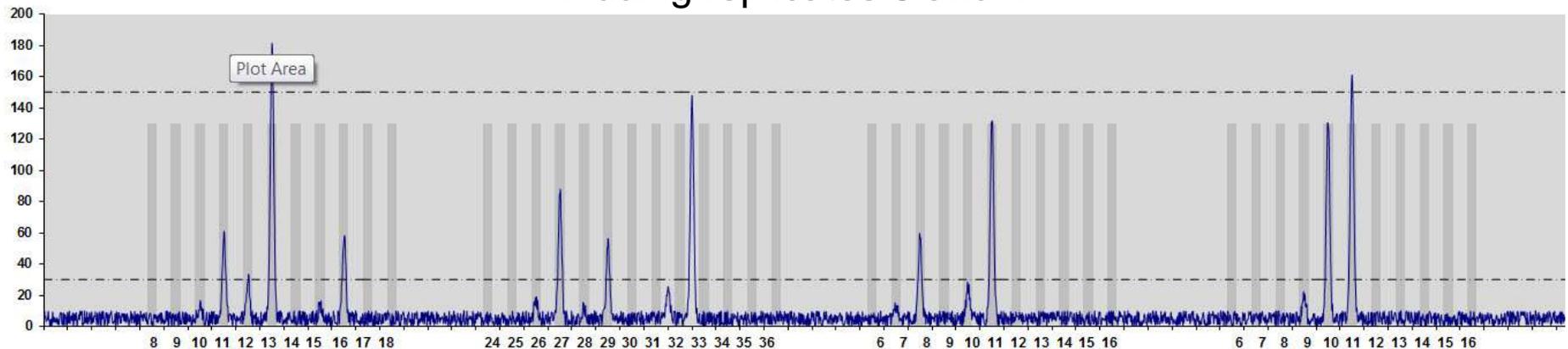


Artificial 1:1 mixtures made from repeat amplifications of 0.125ng of a single sample.

Adding replicates 1 and 2



Adding replicates 3 and 4



Variation is everywhere:

- Without understanding the basics of the PCR and the intrinsic variation we cannot do the complicated
- We cannot do the complicated by “analyst experience”
- For many mixtures our experience can no longer account for all the variables

Slight digression: where have we been?

- 1998-2000 large STR multiplexes are developed & begin to be used
- 1998 two papers by Gill and co-authors
 - Lay out the basics of interpretation of two person mixtures
 - Introduce and describe a method for computer analysis of 2 person mixtures

Gill, P., et al. (1998). Interpreting simple STR mixtures using allelic peak areas. *Forensic Science International*, 91, 41-53.

Gill, P., et al. (1998). Interpretation of simple mixtures when artifacts such as stutters are present—with special reference to multiplex STRs used by the Forensic Science Service. *Forensic Science International*, 95, 213-224.

Moving ahead a little:

- 2001- Perlin and Szabady publish a framework for mathematical approaches to mixture analysis
 - Perlin, MW, Szabady, B (2001). Linear mixture analysis: a mathematical approach to resolving mixed DNA samples. *JFS* 46: 1372-1377
- 2005- Gill publishes a simulation model of the DNA process that describes the impact of probability on the final results of DNA testing
 - Gill, P., et al. (2005). A graphical simulation model of the entire DNA process associated with the analysis of short tandem repeat loci. *Nucleic Acids Research*, 33, 632-643.
- 2006-ISFG guidelines on Mixture Analysis
 - Gill, P., et al. (2006). DNA commission of the International Society of Forensic Genetics: Recommendations on the interpretation of mixtures. *Forensic Science International*, 160, 90-101.

2006 to 2012

- Many articles related to mixture analysis presenting a picture of ever increasing complexity:
 - Variability in stutter %
 - Probability of drop out
 - Detailed analysis of analytical threshold which may vary with dye color and amount of template
 - Analysis and impact of allele sharing when attempting to determination number of contributors to a mixture
 - Further software development and validation

What have we been doing in this time frame:

- Labs rapidly converted to STR analysis
- Accreditation became the norm
- CODIS database has grown from zero to 10,718,700 samples
- Case samples in the database are now 427,500
- Hits have grown from zero to a total of 178,300
- More hits ---- more successes ---- more samples
---- more mixtures!

Analysis of backlog rape kits

- Massively supported by NIJ
- Begins about 2003 and still continues
 - Many cases done in private laboratories
- Many samples contain two person mixtures
- Subtraction of victim's known type allows deduction of unknown contributor and upload to CODIS
 - No need to set aside suspect's profile, there was no suspect
- More success ---- **more samples** ---- **more mixtures!**

Following successes in Britain:

- DNA is extended to less serious crimes
 - Burglaries
 - Car thefts
 - Analysis of weapons
 - Clothes
- This produces
 - Low template DNA &
 - **More mixtures**



Everyone makes **The Leap**

- If we can do two person mixtures we can also do “**more** person” mixtures!
- And.....it can still be simple! All we need is-
 - A stochastic threshold and
 - A CPI stat

Thresholds

- 2009-Budowle et al. publish mixture interpretation paper advocating use of **PAT** (Peak Amplitude Threshold) and **MIT** (Match Interpretation Threshold)
- When analyzing mixtures, if all peaks above **PAT** are not also above **MIT** then do not use data, stochastic effect possible
- 2010 SWGDAM Interpretation Guidelines follow these concepts but allow other approaches as an alternative to using an analytical and a stochastic threshold.

Budowle, B., et al. (2009). Mixture interpretation: defining the relevant features for guidelines for the assessment of mixed DNA profiles in forensic casework. *Journal of Forensic Sciences*, 54, 810-821.

What's wrong with this picture?

- There is nothing simple about the variation which is observed in mixtures from multiple contributors
- “The use of bounds **applied to data that show continuous variation** is common in forensic science and is often a pragmatic decision. However it should be borne in mind that applying such bounds has arbitrary elements to it and that **there will be cases where the data lie outside these bounds.**”

Why are we reluctant to embrace the complexities of our system?

- The courts do not appear to embrace complexity; the lawyers and judges want us to make the complicated into the simple
- Many lab directors would prefer something simple --- complexity and production do not easily go hand in hand
- The NAS doesn't recognize that DNA mixture interpretation procedures used in the US are not generally keeping pace with the literature on the topic or practice in Europe, NZ and Australia. NAS gives DNA a pat on the back for being *scientific*.

And....

- Amount of learning required on our part is, in many cases, extensive
- There is no requirement for serious continuing education
 - Which means there may not be enough funding or time to read, study, take a course, etc.

In every workshop presented and supported by our NIJ grant:

- Participants said they needed more training in
 - Mixture analysis
 - Statistics related to mixtures
- **Every one agrees: more training is needed**
- Like other learning situations - reading or attending a workshop, without further work, may not = learning

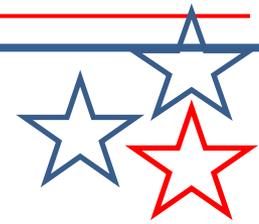
What form should the training take:

- Workshops are good
 - Mixture analysis
 - Statistics
 - Low copy number
 - Difficult samples
 - Testimony skills
- But these are a one day fix to a larger learning gap

Solutions are coming but we're not there yet!

- We need:
 - to have courses available
 - requirement to take the courses
 - time designated for this purpose
- We hope the DNA Mixture Website, funded by NIJ, will be part of the solution:
 - Provide training in basic knowledge and skills
 - Can be accessed anywhere
 - Profiles can be used for any purpose
 - Website available to everyone for analysis, training, validation...

Ultimately



- Implementation of computer software approaches which can model variation, take us away from thresholds and add information for our use in analysis and reporting
- More extensive training in statistical approaches and the use of likelihood ratios will make better use of the data and ultimately benefit the criminal justice system