2012 Mixture Interpretation Workshop:

Mixtures Using *SOUND* Statistics, Interpretation, & Conclusions

Introduction to Workshop

Robin W. Cotton

October 15, 2012

Nashville, TN

BOSTON UNIVERSITY

NIST

Welcome
to Nashville, TN

Thank you Promega for having us back this year!!

Your Presenters are:

Robin Cotton
Boston University

617-638-1952
rwcotton@bu.edu

John Butler
NIST

301-975-4049
john.butler@nist.gov

Catherine Grgicak
Boston University

617-638- 1968
cgrgicak@bu.edu

Mike Coble
NIST

301-975-4330
michael.coble@nist.gov

Charlotte Word
Consultant

301-527-1350
cjword@comcast.net

Presenters

- John Butler
  - Ph.D. in Analytical Chemistry, University of Virginia
  - 20 years experience
  - Writes books on the side
  - The engine behind STRBase

- Mike Coble
  - Ph.D. in Genetics, George Washington University
  - 15 years DNA experience
  - Mitochondrial DNA and STRs at AFDIL
  - Now working even harder at NIST

Presenters

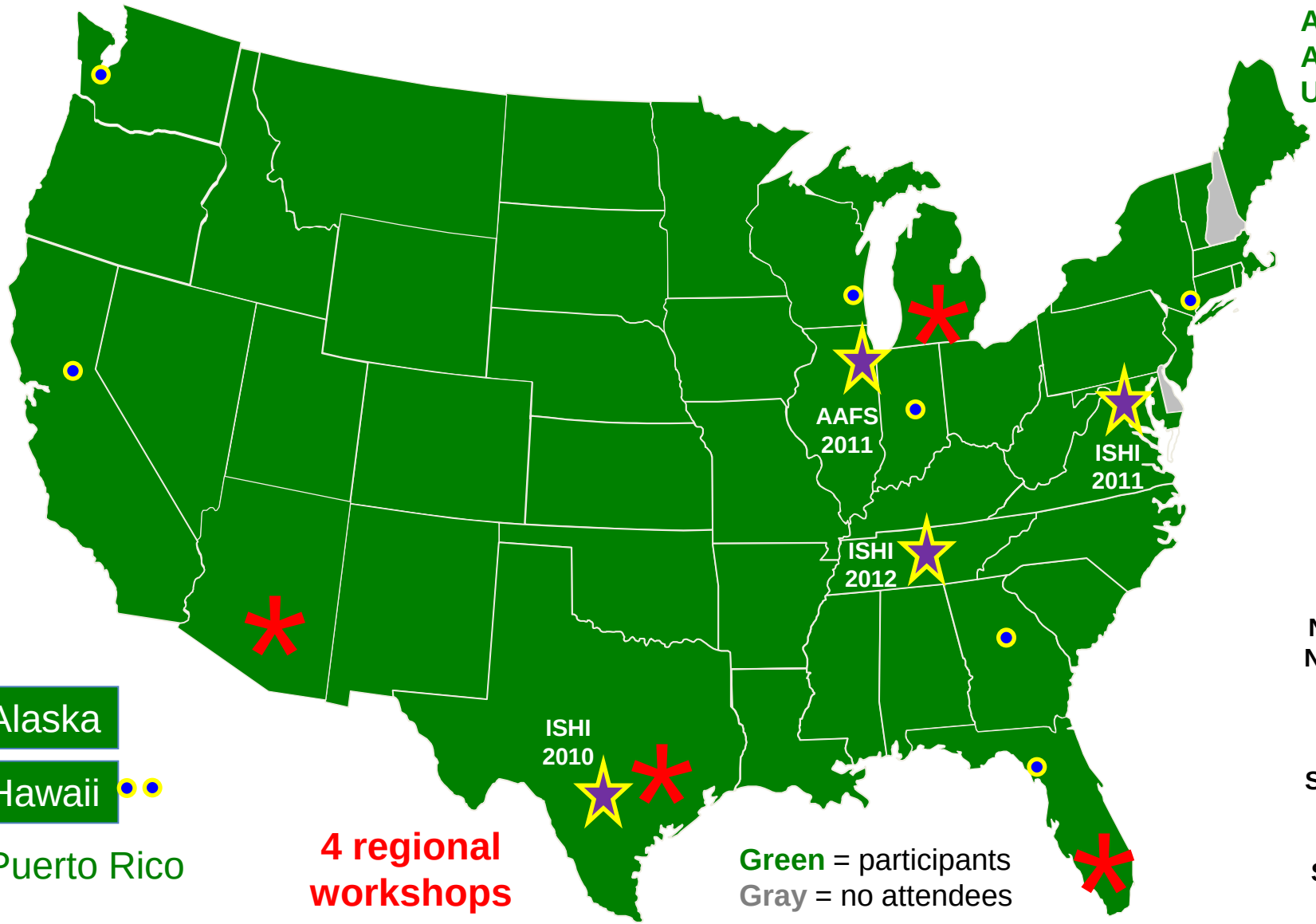
- Charlotte Word
  - Ph.D. in Microbiology, University of Virginia
  - 22 years casework and technical review experience for both public and private laboratories
  - Well over 200 court testimonies in admissibility hearings and trials
  - Currently a private consultant in the Washington DC area

Presenters

- Robin Cotton
  - Ph.D. Molecular Biology and Biochemistry, University of California at Irvine
  - 18 years casework and testimony experience
  - Boston University School of Medicine since 2006
  - Program Director, Biomedical Forensic Sciences
- Catherine Grgicak
  - M.S. Forensic Science, University of Alabama
  - Ph.D. Chemistry, University of Ottawa
  - 3 years experience as DNA Analyst
  - Boston University School of Medicine since 2007

Mixture Workshop Attendees

49 states and 25 other countries, so far:



Alaska

Hawaii

Puerto Rico

4 regional workshops

Green = participants
Gray = no attendees

Federal Labs
FBI
ATF
AFDIL
USACIL

Algeria
Argentina
Bahamas
Belgium
Brazil
Canada
Croatia
Finland
France
Israel
Italy
Jamaica
Japan
Korea
Mexico
Netherlands
New Zealand
Panama
Peru
Russia
Saudi Arabia
Singapore
Spain
Switzerland
UK

Why are mixtures difficult?

- It seems that the more you know, the harder they get!

- **The answer is twofold:**
  - We are working with **evidence**,
  - We do not know the number or ratio of contributors before testing the sample

Why are mixtures difficult?

- **The answer is: We are working with evidence**

A. We do not know the **number** or **ratio** of contributors before testing the sample

– **and**

B. 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**

Variation is everywhere:

- Without understanding the basics of the PCR and the intrinsic variation produced, we cannot interpret the complicated profiles.

- We cannot interpret the complicated profiles using “analyst experience”.

- For many mixtures our “experience” can no longer account for all the variables.

Slight digression: How did we get in this position? Should we have been smarter?

- 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 2 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

- Most current articles related to mixture analysis present a picture of ever increasing complexity:
  - Variability in stutter %
  - Variation in probability of drop out with amount of DNA (component) amplified
  - 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 9,812,200 samples
- Case samples in the database are now 441,200
- Hits have grown from zero to a total of 185,000
- 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 **S**tochastic **T**hreshold &
  - a **C**ombined **P**robability of **I**nclusion statistic

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 the 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**."

Bright, J.A., et al. (2010). Examination of the variability in mixed DNA profile parameters for the Identifiler multiplex. *Forensic Science International: Genetics, 4,* 111-114.

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

- The courts do not appear to embrace complexity; 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 does not 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, New Zealand and Australia. NAS gives DNA a pat on the back for being *scientific*.

And….

- The amount of learning required on our part is, in many cases, extensive

- There is no requirement in the FBI QA Standards for serious continuing education
  - Which means there may not be enough funding for additional training, time to read, study or take a course, etc.

What forms should training take:

- Workshops are good
  - Mixture analysis
  - Statistics
  - Low copy number
  - Difficult samples
  - Testimony skills

Neither

- But these are a one-day fix to a larger learning gap

(or John Butler) is the real solution to the learning gap!

Solutions are coming but we're not there yet!

- Implementation of computer software approaches which model variation & remove the need for "line in the sand" thresholds will 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 data and ultimately benefit the criminal justice system.

- We need training courses available, requirement to take the courses & time designated for this purpose

Today's Workshop incorporates an "Audience Response System"

- The system comes 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

NIST Disclaimer

Points of view are those of presenters and do not necessarily represent the official position or policies of 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.