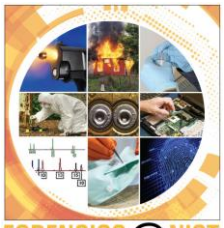
NIST FORENSIC SCIENCES



FORENSICS@NIST

NIST Applied Genetics

BOSTON UNIVERSITY

NIST National Institute of Standards and Technology • U.S. Department of Commerce

FORENSICS@NIST

DNA Mixture Interpretation Webcast
April 12, 2013

http://www.nist.gov/oles/forensics/dna-analyst-training-on-mixture-interpretation.cfm

http://www.cstl.nist.gov/strbase/mixture.htm

Welcome and Opening Remarks

John Paul Jones II
National Institute of Standards and Technology

Important Information

- >1000 registered for this event
- Printable slides are available on conference website
- Potential screen resolution issues
- iPhones & iPads – potential viewing challenges because of Flash requirement – can answer poll questions though
- 21 second delay in broadcast
- Survey Monkey – do you have your cell phone handy?
- Scheduled times are approximate
- Questions – email to forensic@nist.gov (may be read by moderator during webcast – will keep source anonymous)
- Twitter Chat: #NISTForensics
- Certificates of Completion – follow online instructions (TL)
- Webcast Archive: recording webcast and be available for on-demand viewing in a few weeks following transcription

Lets Try a Sample Survey Monkey Question

(Remember there is a 21 second delay)

- Please use your computer or cell phone web browser to click on the link to access our Polls:
- http://go.usa.gov/TaGB
- Poll Question 1: Please tell us what type of laboratory you work for (select the best single answer)
  - Federal
  - State
  - Local
  - Municipal
  - University
  - Private
  - Other (including individuals not employed by a laboratory)

Webcast Format for Training

- **With cuts in federal budgets, webcasts or webinars may become more appealing in the future to reduce costs in providing training**
- Please let us know about any technical difficulties that you may have faced so that we can improve future webcasts
- We welcome suggestions for additional content or topics to cover in future webcast training events
- Please contact John Paul Jones at 301-975-2782 or john.jones@nist.gov

Posting of Video from this Event

- Following transcription of this webcast (this process takes up to a month), **we plan to post videos of each presentation on the conference website**
- All those who registered for this event (onsite or online) will receive email notification when the material is posted.
- Due to costs of maintaining large video files on NIST servers, **webcast videos may only be available for a limited time** (we are planning on at least six months)
- A link to the webcast video website will also be available from the STRBase mixture website to enable future viewing or downloading of video or presentation materials

Concern for Potential Misuse of Webcast Presentations

- We remind current and future viewers that presentations reflect the presenters' opinions at the time they were given on April 12, 2013
- Please do not take any specific comments of the webcast presenters out of context in order to advance either scientific or legal arguments
- Science advances with new discoveries and therefore scientific opinions may change over time given exposure to new ideas or techniques

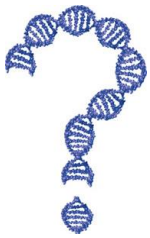
Disclaimer

NIST Disclaimer: 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 it imply that any of the materials, instruments or equipment identified are necessarily the best available for the purpose.

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

http://www.nist.gov/oles/forensics/dna-analyst-training-on-mixture-interpretation.cfm

On behalf of the team that put this together:
We hope you benefit from this webcast!!!



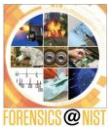
Contact Information

John Paul Jones II
NIST Law Enforcement Standards Office
john.jones@nist.gov
301-975-2782

http://www.nist.gov/oles/forensics/

Additional DNA mixture information available at:
http://www.cstl.nist.gov/strbase/mixture.htm

NIST National Institute of Standards and Technology • U.S. Department of Commerce



DNA Mixture Interpretation Webcast
April 12, 2013

http://www.nist.gov/oles/forensics/dna-analyst-training-on-mixture-interpretation.cfm

http://www.cstl.nist.gov/strbase/mixture.htm

Introduction

Robin W. Cotton
Boston University School of Medicine
Biomedical Forensic Sciences

BOSTON UNIVERSITY

Introduction

- We are all seeing or being asked questions that show limited understanding of the science involved in reliable DNA interpretation
- Need to be prepared to go back and examine old cases with new SOPs to test reliability
- People are making decisions based on reports – there are scientific and ethical issues involved
- We have to be scientists first – then we can transition it into the legal realm of the court room
- Whatever our background, we need to seek help from others to do our job well
- The samples being tested now are not what have been validated in many labs (single-source or 2-person mixtures)

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” and our original kit validations can no longer account for all the variables.

In the last 15 years:

- From 1998-2000 large STR multiplex kits were developed and put into use for forensic casework.
- Labs rapidly converted to STR analysis
- Accreditation became the norm
- CODIS (NDIS) database has grown from zero to 10,142,600 offender samples (as of Jan 2013)
- Case samples in the database are now 422,500
- Hits have grown from zero to a total of 200,300
- 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 **S**tochastic **T**hreshold &
  - a **C**ombined **P**robability of **I**nclusion statistic

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, is extensive.
- The FBI QA Standards require 8 hours of continuing education/year which is not enough.
- 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. (This will also require training.)
- 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.
- Math phobia is out-get rid of it!

Lastly…

- Collectively, in talking to people across the country, we see a continued need for improvement.
- Of course there will be cases that were reported using an older SOP after the lab has implemented a more "mixture savvy" SOP.
- There will be instances when old reports need to be updated with new interpretation.
- This is the only scientifically appropriate route.
- These changes and adjustments are manageable and within our collective capability.

NIST National Institute of Standards and Technology • U.S. Department of Commerce

DNA Mixture Interpretation Webcast
April 12, 2013

http://www.nist.gov/oles/forensics/dna-analyst-training-on-mixture-interpretation.cfm

http://www.cstl.nist.gov/strbase/mixture.htm

Introduction to Interpretation Issues

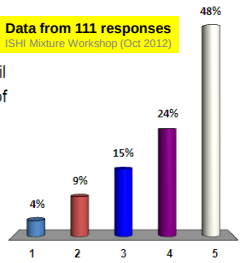
John M. Butler
National Institute of Standards and Technology

2012 Response at ISHI Workshop

Which of the topics below would be your first choice for additional training?

1. Relevant literature
2. How to validate thresholds in more detail
3. Reporting and the use of assumptions
4. **Interpretation of low level mixtures**
5. **Likelihood ratios and other statistical approaches**

Data from 111 responses
ISHI Mixture Workshop (Oct 2012)



~75% want more information on these topics

Planned Presentation Outline

- Overview/thoughts on interpretation & statistics
- SWGDAM 2010 interpretation guidelines
- Thoughts on setting thresholds
- Problems with CPI/CPE statistics
- Take home messages

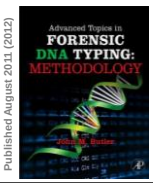
Steps Involved in Process of Forensic DNA Typing

1) Data Interpretation
2) Statistical Interpretation

Gathering the Data | **Understanding the Data**

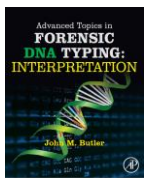
Collection/Storage/ Characterization → Extraction/ Quantitation → Amplification/ Marker Sets → Separation/ Detection → **Interpretation** → **Report**

Advanced Topics: Methodology



Published August 2011 (2012)

Advanced Topics: Interpretation



Hopefully Fall 2014 (delayed)

Importance of Improved Understanding Regarding DNA Mixture Interpretation

- Each DNA analyst may think his or her approach is correct – but misinterpretations have given rise to a variety of approaches being undertaken today, some of which are not correct…
- I believe that **a better understanding of general principles will aid consistency and quality of work being performed**

What We Hope to Accomplish with this NIST Webcast

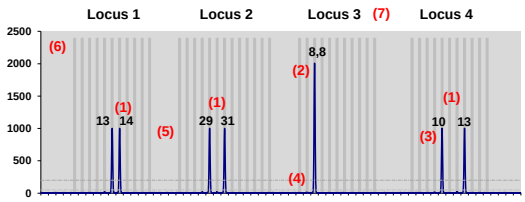
Desired Learning Outcomes:

- Explore how the analytical threshold and stochastic threshold affect data analysis, interpretation, conclusions and statistical calculations in mixed DNA profiles
- Examine approaches for establishing one or more analytical thresholds and stochastic thresholds for casework
- Enhance knowledge of mixture interpretation and presentation of results, conclusions and opinions

Many Labs are in the Process of Changing their Protocols

Perhaps lowering the expected peak height ratio (PHR) from 70% down to 55% when interpreting DNA mixtures?

Using **Ideal Data** to Discuss Principles



(1) 100% PHR between heterozygous alleles
(2) Homozygotes are exactly twice heterozygotes due to allele sharing
(3) No peak height differences exist due to size spread in alleles (any combination of resolvable alleles produces 100% PHR)
(4) No stutter artifacts enabling mixture detection at low contributor amounts
(5) Perfect inter-locus balance
(6) Completely repeatable peak heights from injection to injection on the same or other CE instruments in the lab or other labs
(7) Genetic markers that are so polymorphic all profiles are fully heterozygous with distinguishable alleles enabling better mixture detection and interpretation

Challenges in Real-World Data

- **Stochastic (random) variation** in sampling each allele during the PCR amplification process
  - This is highly affected by DNA quantity and quality
  - Imbalance in allele sampling gets worse with low amounts of DNA template and higher numbers of contributors
- **Degraded DNA** template may make some allele targets unavailable
- **PCR inhibitors** present in the sample may reduce PCR amplification efficiency for some alleles and/or loci
- **Overlap of alleles** from contributors **in DNA mixtures**
  - Stutter products can mask true alleles from a minor contributor
  - Allele stacking may not be fully proportional to contributor contribution

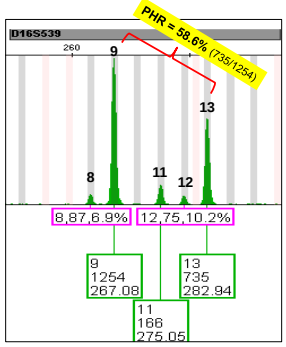
D.N.A. Approach to Understanding

- **D**octrine or Dogma (why?)
  - A fundamental law of genetics, physics, or chemistry
    - Offspring receive one allele from each parent
    - Stochastic variation leads to uneven selection of alleles during PCR amplification from low amounts of DNA templates
    - Signal from fluorescent dyes is based on …
- **N**otable Principles (what?)
  - The amount of signal from heterozygous alleles in single-source samples should be similar
- **A**pplications (how?)
  - Peak height ratio measurements can associate alleles into possible genotypes

Results Depend on Assumptions

- "Although courts expect one simple answer, statisticians know that **the result depends on how questions are framed and on assumptions tucked into the analysis**."
  - Mark Buchanan, Conviction by numbers. *Nature* (18 Jan 2007) 445: 254-255
- SWGDAM 2010 Interpretation Guideline 3.6.5
  - "Because **assumptions** regarding the origin of evidence or the number of contributors to a mixture **can impact comparisons**, the laboratory should **establish guidelines for documenting any assumptions that are made when formulating conclusions**"

Example: D16S539 from Profile 1

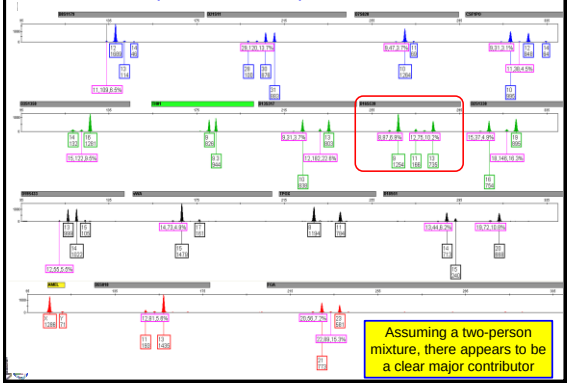


PHR = peak height ratio; also known as heterozygote balance (Hb)

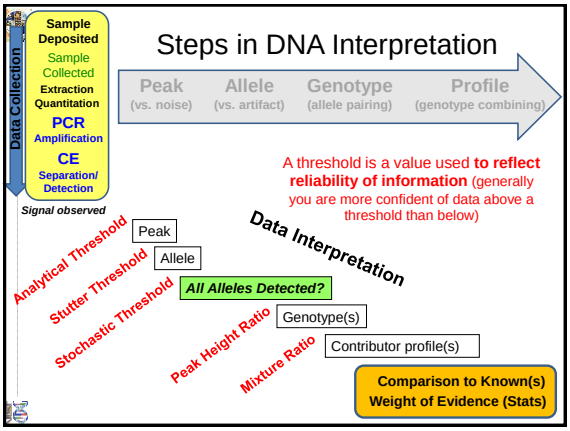
Some Observations:

- Depending on expected PHR, alleles 9 and 13 may or may not be associated into a genotype (<60%)
- Allele 11 could be paired with 8, 9, 12, or 13 or itself (11,11 homozygote) depending on stochastic threshold
- Alleles 8 and 12 could be stutter products or possibly be paired with allele 11

Profile 1 (stutter filter off)



Steps in DNA Interpretation



Overview of the SWGDAM 2010 Interp Guidelines

http://www.swgdam.org/Interpretation_Guidelines_January_2010.pdf

1. Preliminary evaluation of data – **is something a peak and is the analysis method working properly?**
2. Allele designation – **calling peaks as alleles**
3. Interpretation of DNA typing results – **using the allele information to make a determination about the sample**
   1. Non-allelic peaks
   2. Application of peak height thresholds to allelic peaks
   3. Peak height ratio
   4. Number of contributors to a DNA profile
   5. Interpretation of DNA typing results for mixed samples
   6. Comparison of DNA typing results
4. Statistical analysis of DNA typing results – **assessing the meaning (rarity) of a match**

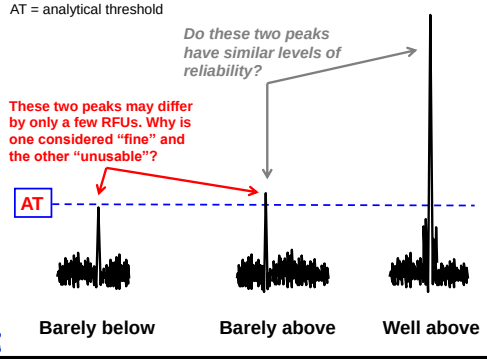
Other supportive material: statistical formulae, references, and glossary

Principles Behind Thresholds

Thresholds (example values)	Principles Behind *(if properly set based on lab- & kit-specific empirical data)*
Analytical Threshold (e.g., 50 RFU)	Below this value, observed peaks cannot be reliably distinguished from instrument noise (baseline signal)
Limit of Linearity (e.g., 5000 RFU)	Above this value, the CCD camera can become saturated and peaks may not accurately reflect relative signal quantities (e.g., flat-topped peaks) and lead to pull-up/ bleed-through between dye color channels
Stochastic Threshold (e.g., 250 RFU)	Above this peak height value, it is reasonable to assume that allelic dropout of a sister allele of a heterozygote has not occurred at that locus; single alleles above this value in single-source samples are assumed to be homozygous
Stutter Threshold (e.g., 15%)	Below this value, a peak in the reverse (or forward) stutter position can be designated as a stutter artifact with single-source samples or some mixtures (often higher with lower DNA amounts)
Peak Height Ratio (e.g., 60%)	Above this value, two heterozygous alleles can be grouped as a possible genotype (often lower with lower DNA amounts)
Major/Minor Ratio (e.g., 4:1)	When the ratio of contributors is closer than this value in a two-person mixture, it becomes challenging and often impossible to correctly associate genotype combinations to either the major or minor contributor

What is the meaning of a threshold?

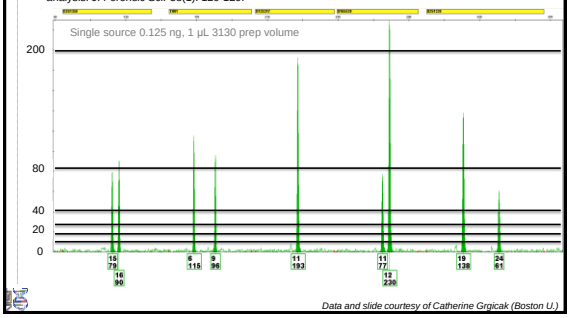
AT = analytical threshold



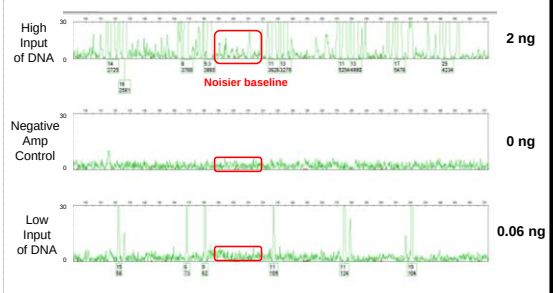
Impact of Various Analytical Thresholds

Rakay et al. (2012) Maximizing allele detection: effects of analytical threshold and DNA levels on rates of allele and locus drop-out. *Forensic Sci. Int. Genet.* 6: 723-728.

Bregu et al. (2013) Analytical thresholds and sensitivity: establishing RFU thresholds for forensic DNA analysis. *J. Forensic Sci.* 58(1): 120-129.



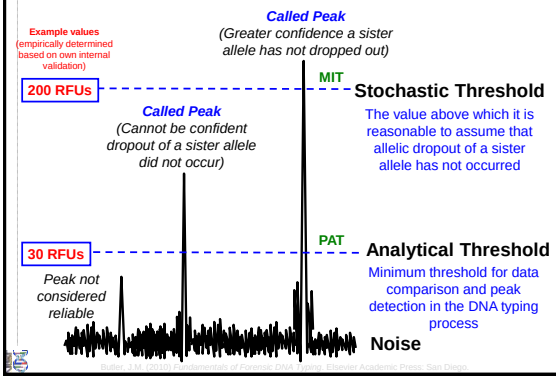
Baselines Positives ≠ Baselines Negatives



Bregu et al. (2013) Analytical thresholds and sensitivity: establishing RFU thresholds for forensic DNA analysis. *J. Forensic Sci.* 58(1): 120-129.

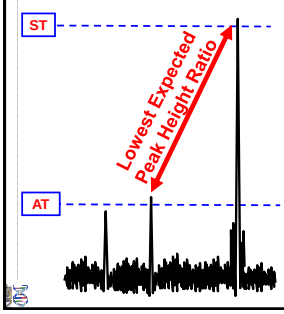
Data and slide courtesy of Catherine Grgicak (Boston U.)

Overview of Two Thresholds



Butler, J.M. (2012) Advanced Topics in Forensic DNA Typing: Interpretation. Elsevier Academic Press, San Diego.

Stochastic and Analytical Thresholds
Impact Lowest Expected Peak Height Ratio



The lower you go trying to analyze low-level data… (i.e., more sensitive STR kits)

the worse your expected peak height ratios for single-source samples

Therefore, there is greater uncertainty with associating genotypes of contributors in mixtures (or even determining that you have a mixture)

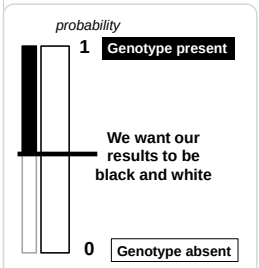
Keep in Mind…

"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**."

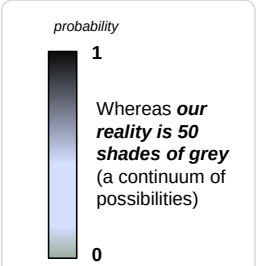
Bright, J.A., et al. (2010). Examination of the variability in mixed DNA profile parameters for the Identifiler multiplex. *Forensic Science International: Genetics, 4,* 111-114.

Adapted from a slide by Peter Gill, Rome meeting, April 27-28, 2012: The hidden side of DNA profiles: artifacts, errors and uncertain evidence

Approaches to Data Interpretation:
Binary vs Probabilistic



Binary Approach



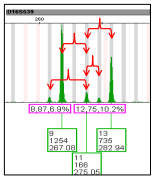
Probabilistic Approach

Is There Uncertainty in the Data?

- **If allele dropout is a possibility** (e.g., in a partial profile), then there is uncertainty in whether or not an allele is present in the sample…and therefore what genotype combinations are possible
- **If different allele combinations are possible** in a mixture, then there is uncertainty in the genotype combinations that may make up the mixture result…

Minor contributor at D16S539 could be:
8,11 or
9,11 or
11,12 or
11,13 or
11,11

Possible allele pairing with the 11 allele



Uncertainty and Probability

- "Contrary to what many people think, **uncertainty is present throughout any scientific procedure**."
  - Dennis V. Lindley, in his foreword to Aitken & Taroni (2004) *Statistics and the Evaluation of Evidence for Forensic Scientists, Second Edition*
- "It is now recognized that **the only tool for handling uncertainty is probability**."
  - Dennis V. Lindley, in his foreword to Aitken & Taroni (2004) *Statistics and the Evaluation of Evidence for Forensic Scientists, Second Edition*

Conference Held in Rome in April 2012

Peter Gill

University of Oslo, Norway

- "If you are going to have a threshold, at least try to associate it with a level of risk. You can have a threshold any where you like, but the lower the [stochastic] threshold, the greater the risk is of wrongful designation [of genotypes]. The higher the threshold, the more likely you will have an inconclusive result."

Rome meeting, April 27-28, 2012: The hidden side of DNA profiles: artifacts, errors and uncertain evidence

David Balding

- "In ideal analysis, we would never use thresholds, but in practice they are useful. I don't think we have sophisticated enough models in many situations to understand all of the details of the data. **Thresholds provide a simplification.** That is reasonable as long as they are backed up by calibration evidence."

Rome meeting, April 27-28, 2012: The hidden side of DNA profiles: artifacts, errors and uncertain evidence

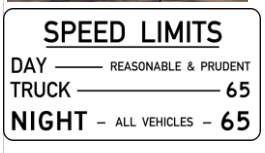
Bruce Budowle

University of North Texas Health Science Center

- "**We put thresholds in place to help protect us from risk of making wrong decisions**. They have value."
- **Compares thresholds to speed limits**, which are set for safety reasons

Rome meeting, April 27-28, 2012: The hidden side of DNA profiles: artifacts, errors and uncertain evidence

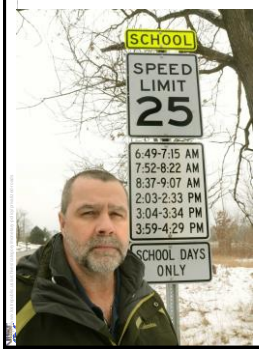
Do you leave thresholds and protocols up to "analysts' discretion"?



Typical speed limit sign that one would see at the Montana state line from December 1995 to June 1999

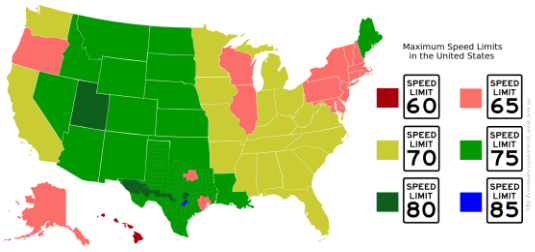
A Potential Outcome!

Do you carefully try to regulate everything with specific protocols?



Truly **a protocol with specificity**…. we even have **an auditor**, the local chief of police!

A variety of approaches exist for how protocols and thresholds are set…

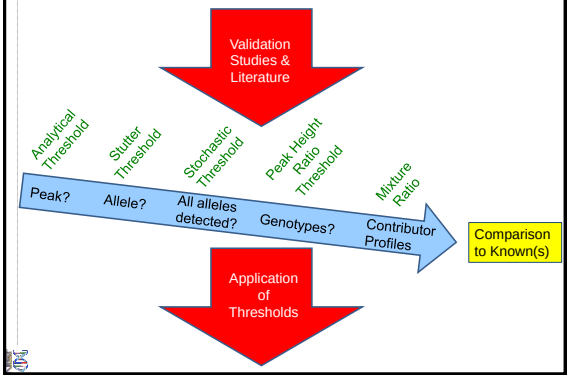


http://en.wikipedia.org/wiki/Speed_limits_in_the_United_States

Threshold Decisions

Thresholds to Determine	Decisions to Make (lab & kit specific)	Useful Validation Data
Analytical = ____ RFU	Single overall value or color specific	Noise levels in negative controls or non-peak areas of positive controls
Stochastic = ____ RFU	Minimum peak height RFU value or alternative criteria such as quantitation values or use of a probabilitistic genotype approach	Level where dropout occurs in low level single-source heterozygous samples under conditions used (e.g., different injection times, post-PCR cleanup)
Stutter filter = ___%	Profile, locus, or allele-specific	Stutter in single-source samples (helpful if examined at multiple DNA quantities)
Peak Height Ratio = ___%	Profile, locus, or signal height (quantity) specific	Heterozygote peak height ratios in single-source samples (helpful if examined at multiple DNA quantities)
Major/Minor Ratio = ____	When will you attempt to separate components of a mixture into major and minor contributors for profile deductions?	Defined mixture ratios (e.g., 1:1, 1:3, 1:9) with known samples to observe consistency across loci and to assess ability to deduce correct contributor profiles

Steps in DNA Interpretation



How Speed Limits Are Set?

http://www.crab.wa.gov/LibraryData/REPORTS/EngineerAnswers/Article03-04SpeedLimits.pdf

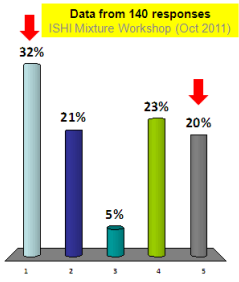
The posted speed limit for a road is set in slightly different ways in different counties. The most common way though, is to use **the "85th percentile" speed**. 85 out of 100 drivers will choose this speed no matter what the signs say. Many studies have shown this method to be safe, practical and enforceable. It also doesn't depend on the opinion of one person.

The 85th percentile speed is easily determined with special traffic counters that check the traffic on the roadway. The speed limit can then be set at the next lower 5 miles per hour. For example, if the traffic counters show 38 mph, the limit would be set at 35 mph. The speed limit may be set another 5 mph lower if there are features not obvious to the driver. These may include unusual roadside or traffic conditions including a high number of accidents.

2011 Response from ISHI Workshop

If your laboratory uses a stochastic threshold (ST), it is:

1. Same value as our analytical threshold (**we don't use a ST**)
2. About twice as high as our AT (e.g., AT = 50 and ST = 100 RFU)
3. Less than twice as high as our AT
4. Greater than twice as high as our AT
5. I don't know!

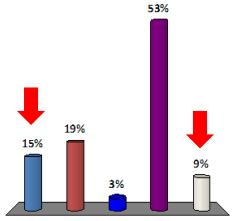


2012 Response from ISHI Workshop

If your laboratory uses a stochastic threshold (ST), it is:

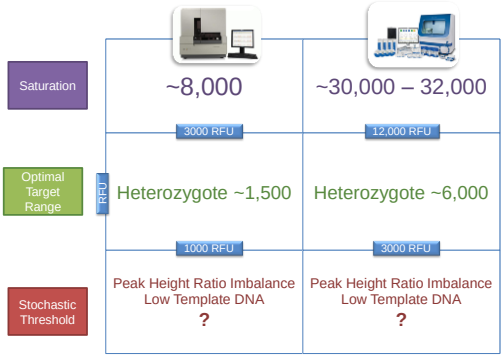
1. Same value as our analytical threshold (**we don't use a ST**)
2. About twice as high as our AT (e.g., AT = 50 and ST = 100 RFU)
3. Less than twice as high as our AT
4. Greater than twice as high as our AT
5. I don't know!

Data from 120 responses
ISHI Mixture Workshop (Oct 2012)



A Few Slides Were Kindly Provided by the Life Technologies/Applied Biosystems Validation Group Showing Data Variation between ABI 3130xl and ABI 3500

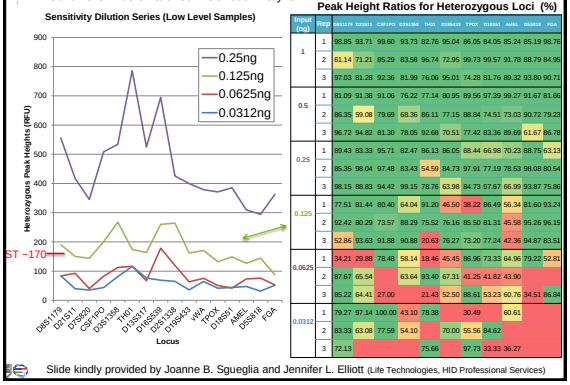
Dynamic Range of 3130*xl* vs. 3500 Genetic Analyzer



Slide kindly provided by Joanne B. Sgueglia and Jennifer L. Elliott (Life Technologies, HID Professional Services)

Stochastic Threshold Considerations

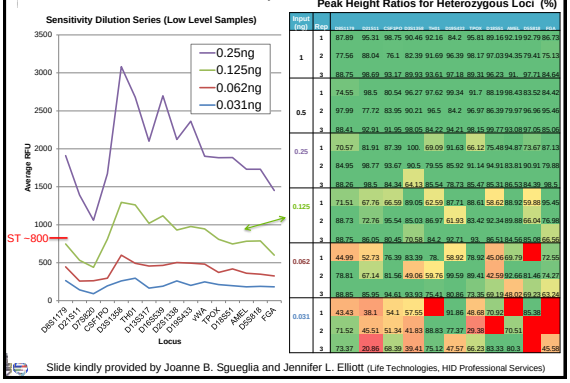
Identifiler® Plus on a 3130*xl* Genetic Analyzer



Slide kindly provided by Joanne B. Sgueglia and Jennifer L. Elliott (Life Technologies, HID Professional Services)

Stochastic Threshold Considerations

Identifiler® Plus on a 3500 Genetic Analyzer

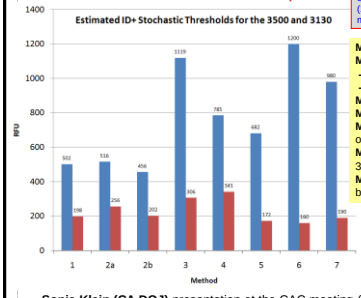


Slide kindly provided by Joanne B. Sgueglia and Jennifer L. Elliott (Life Technologies, HID Professional Services)

Comparison of Different Approaches to Determining a Stochastic Threshold

Results from CA DOJ Identifiler Plus validation experiments

Studied 3 DNA samples with serial dilutions (1, 0.5, 0.25, 0.125, 0.062, 0.031, 0.016 ng), multiple amps of each template quantity



Method 1: tallest false homozygote
Method 2: false homo. ave. +3SD
- **2a:** using most relevant input amount
- **2b:** using all observed false homo.

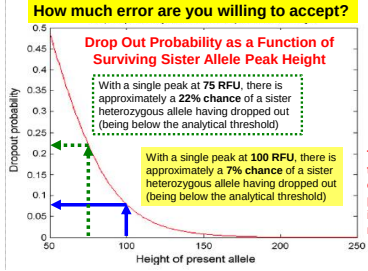
Method 3: average PH het. +3 SD
Method 4: ave. PHR -3 SD vs. signal
Method 5: AT divided by minimum observed PHR
Method 6: partial profile at ~150 pg and 3x AT
Method 7: where majority of PHRs fall below 60%

Blue bars: 3500 ST
Red bars: 3130 ST

Sonja Klein (CA DOJ) presentation at the CAC meeting (Sacramento, CA), October 25, 2011: "Approaches to estimating a stochastic threshold"

Setting a Stochastic Threshold is Essentially Establishing a Risk Assessment

How much error are you willing to accept?



"Currently, most laboratories use an arbitrary stochastic threshold. **When a protocol is changed, especially if it is made more sensitive to low-level DNA, then the stochastic threshold must also change**."
Puch-Solis R, et al. (2011). Practical determination of the low template DNA threshold. *Forensic Sci. Int. Genet.* 5(5): 422-427.

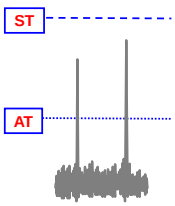
The position and shape of this curve may change based on anything that can impact peak detection (e.g., CE injection time, PCR cycle number, post-PCR cleanup).

Gill, P., et al. (2009). The *low-template* (stochastic) threshold-Its determination relative to risk analysis for national DNA databases. *FSI Genetics*, 3, 104-111.

Limitations of Stochastic Thresholds

- The possibility of allele sharing with a complex mixture containing many contributors may make a stochastic threshold meaningless
- "Enhanced interrogation techniques" to increase sensitivity (e.g., increased PCR cycles) may yield false homozygotes with >1000 RFU
- **New turbo-charged kits with higher sensitivity will need to be carefully evaluated to avoid allele drop-out and false homozygotes**

Can This Locus Be Used for Statistical Calculations?



It depends on your assumption as to the number of contributors!

If you assume a single-source sample, then you can assume that the detection of two alleles fully represents the heterozygous genotype present at this locus.

If you assume (from examining other loci in the profile as a whole) that the sample is a mixture of two or more contributors, then there may be allele drop-out and all alleles may not be fully represented.

Stochastic Threshold Summary

- A stochastic threshold (ST) may be established for a specific set of conditions to reflect possibility of allele drop-out, which is essential for a CPE/CPI stats approach
- ST should be re-examined with different conditions (e.g., higher injection, sample desalting, increase in PCR cycles)
- ST will be dependent on the analytical threshold set with a method and impacts the lowest expected peak height ratio
- Assumptions of the number of contributors is key to correct application of ST

Stats Required for Inclusions

SWGDAM Interpretation Guideline 4.1:

"**The laboratory must perform statistical analysis in support of any inclusion** that is determined to be relevant in the context of a case, irrespective of the number of alleles detected and the quantitative value of the statistical analysis."

Buckleton & Curran (2008): "There is a considerable aura to DNA evidence. Because of this aura **it is vital that weak evidence is correctly represented as weak or not presented at all**."

Buckleton, J. and Curran, J. (2008) A discussion of the merits of random man not excluded and likelihood ratios. *Forensic Sci. Int. Genet.* 2: 343-348.

Coupling of Statistics and Interpretation

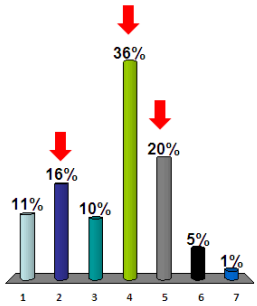
- **The CPE/CPI approach** for reporting an inclusionary statistic **requires that all alleles be observed** in the evidence sample
- If allele drop-out is suspected at a locus, then any allele is possible and the probability of inclusion goes to 100% -- in other words, the locus is effectively dropped from consideration for statistical purposes
- If alleles are seen below the established stochastic threshold, then the locus is typically eliminated ("INC" – declared inconclusive) in many current lab SOPs

2011 Response from ISHI Workshop

What kind of mixture statistic does your lab use?

72% using CPI

1. LR
2. CPE (RMNE, CPI)
3. RMP
4. CPE or RMP
5. Other combinations
6. Probabilistic modeling (e.g., TrueAllele)
7. We don't use stats (contradicting the new guidelines – section 4.1)



Data from 138 responses
ISHI Mixture Workshop (Oct 2011)

CPE/CPI (RMNE) Limitations

- A CPE/CPI approach assumes that ***all alleles*** **are present** (i.e., cannot handle allele drop-out)
- Thus, statistical analysis of low-level DNA CANNOT be correctly performed with a CPE/CPI approach because some alleles may be missing
- Charles Brenner in his AAFS 2011 talk addressed this issue
- Research is on-going to develop allele drop-out models and software to enable appropriate calculations

Notes from Charles Brenner's AAFS 2011 talk

The Mythical "Exclusion" Method for Analyzing DNA Mixtures – Does it Make Any Sense at All?

1. The claim that it requires **no assumption about number of contributors** is mostly wrong.
2. The supposed **ease of understanding** by judge or jury is really an illusion.
3. **Ease of use** is claimed to be an advantage particularly for complicated mixture profiles, those with many peaks of varying heights. The truth is the exact opposite. **The exclusion method is completely invalid for complicated mixtures.**
4. The exclusion method is only **conservative** for guilty suspects.

Conclusion: "Certainly no one has laid out an explicit and rigorous chain of reasoning from first principles to support the exclusion method. It is at best guesswork."

Brenner, C.H. (2011). The mythical "exclusion" method for analyzing DNA mixtures – does it make any sense at all? *Proceedings of the American Academy of Forensic Sciences*, Feb 2011, Volume 17, p. 79

ISFG Recommendations on Mixture Interpretation

http://www.isfg.org/Publication;Gill2006

1. The likelihood ratio (LR) is the preferred statistical method for mixtures over RMNE
2. Scientists should be trained in and use LRs
3. Methods to calculate LRs of mixtures are cited
4. Follow Clayton et al. (1998) guidelines when deducing component genotypes
5. Prosecution determines H_p and defense determines H_d and multiple propositions may be evaluated
6. When minor alleles are the same size as stutters of major alleles, then they are indistinguishable
7. Allele dropout to explain evidence can only be used with low signal data
8. No statistical interpretation should be performed on alleles below threshold
9. Stochastic effects limit usefulness of heterozygote balance and mixture proportion estimates with low level DNA

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

Clayton et al. (1998)
ISFG (2006) Rec. #4

Step #1 Identify the Presence of a Mixture

Step #2 Designate Allele Peaks

Step #3 Identify the Number of Potential Contributors

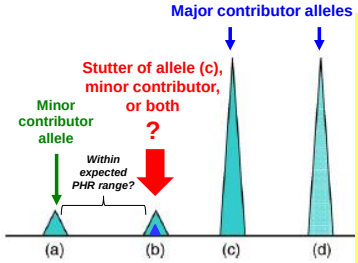
Step #4 Estimate the Relative Ratio of Contributors

Step #5 Consider All Possible Genotype Combinations

Step #6 Compare Reference Samples

ISFG Recommendation #6

Indistinguishable from Stutter



Possibilities for Minor
a,a
a,b
a,c
a,d

Probability of Inclusion = $(f_a + f_b + f_c + f_d)^2$

If peak height of peak a is within established PHR of peak b peak height

Fig. 4. *c* and *d* are unambiguous alleles, *b* is a minor allele in a stutter position and *a* is an unambiguous minor allele.

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

Likelihood Ratio (LR)

- Provides ability to express and evaluate both the prosecution hypothesis, H_p (the suspect is the perpetrator) and the defense hypothesis, H_d (an unknown individual with a matching profile is the perpetrator)

$$LR = \frac{\Pr(E \mid H_p)}{\Pr(E \mid H_d)}$$

- **In the simplest case, the numerator, H_p, is 1** – since in theory the prosecution would only prosecute the suspect if they are 100% certain the suspect is the perpetrator
- The denominator, **H_d**, is typically the profile frequency in a particular population (based on individual allele frequencies and assuming unrelated individuals in Hardy-Weinberg equilibrium) – i.e., **the random match probability**

Take Home Messages

- Inclusionary statements (including "cannot exclude") need statistical support to reflect the relevant weight-of-evidence
- Stochastic thresholds are necessary if using CPI statistics to help identify possible allele dropout
- CPI is only conservative for guilty suspects as this approach does a poor job of excluding the innocent
- Uncertainty exists in scientific measurements and increases with complex mixtures (low level DNA and/or >2 contributors)
- An increasing number of poor samples are being submitted to labs – labs may benefit from developing a complexity threshold

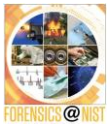
President John F. Kennedy

Yale University commencement address (June 11, 1962)

"For the greatest enemy of truth is very often not the lie – deliberate, contrived and dishonest – but the myth – persistent, persuasive, and unrealistic. Too often we hold fast to the clichés of our forebears. We subject all facts to a prefabricated set of interpretations. **We enjoy the comfort of opinion without the discomfort of thought**."

http://www.jfklibrary.org/Research/Ready-Reference/Kennedy-Library-Miscellaneous-Information/Yale-University-Commencement-Address.aspx

NIST National Institute of Standards and Technology • U.S. Department of Commerce



DNA Mixture Interpretation Webcast
April 12, 2013

http://www.nist.gov/oles/forensics/dna-analyst-training-on-mixture-interpretation.cfm

http://www.cstl.nist.gov/strbase/mixture.htm

Statistical Approaches

Michael D. Coble
National Institute of Standards and Technology

In every workshop presented and supported by the NIJ Training Grant (2008-DN-BX-K158)

- Participants said they needed more training in…
  - – Mixture analysis
  - – **Statistics** related to mixtures

This doesn't have to be a Shakespearean Tragedy!

Stats Required for Inclusions

SWGDAM Interpretation Guideline 4.1:

"**The laboratory must perform statistical analysis in support of any inclusion** that is determined to be relevant in the context of a case, irrespective of the number of alleles detected and the quantitative value of the statistical analysis."

Buckleton & Curran (2008): "There is a considerable aura to DNA evidence. Because of this aura **it is vital that weak evidence is correctly represented as weak or not presented at all**."

Buckleton, J. and Curran, J. (2008) A discussion of the merits of random man not excluded and likelihood ratios. *Forensic Sci. Int. Genet.* 2: 343-348.

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. (1993) 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.

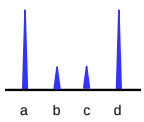
Statistical Approaches with Mixtures

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

"Exclusionary" Approach	"Inferred Genotype" Approach
Random Man Not Excluded (RMNE)	Random Match Probability [modified] (mRMP)
Combined Prob. of Inclusion (CPI)	Likelihood Ratio (LR)
Combined Prob. of Exclusion (CPE)	
"Allele-centric"	"Genotype-centric"

Statistical Approaches with Mixtures

- **Random Man Not Excluded (CPI)** - The probability that a random person (unrelated individual) would not be excluded as a contributor to the observed DNA mixture.

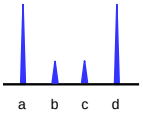


$$PI = (f(a) + f(b) + f(c) + f(d))^2$$

$$CPI = PI_{M1} \times PI_{M2} \cdots$$

$$CPE = 1 - CP1$$

Breaking down the math…



CPI – tries to find all possible "random" persons included in this mixture…

$$(a + b + c + d)^2$$

$$= (a + b + c + d)\,(a + b + c + d)$$

"FOIL"

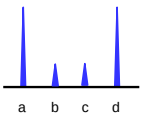
Breaking down the math…

"FOIL"

$$= (a + b + c + d)\,(a + b + c + d)$$

$$= (a^2 + 2ab + 2ac + 2ad + b^2 + \ldots)$$

RMNE Statistics

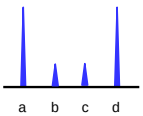


CPI – tries to find all possible "random" persons included in this mixture…

"Included Genotypes"

AA	BB	CC	DD
AB	BC	CD	
AC	BD		
AD			

RMNE Statistics



An "Illogicality" of using RMNE

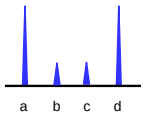
AA + BCD ???

Sure, why not? It fits!

Risk of including individuals ***not*** in the mixture

Statistical Approaches with Mixtures

- **modified Random Match Probability (mRMP)** – The major and minor components can be successfully separated into individual profiles. A random match probability is calculated on the evidence as if the component was from a single source sample.

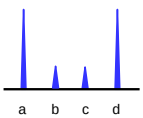


$$mRMP_{minor} = 2pq$$

$$= 2f(b)f(c)$$

Statistical Approaches with Mixtures

- **Likelihood Ratio** - Comparing the probability of observing the mixture data under two (or more) alternative hypotheses; in its simplest form LR = 1/RMP



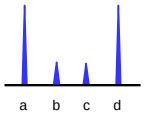
$$\frac{P(E \mid H_1)}{P(E \mid H_2)}$$

E = Evidence
H_1 = Prosecutor's Hypothesis (the suspect did it) = 1
H_2 = Defense Hypothesis (the suspect is an unknown, random person)

$$= \frac{1}{P(E \mid H_2)} = \frac{1}{2pq} = 1/RMP$$

Comparison of the Methods

"Included Genotypes" RMNE



AA BB CC DD
AB BC CD AD
AC BD

"Included Genotypes" LR/mRMP

~~AA BB CC DD~~
~~AB~~ BC ~~CD~~ AD
~~AC~~ ~~BD~~

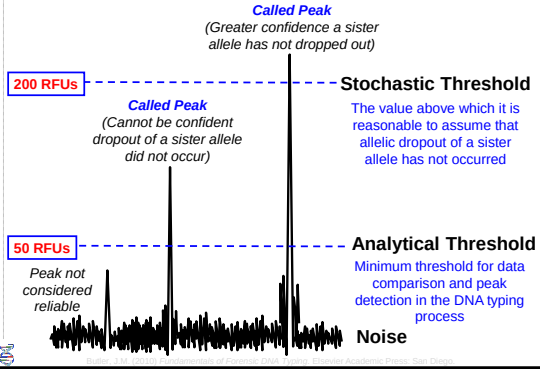
FSI GENETICS
Forensic Science International: Genetics 2 (2008) 343–348

A discussion of the merits of random man not excluded and likelihood ratios

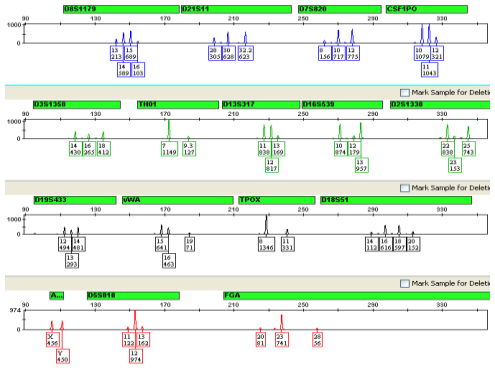
John Buckleton [a,*], James Curran [b]
[a] ESR, PB 92021, Auckland, New Zealand
[b] Department of Statistics, University of Auckland, PB 92019, Auckland, New Zealand
Received 15 January 2008; received in revised form 29 April 2008; accepted 1 May 2008

We conclude that the two matters that appear to have real force are:
(1) LRs are more difficult to present in court and
(2) the RMNE statistic wastes information that should be utilised.

Review of Two Thresholds



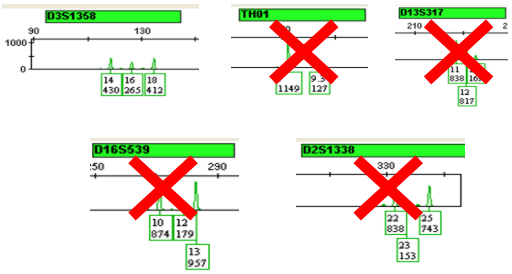
2-Person Mixture



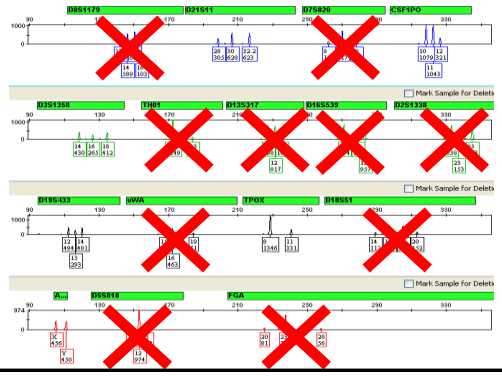
If CPI/CPE Stats are Used

Since exclusionary statistics cannot adjust for the possibility of dropout, and does not take the number of contributors into account, any loci with alleles below the stochastic threshold cannot be used in the CPI statistic.

If CPI/CPE Stats are Used (ST = 200 RFU)



If CPI/CPE Stats are Used



If CPI/CPE Stats are Used

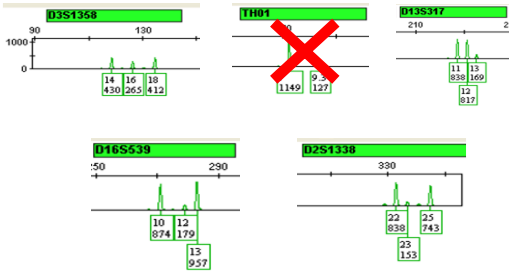
Can use	Cannot use	
D21	D8	D2
CSF	D7	vWA
D3	TH01	D18
D19	D13	D5
TPOX	D16	FGA

Impact: discarding 2/3 of the data

If CPI/CPE Stats are Used

- CPI statistics using FBI Caucasian Frequencies
- 1 in 71 Caucasians included
- 98.59% Caucasians excluded

If CPI/CPE Stats are Used (ST = 150 RFU)

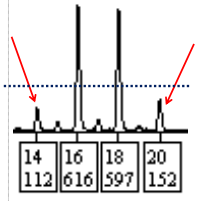


The impact of changing thresholds

If mRMP/LR Stats are Used

- Since there is an assumption to the number of contributors, it is possible to use data that falls below the ST.

mRMP - D18S51



(LR = 113)

If Assume 2 Contributors….

Major	Minor
16,18	14,20

$mRMP_{minor} = 2pq$

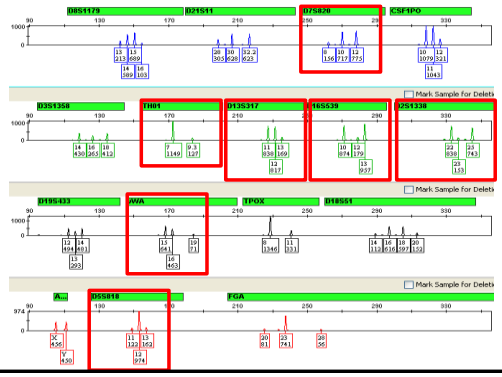
$= 2 \times f(14) \times f(20)$

$= 2 \times (0.1735) \times (0.0255)$

$= 0.00884$ or 1 in 113

mRMP/LR

Potential for Drop-out

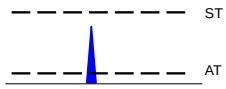


If mRMP/LR Stats are Used

Can use	Loci with potential D-out	
D8	D7	D2
D21	TH01	vWA
D18	D13	D5
D3	D16	
D19		
TPOX		
FGA		
CSF		

The "2p" Rule

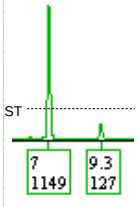
- The "2p" rule can be used to statistically account for zygosity ambiguity – i.e. is this single peak below the stochastic threshold the result of a homozygous genotype or the result of a heterozygous genotype with allele drop-out of the sister allele?



2p – SWGDAM Guidelines

- 5.2.1.3.1. The formula 2p, as described in recommendation 4.1 of NRCII, may be applied to this result.
- 5.2.1.3.2. Instead of using 2p, the algebraically identical formulae **$2p - p^2$** and **$p^2 + 2p(1-p)$** may be used to address this situation without double-counting the proportion of homozygotes in the population.

Macbeth/Duncan Profile - TH01



Major – 7, 7

Possible Minor Contributors

7, 9.3	(2pq)
9.3, 9.3	p^2
9.3, ?	2p (or $p^2 + 2p(1-p)$)

Macbeth/Duncan Profile - TH01

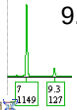
$$\frac{P(E|H_1)}{P(E|H_2)} = \frac{V \,\&\, S}{V \,\&\, U} = \frac{f_7^2 + f_7(1-f_7)\,\theta \;\&\; 1}{f_7^2 + f_7(1-f_7)\,\theta \;\&\; 2p}$$

$p^2 + 2p(1-p)$

V = 7, 7

U = 7, 9.3
9.3, 9.3
9.3, ?

$$= \frac{1}{f_{9.3}^2 + 2f_{9.3}(1-f_{9.3})}$$



$f_{9.3} = 0.3054$

$= 1 / 0.5175 = \boxed{1.93}$

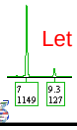
Macbeth/Duncan Profile - TH01

$$\frac{P(E|H_1)}{P(E|H_2)} = \frac{V \& S}{V \& U} = \frac{1}{p^2 + p(1-p)\theta + 2pq}$$

V = 7, 7
U = 7, 9.3
9.3, 9.3

$$= \frac{1}{f_{9.3}^2 + f_{9.3}(1-f_{9.3})\theta + 2f_{9.3}f_7}$$

Let ST = 125 RFU



$f_{9.3}$ = 0.3054
f_7 = 0.1724

= 1 / 0.2007 = **4.98**

Macbeth/Duncan Profile - TH01

	LR
ST = 200 (2p is used)	1.93
ST = 125 (2pq is used)	4.98

2p is conservative…

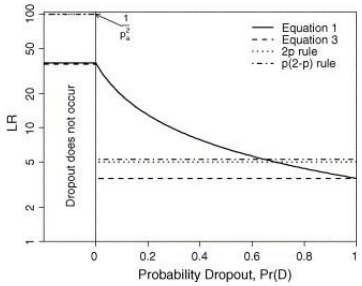
The "2p" Rule

- "This rule arose during the VNTR era. At that time many smaller alleles "ran off the end of the gel" and were not visualised."

- Buckleton and Triggs (2006)

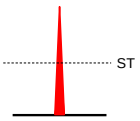
Is the 2*p* rule always conservative?"

The "2p" Rule



Stain = aa

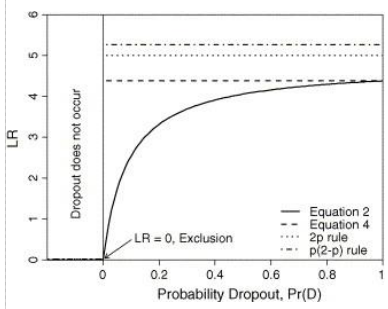
Suspect = aa



LR = 100

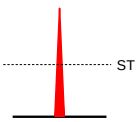
f(a) = 0.10 $1/p^2$ = 100 1/2p = 5

The "2p" Rule



Stain = aa

Suspect = ab



Exclusion

f(a) = 0.10 1/2p = 5

Is there a way forward?

Gill and Buckleton *JFS*
55: 265-268 (2010)

- "The purpose of the ISFG DNA commission document was to provide a way forward to demonstrate the use of ***probabilistic models to circumvent the requirement for a threshold*** and to safeguard the legitimate interests of defendants."

Summary of the Issues

- We need to move away from the interpretation of mixtures from an "allele-centric" point of view.
- Methods to incorporate probability will be necessary as we make this transition and confront the issues of low-level profiles with drop-out.
- "Just as logic is reasoning applied to truth and falsity, probability is reasoning with uncertainty"

-Dennis Lindley

Summary of the Issues

- The LR is a method to evaluate evidence that can overcome many of the limitations we are facing today. ISFG Recommendations are published.
- This will require (obviously) software solutions… however, we need to better understand and be able to explain the statistics as a community.
- "But, for my own part, it was Greek to me"
  — William Shakespeare, *Julius Caesar*
- "We know what we are, but know not what we may be." — William Shakespeare, *Hamlet*

Summary of the Issues

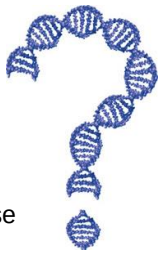
- Extensive training will be necessary – and a single 8 hour workshop will once a year will not suffice.

Thank you for your attention

Contact Information

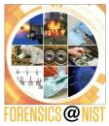
Michael D. Coble
Forensic Biologist
michael.coble@nist.gov
301-975-4330

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



Additional DNA mixture information available at:
http://www.cstl.nist.gov/strbase/mixture.htm

NIST National Institute of Standards and Technology • U.S. Department of Commerce



DNA Mixture Interpretation Webcast
April 12, 2013

http://www.nist.gov/oles/forensics/dna-analyst-training-on-mixture-interpretation.cfm

http://www.cstl.nist.gov/strbase/mixture.htm

Worked Examples & Report Wording

Bruce J. Heidebrecht
Maryland State Police,
Forensic Sciences Division

Process to mixture analysis

1) Look at overall e-gram to make assumptions of number of contributors, ratio of contributors, and if the mixture fits the lab's criteria for major/minor determinations.

2) Identify which alleles are below the stochastic threshold and therefore might have dropout at that locus.

3) For loci without unambiguous minor alleles, determine if minor contributor is reasonable to be considered masked by major, or might be dropping out completely.

4) Analyze mixture for peaks that are "indistinguishable from stutter." ("IFS")

Process to mixture analysis

- At this point, the analysis of the sample may be complete, dependent upon choice of statistics.
  - At this point, all loci should be identified as being useable for major/minor contributor(s) or CPE/CPI statistics.
  - All of this is done independently of the reference standards.
  - The application of which loci are useful for statistics utilizing assumptions (e.g. LR, RMP, and mixture deconvolution) may be influenced by the reference standard of the "known contributor."

Process to mixture analysis

5) Compare any reference standards that are to be considered "known" to the mixture (e.g. victim on own vaginal swab).

6) If doing stats involving a "known" contributor, re-evaluate non-known contribution to mixture for possible dropout and "indistinguishable from stutter".

7) NOTE: this re-evaluation is done without consideration of the probative reference standard.

Process to mixture analysis

8) Compare any reference standards that are to be considered probative to the mixture (e.g. suspect on victim's vaginal swab). If the probative reference standard is excluded from the mixture, declare an exclusion.

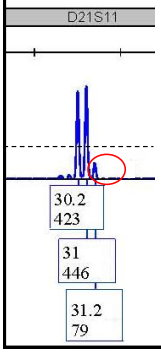
9) If the probative reference standard is not excluded from the mixture, determine the weight of that statement using statistics.

Process to mixture analysis

10) If statistics cannot be applied to support a statement of non-exclusion, then the probative reference standard can not be included, but *might* be able to be excluded, as a potential contributor to the mixture.

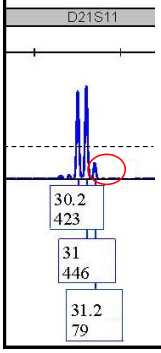
If can not exclude, but can not statistically support an inclusion, the association of the individual to the evidence is inconclusive.

2 person mixture, data below the stochastic threshold, reasonable to assume dropout



- Minor contributor has one detected allele (31.2) below stochastic threshold.
- Reasonable to assume sister allele to the 31.2 may be below the analytical threshold.
- Major alleles 30.2,31 are well balanced (95%PHR) … no indication that a sister allele to the 31.2 must be masked by major contributor.
- Include / exclude to the major based upon genotype 30.2,31 Easy, no need to discuss…
- Include / exclude to the minor based upon a requisite allele 31.2

2 person mixture, data below the stochastic threshold, reasonable to assume dropout



- RMP to probative minor contributor:

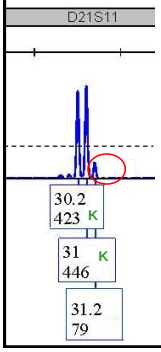
  $2P_{(31.2)}$
- Conclusion statements:

  DNA from two contributors was obtained from the evidence.

  John Q. Suspect cannot be excluded as the minor contributor of this mixture.

  The probability of selecting an unrelated individual at random who cannot be excluded as the minor contributor to the DNA profile obtained from this item is approximately: 1 in 5

2 person mixture, data below the stochastic threshold, reasonable to assume dropout



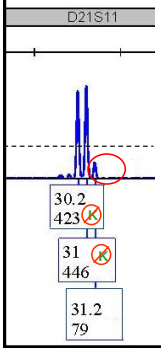
- "Known contributor" = major contributor.
- Likelihood Ratio for probative minor contributor:

  $1 / 2P_{(31.2)}$
- Conclusion statements:

  DNA from two individuals was obtained from the evidence.

  Assuming the presence of Jane K. Victim, the DNA profile is approximately 5 times more likely to occur if it originated from Jane K. Victim and John Q. Suspect than from Jane K. Victim and an unknown individual in the Caucasian population.

2 person mixture, data below the stochastic threshold, reasonable to assume dropout

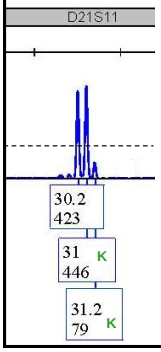


- CPI:

  Cannot perform CPI stats on the minor component due to data below the stochastic threshold.
- Conclusion statements:

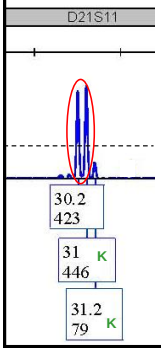
  The minor component of the DNA profile obtained from this item does not satisfy the laboratory's inclusionary reporting criteria.

2 person mixture, data below the stochastic threshold, reasonable to assume dropout?



- What if minor contributor is "known"?
- Victim's fingernail scrapings. Victim's profile is 31,31.2.
- Dropout of the minor contributor is not happening. Victim's DNA is just at low levels.
- This decision is made based upon knowledge of the known contributor's profile in comparison to the mixture.
- This decision is <u>not</u> made based upon any probative reference profile.

2 person mixture, data below the stochastic threshold, unreasonable to assume dropout



- RMP to probative major contributor:

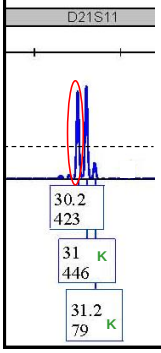
  $2P_{(30.2)}\,P_{(31)}$
- Conclusion statements:

  DNA from two contributors was obtained from the evidence.

  John Q. Suspect cannot be excluded as the major contributor of this mixture.

  The probability of selecting an unrelated individual at random who cannot be excluded as the major contributor to the DNA profile obtained from this item is approximately: 1 in 180

2 person mixture, data below the stochastic threshold, unreasonable to assume dropout



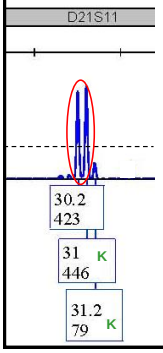
- Unrestricted Likelihood Ratio:

  $1 / \{ [2P_{(30.2)} P_{(31)}] + [2P_{(30.2)} P_{(31.2)}] + [P_{(30.2)}]^2 \}$
- Conclusion statements:

  DNA from two individuals was obtained from the evidence.

  Assuming the presence of Jane K. Victim, the DNA profile is approximately 68 times more likely to occur if it originated from Jane K. Victim and John Q. Suspect than from Jane K. Victim and an unknown individual in the Caucasian population.

2 person mixture, data below the stochastic threshold, unreasonable to assume dropout



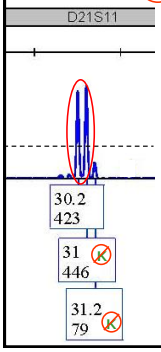
- Restricted Likelihood Ratio:

  $1 / \{2P_{(30.2)} P_{(31)}\}$
- Conclusion statements:

  DNA from two individuals was obtained from the evidence.

  Assuming the presence of Jane K. Victim, the DNA profile is approximately 180 times more likely to occur if it originated from Jane K. Victim and John Q. Suspect than from Jane K. Victim and an unknown individual in the Caucasian population.

2 person mixture, data below the stochastic threshold, unreasonable to assume dropout

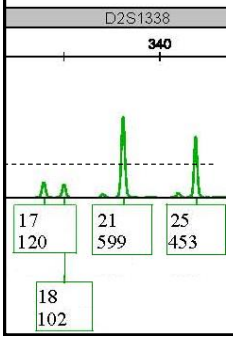


- CPI:

  Cannot perform on locus as a whole due to data below the stochastic threshold.
- Conclusion statements:

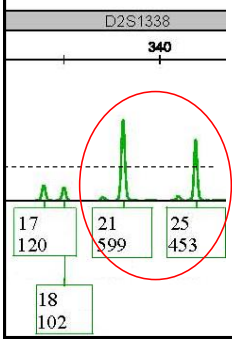
  The DNA profile obtained from this item does not satisfy the laboratory's inclusionary reporting criteria.

2 person mixture, data below the stochastic threshold, unreasonable to assume dropout



- Locus has two detected alleles (17,18) below stochastic threshold.
- Since four alleles detected in a mixture reasoned to be only two contributors, it is unreasonable to assume dropout is occurring.

2 person mixture, data below the stochastic threshold, unreasonable to assume dropout



- RMP to probative major contributor:

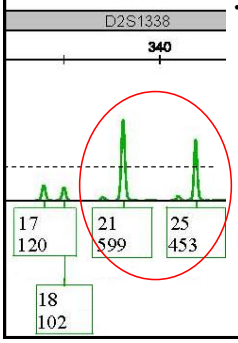
  $2P_{(21)}P_{(25)}$
- Conclusion statements:

  DNA from two contributors was obtained from the evidence.

  John Q. Suspect cannot be excluded as the major contributor of this mixture.

  The probability of selecting an unrelated individual at random who cannot be excluded as the major contributor to the DNA profile obtained from this item is approximately: 1 in 260

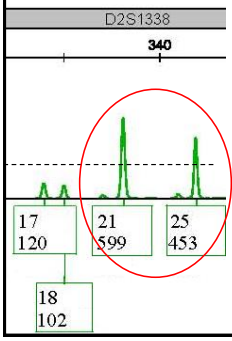
2 person mixture, data below the stochastic threshold, unreasonable to assume dropout



- Unrestricted Likelihood Ratio for major contributor:

$$\frac{\{2P_{(17)}P_{(18)}\}}{[\{2P_{(17)}P_{(18)}\} * \{2P_{(21)}P_{(25)}\}] + [\{2P_{(17)}P_{(21)}\} * \{2P_{(18)}P_{(25)}\}] + [\{2P_{(17)}P_{(25)}\} * \{2P_{(18)}P_{(21)}\}] +}$$

2 person mixture, data below the stochastic threshold, unreasonable to assume dropout

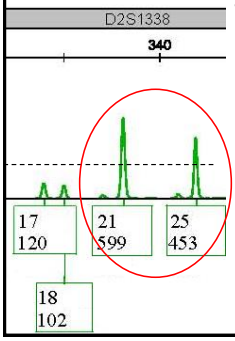


- Unrestricted Likelihood Ratio for major contributor:
- Conclusion statements:

DNA from two individuals was obtained from the evidence.

The DNA profile is approximately 44 times more likely to occur if it originated from John Q. Suspect and an unknown individual in the Caucasian population than from two unknown individuals in the Caucasian population.

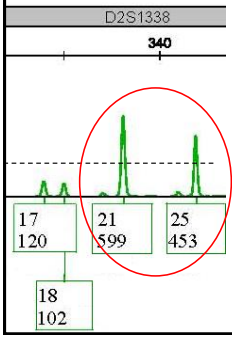
2 person mixture, data below the stochastic threshold, unreasonable to assume dropout



- Restricted Likelihood Ratio for major contributor:

$$\frac{\{2P_{(17)}P_{(18)}\}}{[\,\{2P_{(17)}P_{(18)}\} * \{2P_{(21)}P_{(25)}\}\,]}$$

2 person mixture, data below the stochastic threshold, unreasonable to assume dropout

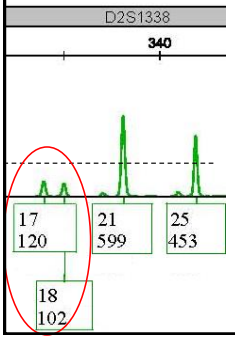


- Restricted Likelihood Ratio for major contributor:
- Conclusion statements:

DNA from two individuals was obtained from the evidence.

The DNA profile is approximately 260 times more likely to occur if it originated from John Q. Suspect and an unknown individual in the Caucasian population than from two unknown individuals in the Caucasian population.

2 person mixture, data below the stochastic threshold, unreasonable to assume dropout



- RMP to probative minor contributor:

$$2P_{(17)}P_{(18)}$$

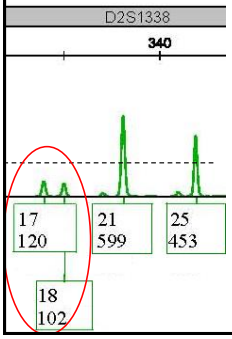
- Conclusion statements:

DNA from two contributors was obtained from the evidence.

John Q. Suspect cannot be excluded as the minor contributor of this mixture.

The probability of selecting an unrelated individual at random who cannot be excluded as the minor contributor to the DNA profile obtained from this item is approximately: 1 in 48

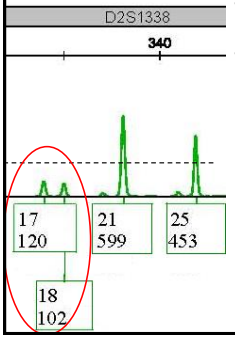
2 person mixture, data below the stochastic threshold, unreasonable to assume dropout



- Unrestricted Likelihood Ratio for minor contributor:

$$\frac{\{2P_{(21)}P_{(25)}\}}{\begin{array}{l}[\,\{2P_{(17)}P_{(18)}\} * \{2P_{(21)}P_{(25)}\}\,] + \\ [\,\{2P_{(17)}P_{(21)}\} * \{2P_{(18)}P_{(25)}\}\,] + \\ [\,\{2P_{(17)}P_{(25)}\} * \{2P_{(18)}P_{(21)}\}\,] +\end{array}}$$

2 person mixture, data below the stochastic threshold, unreasonable to assume dropout

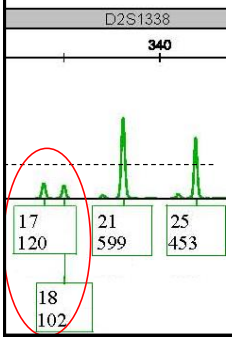


- Unrestricted Likelihood Ratio for minor contributor:
- Conclusion statements:

DNA from two individuals was obtained from the evidence.

The DNA profile is approximately 8 times more likely to occur if it originated from John Q. Suspect and an unknown individual in the Caucasian population than from two unknown individuals in the Caucasian population.

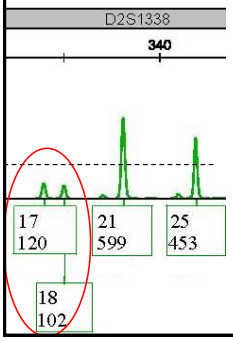
2 person mixture, data below the stochastic threshold, unreasonable to assume dropout



- Restricted Likelihood Ratio for minor contributor:

$$\frac{\{2P_{(21)}P_{(25)}\}}{[\{2P_{(17)}P_{(18)}\} * \{2P_{(21)}P_{(25)}\}]}$$

2 person mixture, data below the stochastic threshold, unreasonable to assume dropout

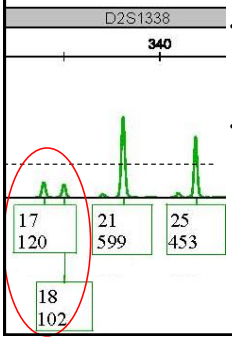


- Restricted Likelihood Ratio for minor contributor:
- Conclusion statements:

  DNA from two individuals was obtained from the evidence.

  The DNA profile is approximately 48 times more likely to occur if it originated from John Q. Suspect and an unknown individual in the Caucasian population than from two unknown individuals in the Caucasian population.

2 person mixture, data below the stochastic threshold, unreasonable to assume dropout

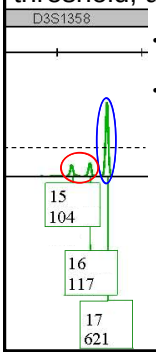


- CPI:

  Cannot perform on locus as a whole due to data below the stochastic threshold.
- Conclusion statements:

  The DNA profile obtained from this item does not satisfy the laboratory's inclusionary reporting criteria.

2 person mixture, data below the stochastic threshold, unreasonable to assume dropout



- Locus has two detected alleles (15,16) below stochastic threshold.
- Although less than four alleles detected in a mixture reasoned to be only two contributors, it is unreasonable to assume dropout is occurring based upon examination for potential genotypes.

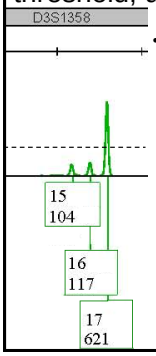
  15,F and 16,17 = unreasonable (19%PHR)

  16,F and 15,17 = unreasonable (17%PHR)

  15,16 and 17,F = unreasonable (allele 17 above stochastic threshold)

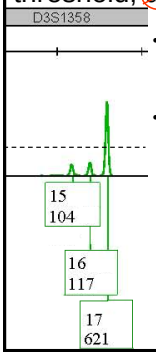
  15,16 and 17,17 = reasonable (89%PHR)

2 person mixture, data below the stochastic threshold, unreasonable to assume dropout



- Since locus has been reasoned to have no dropout and major/minor genotypes have been reasoned, can perform:
  - RMP for major
  - RMP for minor
  - ULR for major
  - ULR for minor
  - RLR for major
  - RLR for minor

2 person mixture, data below the stochastic threshold, unreasonable to assume dropout

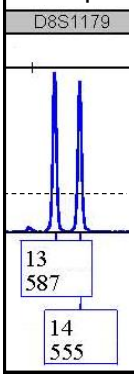


- CPI:

  Cannot perform on locus as a whole due to data below the stochastic threshold.
- Conclusion statements:

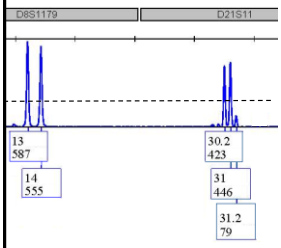
  The DNA profile obtained from this item does not satisfy the laboratory's inclusionary reporting criteria.

2 person mixture, possible dropout?



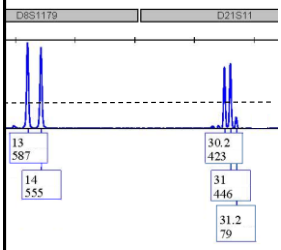
- Ratio of contributors v. stochastic threshold:
  - Is this 1:1? Dropout is unreasonable
  - Is this 10:1? Dropout is reasonable
- Overall height of minor contributor:
  - Minor contributor heterozygous alleles are above stochastic threshold? Dropout is unreasonable
  - Minor contributor heterozygous alleles are below stochastic threshold? Dropout is reasonable
- Molecular weight of locus:
  - Minor contributor alleles are seen in higher molecular weight loci? Dropout is *less* reasonable
  - Minor contributor alleles are not seen in higher molecular weight loci? Dropout is *more* reasonable

2 person mixture, reasonable to assume dropout at D8?



- Ratio of contributors:
  - D21 is ~ 5:1 or 10:1.
    Dropout at D8 is reasonable.
- Overall height of minor contributor:
  - Minor contributor allele at D21 is below stochastic threshold.
    Dropout at D8 is reasonable.
- Molecular weight of locus:
  - Minor contributor allele is seen in D21.
    Dropout is less reasonable.

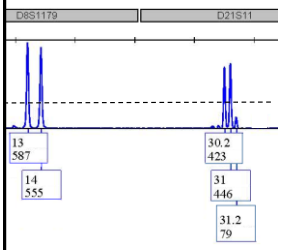
2 person mixture, reasonable to assume dropout at D8?



- Make a decision before comparing the profile of the probative reference standard.
- If declaring possible dropout at D8, then the true minor contributor could be any profile.
- This renders the locus useless for statistics for the minor.

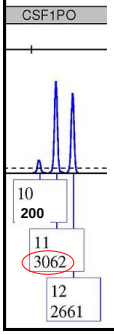
• Even if the probative reference standard is fully represented by the detected alleles!

2 person mixture, reasonable to assume dropout at D8?



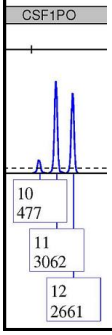
- Make a decision before comparing the profile of the probative reference standard.
- If declaring dropout is unreasonable at D8, then the true minor contributor must be masked.
- This renders the locus useful for statistics for the minor.
- However, if the probative reference standard is not masked by the detected alleles, then exclude!

Minor allele in stutter position (consider stutter percentage)



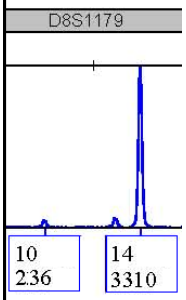
- Major alleles 11,12 are well balanced (87%PHR) … no indication that a sister allele to the 10 must be masked by major contributor.
- 200rfu stochastic threshold.
- 477-277 = 200.
- If 277rfu in bin 10 is stutter, the true value of allele 10 may be below stochastic threshold; sister allele to the 10 may be undetected.
- 277 / 3062 = 9%
- Is 9% stutter reasonable?

Minor allele in stutter position (consider stutter percentage)



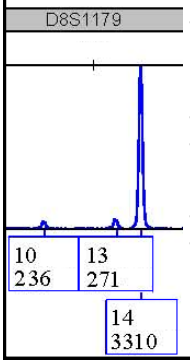
- Even though all data is above stochastic threshold, a thorough interpretation may show that dropout is still reasonable.
- Even if the probative reference standard is fully represented within the detected alleles, the true minor contributor may have an undetected allele.
- Stats for the minor contributor:
  RMP using a "2P" calculation
  ULR using a "2P" calculation
  RLR using a "2P" calculation
  CPI is not appropriate.

Minor allele indistinguishable from stutter



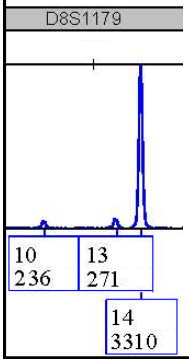
- Since unambiguous minor (10) is above stochastic threshold, and not in stutter position, unreasonable to assume dropout of a sister allele to the allele 10.
- However, the sister allele to the allele 10 may be in the stutter bin 13 and have been filtered out by the software.
- Reanalyze the mixture with stutter filters set to 0%.

Minor allele indistinguishable from stutter



- Upon examining all data above analytical, without regards to stutter filters, compare peaks in stutter positions to unambiguous minor alleles.
- 271 / 236 = 115% PHR
- Even with some amount of stutter present in bin 13, this peak may contain a true sister allele to the allele 13.
- However, it could also be only stutter.
- As such, it is "indistinguishable from stutter" (IFS).

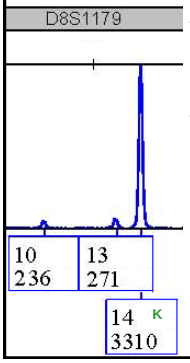
Minor allele indistinguishable from stutter



- Statistics for the minor contributor (or mixture as a whole) must incorporate both ideas of the peak being stutter and being a true minor allele.
- RMP for minor:
  $(P_{(10)}^2) + (2P_{(10)}P_{(14)}) + (2P_{(10)}P_{(13)})$

The probability of selecting an unrelated individual at random who cannot be excluded as the minor contributor to the DNA profile obtained from this item is approximately: 1 in 8

Minor allele indistinguishable from stutter

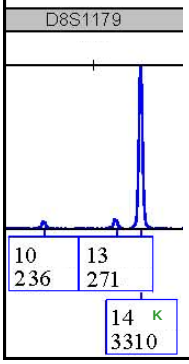


- Statistics for the minor contributor (or mixture as a whole) must incorporate both ideas of the peak being stutter and being a true minor allele.
- ULR for minor (considering major is "known")

$$\frac{1}{P_{(10)}^2 + 2P_{(10)}P_{(14)} + 2P_{(10)}P_{(13)}}$$

The DNA profile is approximately 8 times more likely to occur if it originated from Jane K. Victim and John Q. Suspect than from Jane K. Victim and an unknown individual in the Caucasian population.

Minor allele indistinguishable from stutter

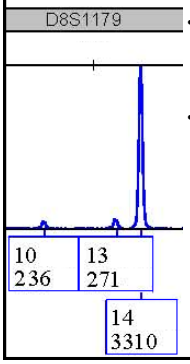


- Statistics for the minor contributor (or mixture as a whole) must incorporate both ideas of the peak being stutter and being a true minor allele.
- RLR for minor (considering major is "known")

$$\frac{1}{P_{(10)}^2 + 2P_{(10)}P_{(14)} + 2P_{(10)}P_{(13)}}$$

The DNA profile is approximately 8 times more likely to occur if it originated from Jane K. Victim and John Q. Suspect than from Jane K. Victim and an unknown individual in the Caucasian population.

Minor allele indistinguishable from stutter



- Statistics for the minor contributor (or mixture as a whole) must incorporate both ideas of the peak being stutter and being a true minor allele.
- CPI for mixture as a whole:
  $\{P_{(10)} + P_{(13)} + P_{(14)}\}^2$

The probabilities of selecting an unrelated individual at random who cannot be excluded as one of the possible sources of the DNA profile obtained from this item are approximately 1 in 2 in the Caucasian population.

Minor peak distinguishable as stutter

- Not every peak in every stutter bin is worthy of being designated as IFS.
- If the mixture has no distinction of major and minor, then there is no minor contributor at the rfu level of stutter peaks.
- If a locus has already been declared to have the possibility of dropout, the statistics that incorporate dropout account for IFS peaks. "Dropout trumps IFS."
- If the minor contributor already has a complete genotype defined by the unambiguous alleles.
- If the minor contributor is "known" and that genotype is already defined by the unambiguous alleles.

Documentation

- Documentation of the interpretation within the case folder is crucial:
  - The technical reviewer can understand why the analyst made certain decisions.
  - The analyst can refer to the case notes in court to recall the decisions.
  - The analysis is open to the scrutiny of another expert.

Documentation

- Documentation of the interpretation within the case folder is crucial:
  - Analytical and stochastic thresholds.
  - Number of contributors hypothesized to be present.
  - Presence of any "known" contributors.
  - Reasons to discount dropout when data is present below the stochastic threshold.
  - Reasons to include possible dropout when no data is visible below the stochastic threshold.
  - Reasons to identify a peak as stutter or "indistinguishable from stutter".

Documentation

- Documentation of the assumptions (number of contributors, presence of "known" contributor, etc.) within the case report is crucial:
  - Who may see only the report and never see the case notes?
    - Law enforcement
    - Prosecuting attorney
    - Defense attorney
    - Judge
    - Jury

NIST National Institute of Standards and Technology • U.S. Department of Commerce

DNA Mixture Interpretation Webcast
April 12, 2013

http://www.nist.gov/oles/forensics/dna-analyst-training-on-mixture-interpretation.cfm

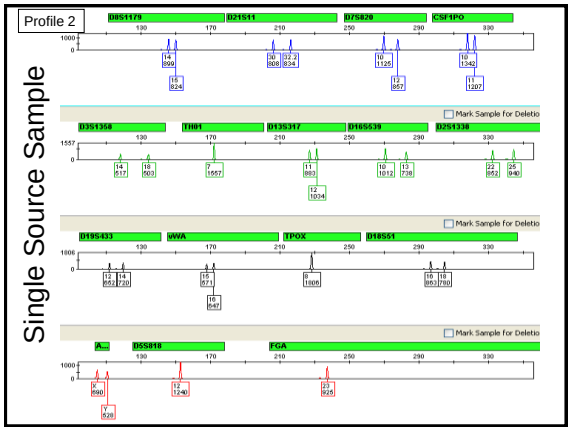
http://www.cstl.nist.gov/strbase/mixture.htm

Different Assumptions & Different Interpretations

Charlotte J. Word
Consultant

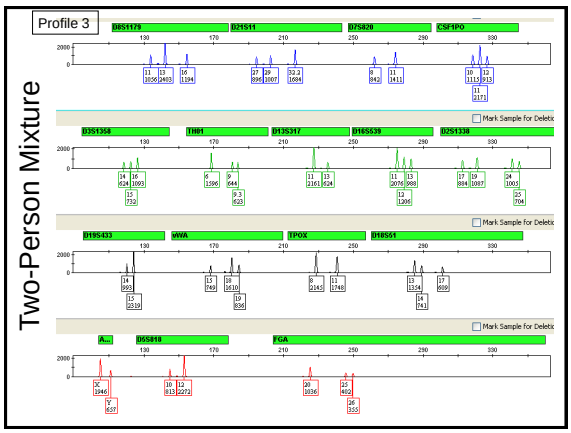
MYTH

No assumptions are needed for interpreting DNA profiles from good quality single source samples.



Assumptions Made
Single Source

- Peaks above the analytical threshold are alleles from the contributor
  - Stutter peaks, other peaks are assumed to be artifacts and can be ignored
- All alleles from the contributor are present since all peaks are above the stochastic threshold
- There is a single DNA contributor
  - No more than two alleles at any locus
  - Genotypes are easy to assume
    - Balanced peak heights where heterozygous
    - Double peak height where homozygous



Assumptions Made
Two Person Mixture

- Peaks above the analytical threshold are alleles from the contributors
  - Stutter peaks, other peaks are assumed to be artifacts and can be ignored
- All alleles from the contributors are present since all peaks are above the stochastic threshold
- There are (only) two DNA contributors
  - No more than four alleles at any locus
  - Data consistent with mixture validation studies and experience with two person mixtures

Assumptions Made
Two Person Mixture

- Genotypes may be easily assumed
  - If have major:minor scenario, can use mixture ratio and peak height ratios to associate alleles into genotypes and associate genotypes into complete profiles
  - Can assume one known is a contributor and deduce the second contributor
  - If have indistinguishable mixture, can assume a limited number of possible genotypes and genotype combinations at each locus: (e.g., alleles 13,14,15,16 = genotypes of 13,14 + 15,16 or 13,15 + 14,16 or 13,16 + 14,15)

Assumptions

- Assumptions are made with all data analyses and with all interpretations of data
- We may not always clearly state those assumptions or even be aware that we are making those assumptions
- We may not always report those assumptions

But we MUST be aware of what assumptions we are making

MYTH

No assumptions are needed for interpreting DNA profiles from good quality single source samples.

Assumptions

- We have a lot of familiarity and experience making reasonable assumptions for high quality single source and two person mixtures
- High quality profile leads to high confidence in data and high certainty regarding interpretations and conclusions

But what about REAL Casework Profiles?!

REAL Casework

Situations with increased uncertainty, and therefore decreased confidence:

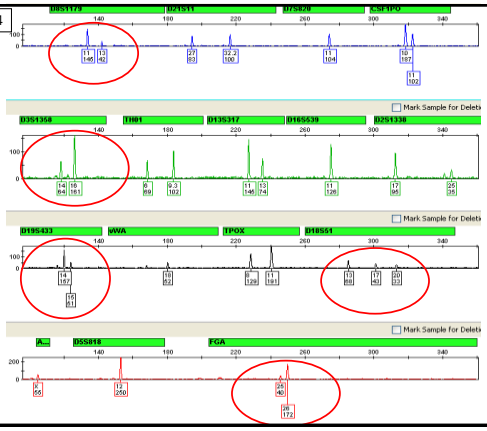
- Alleles vs. artifacts? (LT or high level DNA)
- Stochastic effects possible? (Low peak heights; all or some below stochastic threshold)
  - Sure all alleles are present (drop-out)?
  - Elevated stutter & drop-in present?
- Number of contributors? 1, 2, 3 or more?
- Inability to associate all alleles into reasonable genotypes with high confidence
- Degradation?

MYTH

It may be useful to consider some DNA profiles under different assumptions.

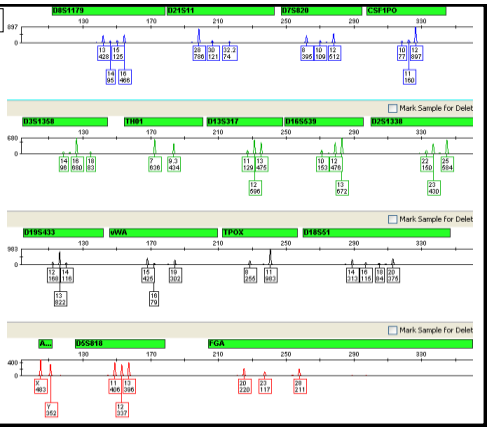
Profile 4

Single Source vs. Mixture?



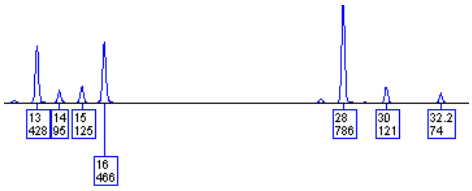
Profile 5

All Alleles vs. Allelic Drop-out?
Two or More Contributors?



Profile 5

Is Known Individual Included or Excluded?

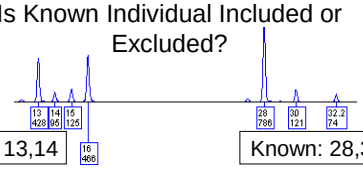


Known: 13,14

Known: 28,30

Profile 5

Is Known Individual Included or Excluded?



Known: 13,14

Known: 28,30

Assumptions:

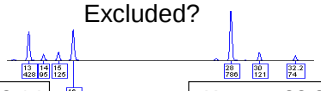
1) 2 contributors *and* all data are present →
2) 1 major and 1 minor contributor →
3) Major must have 13,16 and 28,28 genotypes and
4) Minor must have 14,15 and 30,32.2 genotypes

Based on these assumptions, the individual is excluded

Genotype is excluded even if alleles are included

Profile 5

Is Known Individual Included or Excluded?



Known: 13,14

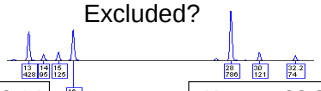
Known: 28,30

New Assumption:

1) 3 contributors *and* possible LT DNA →
2) 1 major and 2 minor contributors →
3) Major must have genotype of 13,16 and 28,28
4) One or other or both minor contributors have 14 and/or 15 and 30 and/or 32.2, but cannot associate alleles to genotypes
5) Possible genotype list is long due to stochastic effects

Profile 5

Is Known Individual Included or Excluded?



Known: 13,14

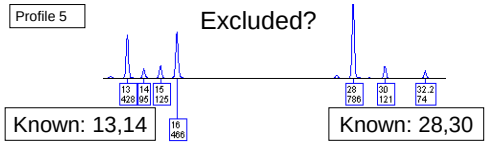
Known: 28,30

Other Possible Assumptions:

1) 3 or more contributors *and* possible LT DNA→
2) 1 major and 3 or more minor contributors → OR
3) 2 major contributors and 1 minor contributor → OR
4) 2 majors and 2 or more minor contributors →
5) Decreased ability to associate alleles to genotypes
6) Possible genotype list is long

Is Known Individual Included or Excluded?

Profile 5



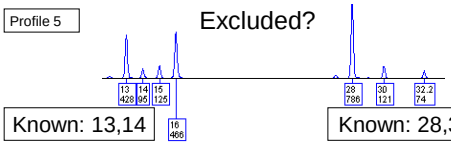
Known: 13,14 Known: 28,30

Based on the assumption of 3 or more contributors, there is insufficient information to exclude known genotypes. What do you report?

Inclusion – but statistics MUST take into account possible stochastic effects (may not be meaningful)
Inconclusive – but throwing away possibly exculpatory or inculpatory data

Is Known Individual Included or Excluded?

Profile 5



Known: 13,14 Known: 28,30

Which set of assumptions is "correct "?

May need to report using more than one assumption set!

Reporting Multiple Conclusions

Different conclusions may result from using different assumptions.

If 2 contributors:  **EXCLUDED**

BUT

If ≥3 contributors:  **INCLUDED**
INCONCLUSIVE

REPORT ALL CONCLUSIONS!

MYTH

It may be useful to consider some DNA profiles under different assumptions.

Profile 6

Indistinguishable Mixture Profile

Known: 13,14 Known: 28,30

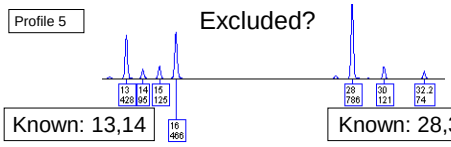
What if the genotypes CANNOT be distinguished?

Alleles are included, BUT are genotypes?

We know from previous data this person is excluded!
(assuming 2 contributors)

Is Known Individual Included or Excluded?

Profile 5

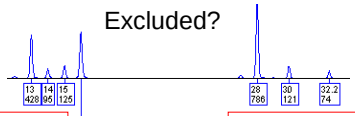


Known: 13,14 Known: 28,30

Which set of assumptions is "correct "?

Profile 5

Is Known Individual Included or Excluded?



Known: 14,15 16 466 **Known: 30,32.2**

Which set of assumptions is "correct"?

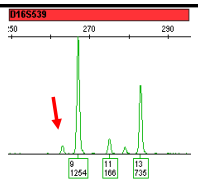
What if known genotypes are different and included as the single minor contributor under the assumption of only two contributors?
Include with appropriate statistics
What if ≥3 contributors? Include? Exclude? Inconclusive?

Profile 1

Stutter or true allele?
All alleles present?

Profile 1

Stutter or true allele?
All alleles present?



8,11 = true minor contributor

8 allele filtered out by software

If assume 8 is a stutter peak and assume all peaks are present, would exclude the true contributor!

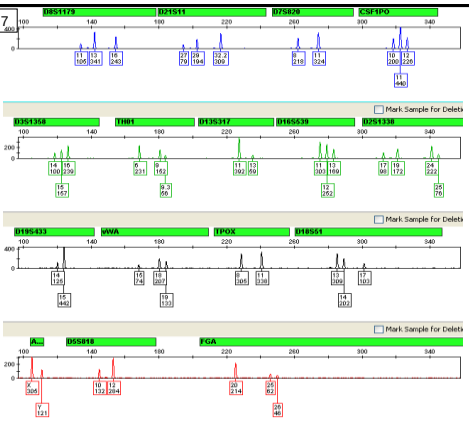
Uncertainty in evaluating the presence or absence of alleles leads to false inclusions and exclusions

Inclusion/Exclusion Criteria

- Must have a good interpretation procedure for excluding individuals who are non-contributors to the DNA sample
- If fail to exclude an individual as a possible contributor, you **MUST** have a statistical approach that embraces all of the possible included alleles and genotypes
- Must consider possible reasonable alternatives

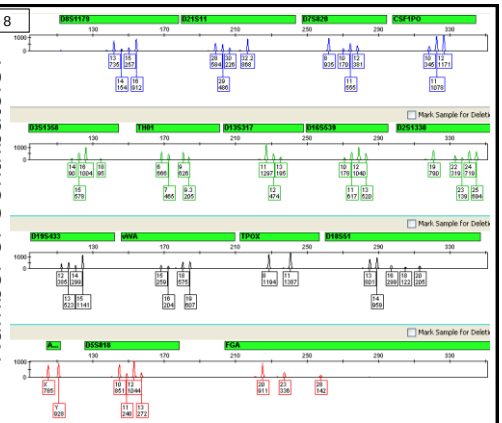
Profile 7

Major vs. Indistinguishable?
2 contributors?



Profile 8

Major vs. Indistinguishable?
Number of Contributors?



When to Consider Different Assumptions

May need to consider multiple assumptions for data interpretation when:

- Possible LT DNA profile
  - Stochastic effects (allelic drop-in, allelic drop-out, elevated stutter)
- Possible artifact vs. true allele
- Possible minor contributor in mixed DNA profile
- Possible known contributor(s) and deducing
- More than 2 contributors (later today)

What do you do when…

You have increased uncertainty, and therefore decreased confidence?

Options for interpreting and reporting:

1. Do not interpret the data → report inconclusive
   - When uncertainty is too high
2. Pick one interpretation to report
   - When have minimal uncertainty
3. Interpret and report the data under two or more different assumptions
   - When certainty is medium-to-high but possible scientifically sound alternatives exist

Different Experts→ Different Opinions

- Are the experts asking/answering the same question?
- Are they using the same information and data?
- Are they using the same interpretation methods?
- Are they using good scientific practices?
- Any possibility of bias?
- Are the differences meaningful or trivial?

Reporting

- Consider the data from several scientific perspectives – for conclusions *and* statistical calculations
- Report all appropriate scientific conclusions and opinions in the laboratory report
- ESPECIALLY if the conclusions differ under different reasonable assumptions

Why Report?

- Opinions may be important to different individuals reading the report (e.g., law enforcement, prosecutor, defense attorney, client, judge, jury)
- Reports should be **neutral** to the case yet address the question(s) asked by the client

Why Report?

- Not all cases (<10%) make it to court
- Critical decisions often based on report and (mis)understandings alone
- If not provided in advance to all parties, opinions may not be admissible in court

Summary

- EVERY interpretation requires assumptions
- Assumptions MUST be made from the data alone and prior to knowing the profiles of the known contributors
  - Artifact, stutter vs. true alleles
  - Number of contributors
  - Major:minor contributors
- All assumptions must be documented and should be reported
- Just because the known profile "fits" the data under one assumption set does not mean those are the correct assumptions and the correct conclusion

THANK YOU!!

John Butler
Mike Coble
Robin Cotton
Catherine Grgicak
Bruce Heidebrecht
& Workshop attendees

For many hours of discussions!

Catherine Grgicak
Robin Cotton
NIJ Grant to Boston University

For all of the profiles!

NIST National Institute of Standards and Technology • U.S. Department of Commerce

FORENSICS@NIST

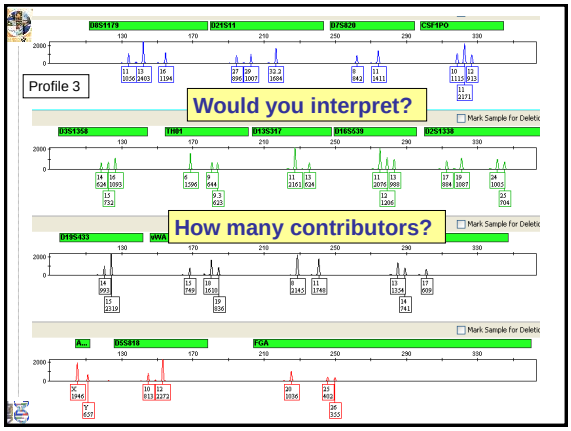
DNA Mixture Interpretation Webcast
April 12, 2013

http://www.nist.gov/oles/forensics/dna-analyst-training-on-mixture-interpretation.cfm

http://www.cstl.nist.gov/strbase/mixture.htm

Complex Mixtures

Charlotte J. Word
Consultant



Two-Person Mixtures

- **Lots of experience and familiarity with two-person mixtures, literature, validation studies, training samples**
- **Published guidelines for interpretation**
- **Well developed SOPs for interpretation**
- **Routine amount of input DNA in amplification generally leads to nice profiles**

Two-Person Mixtures

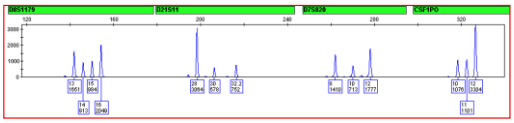
High Certainty Leads to High Confidence

- **Only two contributors present**
- **Distinguishing stutter/artifacts from true alleles**
- **Use stochastic threshold to assess if all alleles are likely present vs. LT DNA with stochastic effects**
- **Assessing mixture ratio (distinguishable/ major:minor or indistinguishable mixture)**
- **Deducing second contributor if one contributor is known**

Two-Person Mixtures

Assume number of contributors is two:

- **Aids in allele association at each locus based on peak height ratios**
- **May aid in genotype association for full profile based on mixture ratio**
- **Statistics calculations often straight forward**

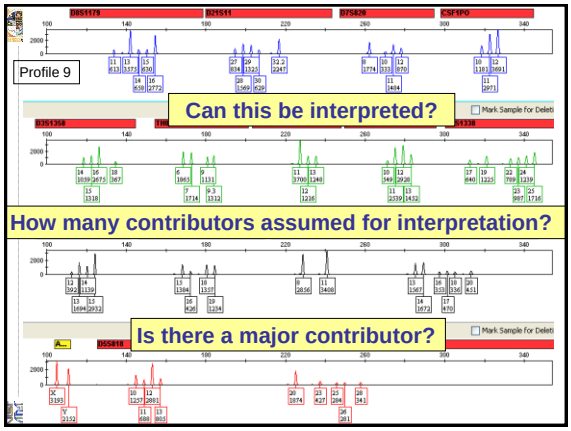


Complex Mixtures

- **Multiple contributors**
  - **3- & 4- person (or more!)**
- **Relatives in the Mixtures**

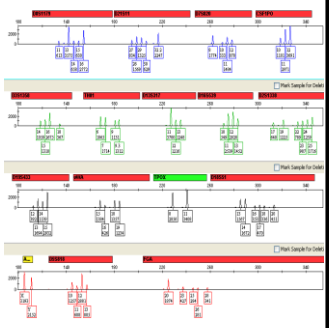
MYTH

It is easy to determine the number of contributors to a DNA profile.



Complex Mixture – Allele Summary

- 6 alleles at 2 loci
- 5 alleles at 3 loci
- 4 alleles at 7 loci
- 3 alleles at 2 loci
- 2 alleles at 1 locus
- 1 allele at 0 loci
- 63 total alleles



Two-Person Mixtures

Observed profile A B

14 total combinations

4 alleles
All heterozygotes and non-overlapping alleles

3 alleles
Heterozygote + heterozygote, one overlapping allele
Heterozygote + homozygote, no overlapping alleles

2 alleles
Heterozygote + heterozygote, two overlapping alleles
Heterozygote + homozygote, one overlapping allele
Homozygote + homozygote, no overlapping alleles

1 allele
Homozygote + homozygote, overlapping allele

Three-Person Mixtures

Observed profile

150 total combinations

6 alleles
All heterozygotes and non-overlapping alleles

5 alleles
Two heterozygotes and one homozygote
Three heterozygotes, one overlapping allele

4 alleles
Six combinations of heterozygotes, homozygotes and overlapping alleles

3 alleles
Eight combinations of heterozygotes, homozygotes, and overlapping alleles

2 alleles
Five combinations of heterozygotes, homozygotes, and overlapping alleles

1 allele
All homozygotes, overlapping allele

Four-Person Mixtures

Observed profile

MANY combinations

8 alleles
All heterozygotes and non-overlapping alleles

7 alleles
Several combinations of heterozygotes, homozygotes, and overlapping alleles

6 alleles
Many combinations

5 alleles
Many combinations

4 alleles
Many combinations

3 alleles
Many combinations

2 alleles
Many combinations

1 allele
All homozygotes, overlapping allele

Available online at www.sciencedirect.com
ScienceDirect
Forensic Science International: Genetics 1 (2007) 20–28
ELSEVIER
FSI GENETICS
www.elsevier.com/locate/fsig

Towards understanding the effect of uncertainty in the number of contributors to DNA stains

John S. Buckleton[a], James M. Curran[b,*], Peter Gill[c]

[a] The Institute of Environmental Science and Research Ltd., Private Bag 92021, Auckland, New Zealand
[b] Department of Statistics, University of Auckland, Private Bag 92019, Auckland, New Zealand
[c] The Forensic Science Service, Trident Court, Solihull Parkway, Birmingham Business Park, Solihull B37 7YN, UK

Received 31 May 2006; received in revised form 12 September 2006; accepted 13 September 2006

Abstract

DNA evidence recovered from a scene or collected in relation to a case is generally declared as a mixture when more than two alleles are observed at several loci. However, in principle, all DNA profiles may be considered to be potentially mixtures, even those that show not more than two alleles at any locus. When using a likelihood ratio approach to the interpretation of mixed DNA profiles it is necessary to postulate the number of potential contributors. However, this number is never known with certainty. The possibility of, a, say three-person mixture, presenting four or fewer peaks at each locus of the CODIS set was explored by Paoletti et al. [D.R. Paoletti, T.E. Doom, C.M. Krane, M.L. Raymer, D.E. Krane, Empirical analysis of the STR profiles resulting from conceptual mixtures, J. Forensic Sci. 50 (2005) 1361–1366]. In this work we extend this analysis to consider the profiler plus and SGM plus multiplices. We begin the assessment of the risk associated with current practice in the calculation of LR's. We open the discussion of possible ways to surmount this ambiguity.

Forensic Science International: Genetics 1 (2007) 20–28

Two-Person Simulated Mixtures – SGM+ Number of Alleles at each Locus

Table 1
The probability of observing a given number of alleles in a two-person mixtures for simulated profiles at the SGM+™ loci

Loci	No. of alleles			
	1	2	3	4
D3	0.011	0.240	0.559	0.190
vWA	0.008	0.194	0.548	0.250
D16	0.016	0.287	0.533	0.164
D2	0.003	0.094	0.462	0.441
D8	0.011	0.194	0.521	0.274
D21	0.007	0.147	0.505	0.341
D18	0.003	0.095	0.472	0.430
D19	0.020	0.261	0.516	0.203
THO	0.016	0.271	0.547	0.166
FGA	0.003	0.116	0.500	0.381

Buckleton et al. Forensic Science International: Genetics 1 (2007) 20–28

Three-Person Simulated Mixtures – SGM+ Number of Alleles at each Locus

Table 2
The probability of observing a given number of alleles in a three-person mixtures for simulated profiles at the SGM+™ loci

Loci	No. of alleles showing					
	1	2	3	4	5	6
D3	0.000	0.053	0.366	0.463	0.115	0.002
vWA	0.000	0.037	0.285	0.468	0.194	0.016
D16	0.001	0.086	0.397	0.411	0.100	0.005
D2	0.000	0.008	0.104	0.385	0.393	0.110
D8	0.001	0.041	0.258	0.436	0.236	0.029
D21	0.000	0.023	0.192	0.428	0.302	0.055
D18	0.000	0.007	0.109	0.392	0.396	0.096
D19	0.003	0.078	0.352	0.401	0.152	0.014
THO	0.001	0.074	0.395	0.439	0.088	0.002
FGA	0.000	0.012	0.144	0.424	0.346	0.074

Buckleton et al. Forensic Science International: Genetics 1 (2007) 20–28

2, 3, 4-Person Simulated Mixtures – CODIS Loci Number of Alleles at each Locus

J Forensic Sci, Nov. 2005, Vol. 50, No. 6
Paper ID JFS2004475
Available online at: www.astm.org

David R. Paoletti,[1] M.S.; Travis E. Doom,[1,2] Ph.D.; Carissa M. Krane,[3] Ph.D.; Michael L. Raymer,[1,2] Ph.D.; and Dan E. Krane,[4] Ph.D.

Empirical Analysis of the STR Profiles Resulting from Conceptual Mixtures

ABSTRACT: Samples containing DNA from two or more individuals can be difficult to interpret. Even ascertaining the number of contributors can be challenging and associated uncertainties can have dramatic effects on the interpretation of testing results. Using an FBI genotypes dataset, containing complete genotype information from the 13 Combined DNA Index System (CODIS) loci for 959 individuals, all possible mixtures of three individuals were exhaustively and empirically computed. Allele sharing between pairs of individuals in the original dataset, a randomized dataset and datasets of generated cousins and siblings was evaluated as were the number of loci that were necessary to reliably deduce the number of contributors present in simulated mixtures of four or less contributors. The relatively small number of alleles detectable at most CODIS loci and the fact that some alleles are likely to be shared between individuals within a population can make the maximum number of different alleles observed at any tested loci an unreliable indicator of the maximum number of contributors to a mixed DNA sample. This analysis does not use other data available from the electropherograms (such as peak height or peak area) to estimate the number of contributors to each mixture. As a result, the study represents a worst case analysis of mixture characterization. Within this dataset, approximately 3% of three-person mixtures would be mischaracterized as two-person mixtures and more than 70% of four-person mixtures would be mischaracterized as two- or three-person mixtures using only the maximum number of alleles observed at any tested locus.

Paoletti et al. J Forensic Sci, Nov. 2005, Vol. 50, No. 6

2- to 5-Person Simulated Mixtures – Identifiler Number of Alleles vs. Likelihood Estimator

J Forensic Sci, January 2011, Vol. 56, No. 1
doi: 10.1111/j.1556-4029.2010.01550.x
Available online at: interscience.wiley.com

PAPER
CRIMINALISTICS

Hinda Haned,[1] M.S.; Laurent Pène,[2] M.S.; Jean R. Lobry,[1] Ph.D.; Anne B. Dufour,[1] Ph.D.; and Dominique Pontier,[1] Ph.D.

Estimating the Number of Contributors to Forensic DNA Mixtures: Does Maximum Likelihood Perform Better Than Maximum Allele Count?

Haned et al. J Forensic Sci, January 2011, Vol. 56, No. 1

Number of Contributors – Total Number of Alleles

314 FORENSIC SCIENCE
doi: 10.3325/cmj.2011.52.314
CMJ

Estimating the number of contributors to two-, three-, and four-person mixtures containing DNA in high template and low template amounts

Jaheida Perez, Adele A. Mitchell, Nubia Ducasse, Jeannie Tamariz, Theresa Caragine
Office of Chief Medical Examiner of the City of New York, The Department of Forensic Biology, New York, NY, USA

Perez et al., Croat Med J. 2011; 52:314-26

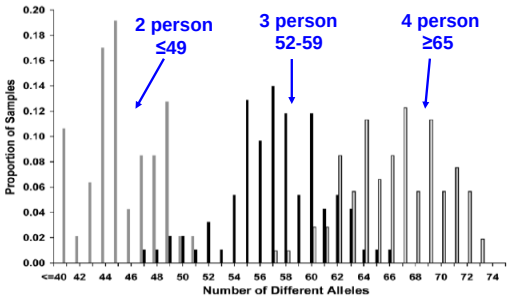


Figure 1. Expected # of different alleles from mixtures.

Estimating the number of contributors to two-, three-, and four-person mixtures containing DNA in high template and low template amounts
Perez et al., Croat Med J. 2011; 52:314-26

Two-Person Mixture Studies Summary

Based on Allele Counts Alone:

- **Always** recognized as a mixture – no risk of confusing as a single-source
  - Loci with 3 or 4 alleles
  - Peak height ratio imbalance at loci with 2 alleles
- Observe more loci with 2 or 3 alleles than 4 alleles – even when DNA from two heterozygous individuals were mixed
- 49 or fewer total alleles

Buckleton et al. Forensic Science International: Genetics 1 (2007) 20–28; Paoletti et al. J Forensic Sci, Nov. 2005, Vol. 50, No. 6; Haned et al. J Forensic Sci, January 2011, Vol. 56, No. 1; Perez et al., Croat Med J. 2011; 52:314-26

Three-Person Mixture Studies Summary

- No risk of confusing as a single-source
- Small risk of confusing with two-person mixture
  - Observe at least one locus with 5 or 6 alleles in ~97% of profiles (3% have ≤4 alleles)
  - Maximum allele count works most of time
  - 3% profiles look like 2-person mixture
  - Risk if LT-DNA, degradation, inhibition, primer mutation to look like 2-person mixture
- Most loci have 3 or 4 alleles
- 52-59 total alleles

Buckleton et al. Forensic Science International: Genetics 1 (2007) 20–28; Paoletti et al. J Forensic Sci, Nov. 2005, Vol. 50, No. 6; Haned et al. J Forensic Sci, January 2011, Vol. 56, No. 1; Perez et al., Croat Med J. 2011; 52:314-26

Four-Person Mixture Studies Summary

- No risk of confusing as a single-source
- Very small risk of confusing with two-person mixture
  - Likely to have peak height imbalance
- **Very small number of loci with 8 alleles and very few with 7 alleles**
  - High risk of confusing with three-person mixture
  - Risk if LT-DNA, degradation, inhibition, primer mutation
- ≥65 total alleles

Buckleton et al. Forensic Science International: Genetics 1 (2007) 20–28; Paoletti et al. J Forensic Sci, Nov. 2005, Vol. 50, No. 6; Haned et al. J Forensic Sci, January 2011, Vol. 56, No. 1; Perez et al., Croat Med J. 2011; 52:314-26

Four-Person Mixture Studies Summary

>70% of 4-person mixtures would NOT be recognized as 4-person mixtures based on maximum number allele count at a locus

Buckleton et al. Forensic Science International: Genetics 1 (2007) 20–28; Paoletti et al. J Forensic Sci, Nov. 2005, Vol. 50, No. 6; Haned et al. J Forensic Sci, January 2011, Vol. 56, No. 1; Perez et al., Croat Med J. 2011; 52:314-26

Five-, Six- Person Mixture Studies Summary

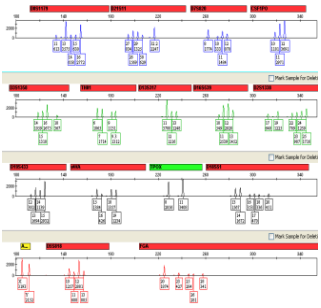
- >99% of 5 person mixtures would look like 4 person mixtures (~60%) or 3-person mixtures (~40%)
- Most 6 person mixtures would look like 5 person mixture (6%), 4-person mixtures (80%) or 3-person mixtures (14%)

Wang, T.W., Kalet, P., Pendleton, J., Gilbert, K., Lucas, L. and Birdwell, J.D. 2005 The probable number of contributors to a STR DNA mixture. http://www.promega.com/products/pm/genetic-identity/ishi-conference-proceedings/16th-ishi-poster-abstracts/; Haned et al. J Forensic Sci, January 2011, Vol. 56,(1), 23-28

Complex Mixture – Allele Summary

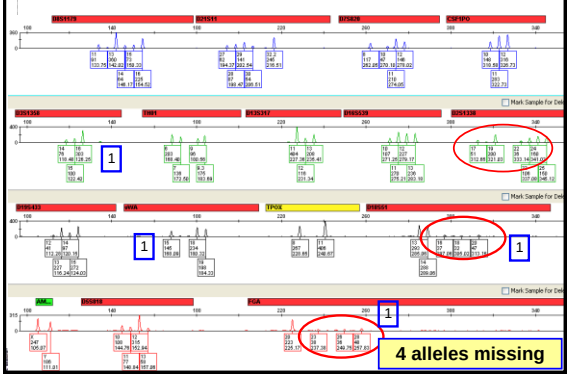
- 6 alleles at 2 loci
- 5 alleles at 3 loci
- 4 alleles at 7 loci
- 3 alleles at 2 loci
- 2 alleles at 1 locus
- 1 allele at 0 loci
- 63 total alleles

No Major Contributor!

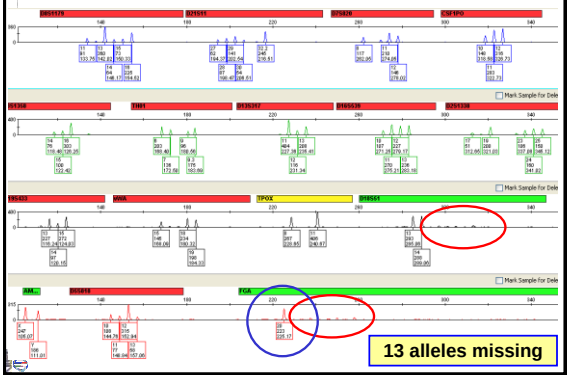


A 4-person mixture @ 3:2:1.6:1 ratio!!

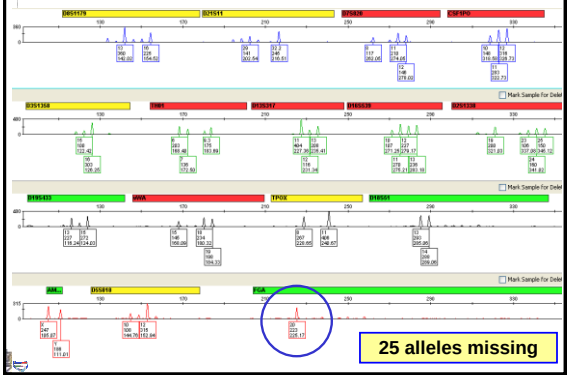
Mixture with 30 RFU Analytical Threshold



Mixture with 50 RFU Analytical Threshold



Mixture with 100 RFU Analytical Threshold



Mixture with 100 RFU Analytical Threshold

- Looks like it could be a two-person mixture
- Looks like it may have a major contributor at some loci, but not all → indistinguishable mixture?
- Many alleles near or above 150-200 RFU

Good to interpret?

Mixture with 100 RFU Analytical Threshold

- If compare this profile to the known contributors:
  - The highest peak or peaks are not always from the person with the most DNA (3:2:1.6:1)
  - The highest peaks are not consistent with any of the known contributors over the profile
  - Cannot correctly "pull out" any one or two of the correct contributors at all loci
  - The "major" contributor is missing an allele from this profile
- Allele shares complicate mixture interpretation
- Allele shares can cause high peaks that are suggestive of major contributor profiles
- Stochastic effects lead to loss of data

CPI Statistical Frequencies with Different Analytical Thresholds

	Frequency of 1 in __ unrelated individuals			
	Full Profile	30 RFU	50 RFU	100 RFU
Caucasian	5,300	45,000	2,400,000*	5.7 billion*
African American	25,000	250,000	290,000,000*	870 billion*
SW Hispanic	4,400	75,000	10,000,000*	20 billion*

*Single allele at one locus; p^2 in calculation rather than 2p

Total # of Alleles	63	59	50	38
# of Alleles Missing	--	4	13	25

Thanks to Liz Benzinger and Kristen Slaper for the PopStats Calculations!

MYTH

It is easy to determine the number of contributors to a DNA profile.

Steps in the interpretation of mixtures

(Clayton *et al. Forensic Sci. Int.* 1998; 91:55-70)

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 → Statistics

Step #6 Compare Reference Samples

Modified from slide from Dr. John Butler

INCREASED COMPLEXITY

↓

HIGH UNCERTAINTY

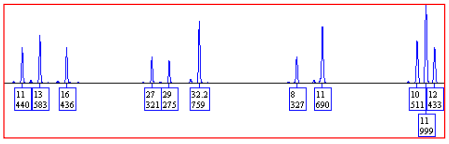
↓

LACK OF CONFIDENCE

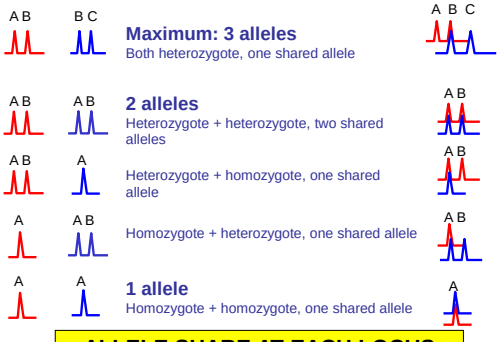
Complex Mixtures

Mixtures with Relatives

Parent-Child
Sibling-Sibling



Parent + Child → Mixture DNA Profile Pattern



P1 + P2	Genotypes of Children	% Sibling Allele Sharing
A B / C D	AC or AD or BC or BD	0%, 50% or 100%
A B / B C	AB or AC or BB or BC	0%, 50% or 100%
A B / A B	AB/BA or AA or BB	0%, 50% or 100%
A B / C	AC or BC	50% or 100%
A B / A	AA or BA	50% or 100%
A / B	AB	100%
A / A	AA	100%

P1 = Parent 1; P2 = Parent 2

Allele Sharing in Relatives

ELSEVIER

Forensic Science International 131 (2003) 85–89

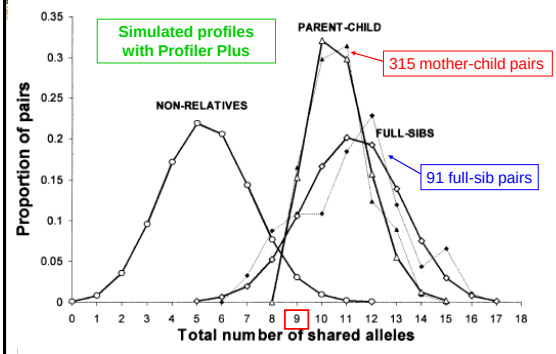
Forensic Science International

www.elsevier.com/locate/forsciint

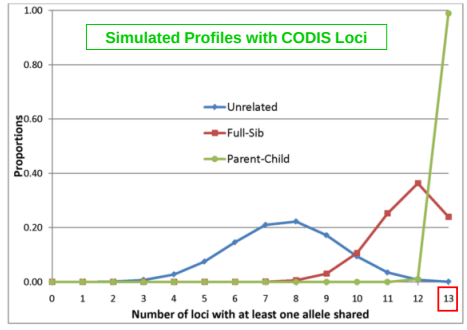
Allele sharing in first-degree and unrelated pairs of individuals in the Ge.F.I. AmpFlSTR® Profiler Plus™ database

Silvano Presciuttini[a,1,*], Francesca Ciampini[a], Milena Alì[b], Nicoletta Cerri[c], Marina Dobosz[d], Ranieri Domenici[e], Gabriella Peloso[f], Susi Pelotti[g], Andrea Piccinini[h], Elena Ponzano[i], Ugo Ricci[l], Adriano Tagliabracci[g], J.E. Baley-Wilson[j], Francesco De Stefano[m], Vincenzo Pascali[g,1]

Presciuttini et al. Forensic Science International 131 (2003) 85-89



Forensic Science International 131 (2003) 85-89



Ge et al. Comparisons of the familial DNA databases searching policies. J. Forensic Sci. 2011;56(6):1448-56.

Mixtures with Relatives – Summary

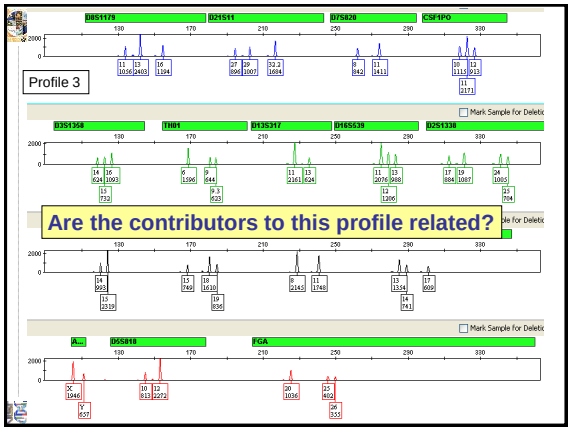
Parent-Child

- Expect at least 50% allele share
- Expect at least one shared allele at each locus
- Maximum 3 alleles per locus (in absence of mutation)
- If test X loci, expect >X allele shares (9-14 Profiler Plus; 13-20 CODIS)

Mixtures with Relatives – Summary

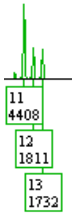
Sibling-Sibling

- Expect at least 50% allele share overall, but variable: 7-16 Profiler Plus; 12-22 CODIS (≥X-1)
- Expect 0, 50 or 100% allele share at each locus
- Expect at least one allele share at 9-13 loci (CODIS data)



Mixtures with Relatives – Working Backwards from Mixed DNA Profile

- With mixed DNA profile from unknowns, may not know if alleles are shared
- Data in the graphs are not helpful



11,12 + 11,13

or

11,11 + 12,13

Relative? Parent-Child? Sibs?

Relative? Sibs?

Unrelated?

True Known Contributors to Previous Profile

- Share 14 alleles over 15 Identifiler loci
  - 8 alleles at 9 Profiler Plus loci
  - 13 alleles at 13 CODIS loci
  - 15 alleles 17 loci (Identifiler + PowerPlex 16 HS)
- One allele in common at each locus, except D2, FGA and Penta E
- Likely not parent, unless mutations occurred
- Sibs?
  - Using known contributors' profiles : Inconclusive from allele #; Ge locus data suggests sibs

Relatives in DNA Testing

What if the true contributors **in a mixed DNA sample** are closely related?

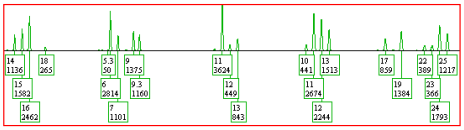
- Significant issue with the types of samples being tested today (e.g., "Touch" DNA)
  - Any item likely to routinely be used/shared by related individuals (e.g., living in same household, driving same car, sharing clothing)
  - Relatives committing crimes together (e.g., shared clothing, weapons)
  - Multiple homicides involving family members
- NO statistical method to address this
- Statistics reported for "random individuals"
  - However, a relative is more likely to be included

Complex Mixture Interpretation

- We have limited experience with known complex mixtures (training, validation, or proficiency tests)
- No or limited published guidelines for interpretation
- Limited interpretation SOPs available
- Routine amount of DNA amplified → poor quality profiles, LT DNA likely for 1 or more contributors
- How do you do the statistical calculations?

Complex Mixture Interpretation

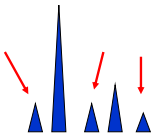
Is hard because the parameters used to interpret two-person mixtures often may not be directly applicable to complex mixtures



Complex Mixtures

More Uncertainty and Lack of Confidence

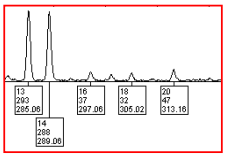
- Peak vs. Artifacts
  - Stutter?
  - Pull-up?
  - True Allelle?



Complex Mixtures

More Uncertainty and Lack of Confidence

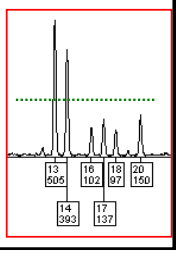
- High likelihood that DNA from one or more contributors is below optimal range
  - LT DNA = stochastic effects
  - Missing alleles? (allele drop out)
  - Elevated Stutter? True allele vs. Stutter?
  - Allele drop-in?



Complex Mixtures

More Uncertainty and Lack of Confidence

- Stochastic threshold
  - Only meaningful for the peaks below the value – may be missing sister allele
  - Only helps with assessing if ALL alleles are likely present

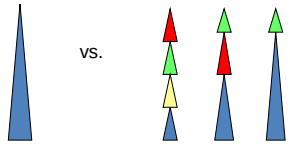


Complex Mixtures

More Uncertainty and Lack of Confidence

- Stochastic threshold
  - NO meaning for peaks above the value –
    - Major contributor?
    - Shared alleles? How many shares? Relatives or unrelated?

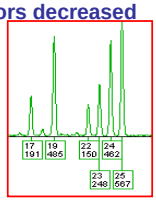
Major vs. Shared alleles



Complex Mixtures

More Uncertainty and Lack of Confidence

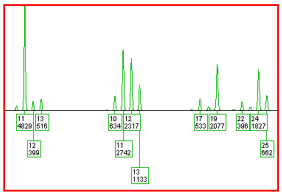
- Peak height ratios have no meaning at most or all loci
  - Cannot use to associate alleles into genotypes
  - Ability to deduce other contributors decreased even if you know one contributor



Complex Mixtures

More Uncertainty and Lack of Confidence

- Mixture ratio cannot be calculated
  - Different amount from each contributor likely with no way to determine
  - Cannot use to associate genotypes into profiles



Complex Mixtures

More Uncertainty and Lack of Confidence

- **Number of contributors – maximum allele count/minimum number often an underestimate**
  - **What number to assume?**
  - **May need to interpret under multiple assumptions (especially if the conclusion changes)**

Complex Mixtures

More Uncertainty and Lack of Confidence

- **"Inclusion" based on alleles NOT based on genotypes → may not be correct inclusion**
- **False Inclusions**
  - Increased risk as # of alleles increase
- **How calculate statistical frequency?**

Complex Mixtures

Exclusions less likely/ Exclusion criteria difficult to develop

- Can anyone be excluded if LT DNA present?
- Partial "inclusions"

Estimate frequency of included individuals can be quite common – can become meaningless (1 in 2 individuals)

Inconclusive reporting increased

What can we do?

- Amplify more DNA?
- Test another portion of the sample?
- Test another sample in the case?
- Probabilistic approaches to interpretation? (stay tuned)

Conclusions

- Criteria routinely used in crime laboratories for the interpretation of two-person mixtures may not apply for most complex mixtures
- LT-DNA, degradation, inhibition play more significant role
- Additional complex mixtures need to be generated and evaluated for establishment of scientifically supported interpretation guidelines

THANK YOU!!

John Butler
Mike Coble
Robin Cotton
Catherine Grgicak
Bruce Heidebrecht
& Workshop attendees

For many hours of discussions!

Catherine Grgicak
Robin Cotton
NIJ Grant to Boston University

For all of the profiles!

NIST National Institute of Standards and Technology • U.S. Department of Commerce

DNA Mixture Interpretation Webcast
April 12, 2013

http://www.nist.gov/oles/forensics/dna-analyst-training-on-mixture-interpretation.cfm

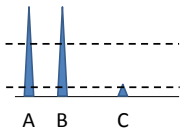
http://www.cstl.nist.gov/strbase/mixture.htm

Probabilistic Genotyping

Michael D. Coble
National Institute of Standards and Technology

What should we do with discordant data?

- Ignore/drop the locus – this is the "most conservative" option.



Complainant = AB
POI = CD

Curran and Buckleton (2010)

JOURNAL OF FORENSIC SCIENCES

J Forensic Sci, September 2010, Vol. 55, No. 5
doi: 10.1111/j.1556-4029.2010.01446.x
Available online at: interscience.wiley.com

PAPER
CRIMINALISTICS; GENERAL

James M. Curran,[1] M.Sc.(Hons.), Ph.D. and John Buckleton,[2] Ph.D.

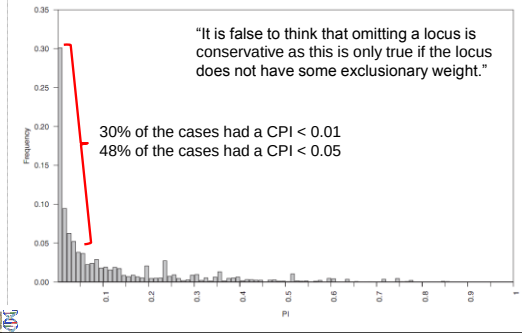
Inclusion Probabilities and Dropout

Created 1000 Two-person Mixtures (Budowle *et al.*1999 AfAm freq.).

Created 10,000 "third person" genotypes.

Compared "third person" to mixture data, calculated PI for included loci, ignored discordant alleles.

Curran and Buckleton (2010)

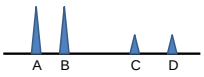


"It is false to think that omitting a locus is conservative as this is only true if the locus does not have some exclusionary weight."

30% of the cases had a CPI < 0.01
48% of the cases had a CPI < 0.05

Curran and Buckleton (2010)

"It is false to think that omitting a locus is conservative as this is only true if the locus does not have some exclusionary weight."



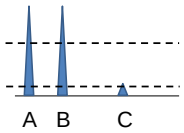
POI = C,D

"Conservative"

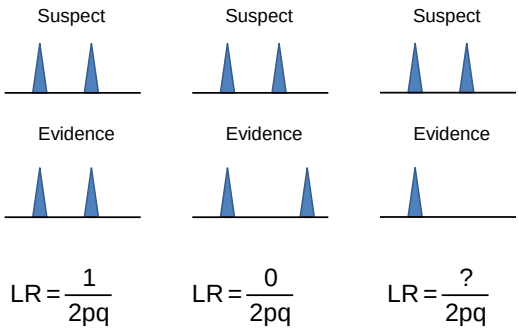
Dropping a locus is beneficial to the "guilty" and detrimental to the "innocent".

What should we do with discordant data?

- Ignore/drop the locus – this is the "most conservative" option.



Complainant = AB
POI = CD



Whatever way uncertainty is approached, probability is the *only* sound way to think about it.

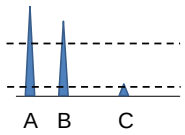
-Dennis Lindley

What should we do with discordant data?

- Continue to use RMNE (CPI, CPE)
- Use the Binary LR with 2p
- Semi-continuous methods with a LR (Drop models)

Drop Models

- Examine the alleles present and include a Pr(D) in the LR calculation



Alleles Present
ABCF

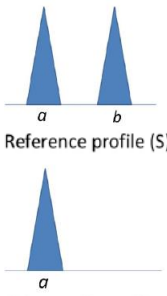
December 2012 Issue of FSI-G


Forensic Science International: Genetics 6 (2012) 679–688
Contents lists available at SciVerse ScienceDirect
Forensic Science International: Genetics
journal homepage: www.elsevier.com/locate/fsig

DNA commission of the International Society of Forensic Genetics: Recommendations on the evaluation of STR typing results that may include drop-out and/or drop-in using probabilistic methods

P. Gill [a,b,*], L. Gusmão [c], H. Haned [d], W.R. Mayr [e], N. Morling [f], W. Parson [g], L. Prieto [h], M. Prinz [i], H. Schneider [j], P.M. Schneider [k], B.S. Weir [l]

ISFG Recommendations



Pr(D) = Prob. Drop-out (het)

Pr($\overline{D}$) = No Prob. Drop-out (het)

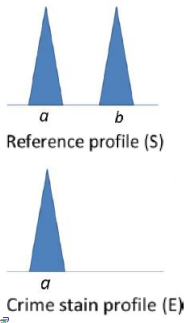
Pr(D_2) = Prob. Drop-out (hom)

Pr($\overline{D_2}$) = No Prob. Drop-out (hom)

Pr(C) = Prob. Drop-in

Pr($\overline{C}$) = No Prob. Drop-in

Prosecutor's Explanation



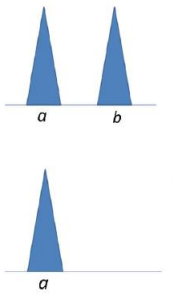
No Drop-out of the "A" allele
The "B" allele dropped out
No other Drop-in

$Pr(\overline{D})\ Pr(D)\ Pr(\overline{C})$

The LR

$$LR = \frac{Pr(\overline{D})\ Pr(D)\ Pr(\overline{C})}{\quad}$$

Defense Explanation

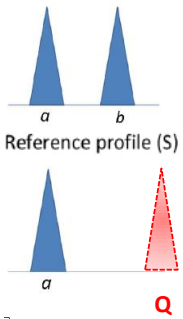


4 possibilities

(1) The real culprit is a homozygote

$p_a^2 Pr(\overline{D}_2)\ Pr(\overline{C})$

Defense Explanation



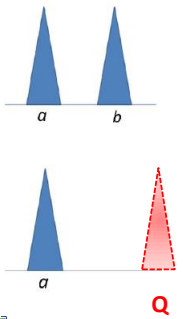
4 possibilities

(2) Drop out of a heterozygote (not B)
No drop-in of "A"

$2p_a p_Q Pr(\overline{D}) Pr(D) Pr(\overline{C})$

Q

Defense Explanation



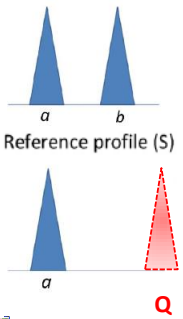
4 possibilities

(3) Drop out of a homozygote (not B)
Drop in of "A"

$p_Q^2 Pr(D_2)\ Pr(C) p_a$

Q

Defense Explanation



4 possibilities

(4) Drop out of a homozygote (not AB)
Drop in of "A"

$2p_Q p_{Q'} Pr(D)^2\ Pr(C) p_a$

Q Q'

The LR

$$LR = \frac{Pr(\bar{D})\ Pr(D)\ Pr(\bar{C})}{\begin{array}{ll} p_a^2 Pr(\bar{D}_2)\ Pr(\bar{C}) & + \\ 2p_a p_Q Pr(\bar{D}) Pr(D) Pr(\bar{C}) & + \\ p_Q^2 Pr(D_2)\ Pr(C) p_a & + \\ 2p_Q p_{Q'} Pr(D)^2\ Pr(C) p_a & \end{array}}$$

Some Drop Model Examples

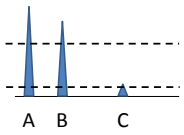
- LR mix (Haned and Gill)
- Balding and Buckleton (R program)
- FST (NYOCME, Mitchell *et al.*)
- Kelly *et al.* (University of Auckland, ESR)
- Lab Retriever (Lohmueller, Rudin and Inman)

What should we do with discordant data?

- Continue to use RMNE (CPI, CPE)
- Use the Binary LR with 2p
- Semi-continuous methods with a LR (Drop models)
- Fully continuous methods with LR

Continuous Models

- Mathematical modeling of "molecular biology" of the profile (mix ratio, PHR (Hb), stutter, etc…) to find optimal genotypes, giving WEIGHT to the results.



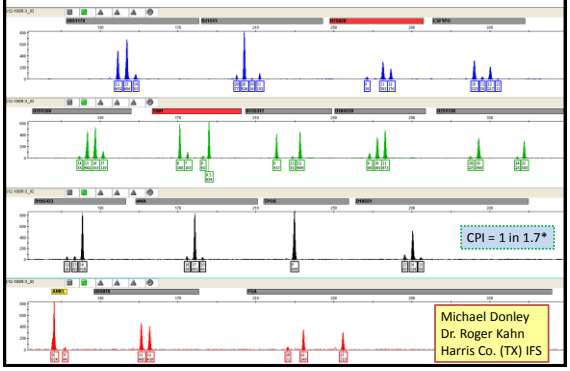
Probable Genotypes

AC – 40%
BC – 25%
CC – 20%
CQ – 15%

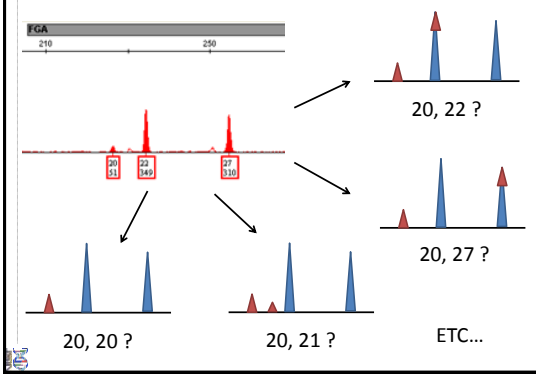
Some Continuous Model Examples

- TrueAllele (Cybergenetics)
- STRmix (ESR [NZ] and Australia)
- Cowell et al. (FSI-G (2011) **5**:202-209)

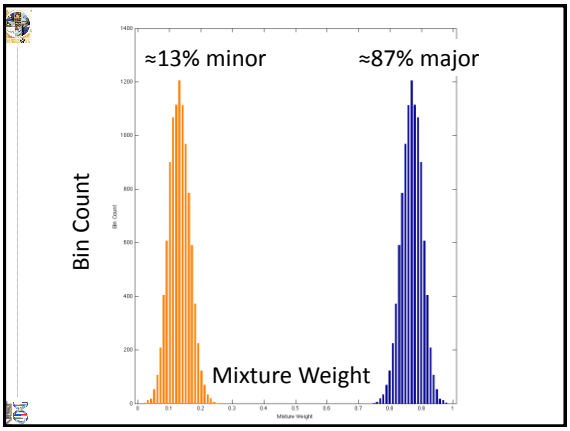
Challenging Mixture

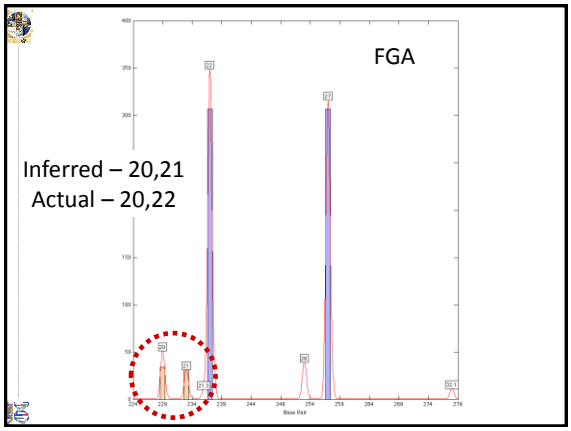


Challenging Mixture



TrueAllele Results





Statistical Calculation

	Inferred	Prob.	HWE	Suspect
FGA	20, 22	0.1474	0.0543	1
	20, 21	0.0722	0.0461	0
	20, 26	0.1309	0.0058	0
	20, 20	0.0882	0.0156	0
	21, 22	0.0056	0.08	0
	21, 26	0.0176	0.0085	0
	22, 26	0.0077	0.01	0
	20, 27	0.0142	0.0008	0
	22, 22	0.001	0.0471	0

H_P

$$LR = \frac{0.1474}{}$$

Statistical Calculation

	Inferred	Prob.	HWE	Pr*HWE
FGA	20, 22	0.1474	0.0543	0.008
	20, 21	0.0722	0.0461	0.0033
	20, 26	0.1309	0.0058	0.0008
	20, 20	0.0882	0.0156	0.0014
	21, 22	0.0056	0.08	0.0004
	21, 26	0.0176	0.0085	0.0001
	22, 26	0.0077	0.01	0.0001
	20, 27	0.0142	0.0008	0
	22, 22	0.001	0.0471	0

$\sum$ 0.0143

H_D

$$LR = \frac{0.1474}{0.0143}$$

LR = 10.33

STRmix

Mixture Proportions
Contributor 1 - 87%
Contributor 2 - 13%

```
-------------------------------------------------------
GENOTYPE PROBABILITY DISTRIBUTION
FGA
[22,27]     [20,20]          0.3293750584474653
[22,27]     [-1,20]          0.17910853800572998
[22,27]     [20,22]          0.23609375398505444
[22,27]     [20,27]          0.25542264956175026
-------------------------------------------------------

Locus 15(FGA): Pr(E|Hp) = 0.05152, Pr(E|Hd) = 5.0E-4, LR = 103.29885
```

Summary of the Issues

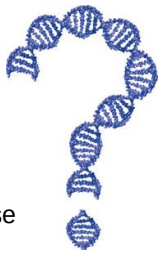
- New kits, new instruments will only increase the difficulties of interpreting low-level, challenging samples.
- If we are really serious about properly interpreting low level and complex mixtures, we must move away from the RMNE mentality. POPSTATS will not do!!
- Probabilistic methods are the way forward and a number of software programs are available ranging from “open source” to commercial packages.

Thank you for your attention

Contact Information

Michael D. Coble
Forensic Biologist
michael.coble@nist.gov
301-975-4330

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



Additional DNA mixture information available at:
http://www.cstl.nist.gov/strbase/mixture.htm

NIST National Institute of Standards and Technology • U.S. Department of Commerce

FORENSICS@NIST

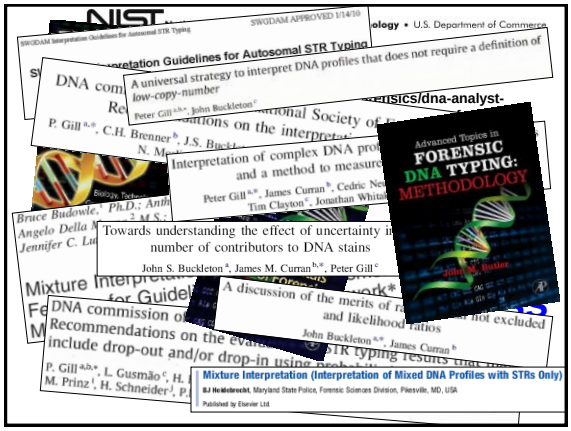
DNA Mixture Interpretation Webcast
April 12, 2013

http://www.nist.gov/oles/forensics/dna-analyst-training-on-mixture-interpretation.cfm

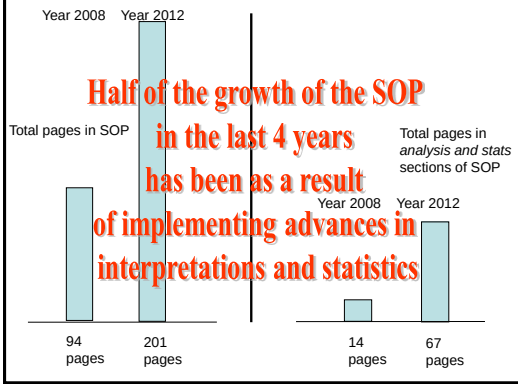
http://www.cstl.nist.gov/strbase/mixture.htm

Updating SOPs
&
improving the lab process

Bruce J. Heidebrecht
Maryland State Police,
Forensic Sciences Division



Maryland State Police DNA casework SOP



Attitude towards change

"When I go back to my lab with these changes the analysts are going to come at me with pitchforks!"

Attitude towards change

"It's going to be difficult…

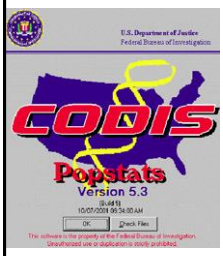
But we know that we'll be better analysts."

Meetings, not just articles

- Forensic scientist meetings:
  AAFS, LAFS, MAAFS, MAFS, NEAFS, NJAFS, NWAFS, SAFS, SWAFS, etc…….
- DNA Technical Leader Summit to be held November 20-21, 2013.

(hopefully…..$$$$$$$)

Know your software

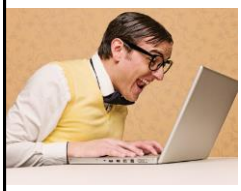


- Popstats is not programmed to perform Likelihood Ratio stats that include the possibility of undetected data.
- Popstats is not programmed to perform Likelihood Ratio stats that include peaks "indistinguishable from stutter".
- Popstats is not programmed to perform Restricted Likelihood Ratios.

Purchase your software

- Be aware of what the software can and cannot do.
- Be aware of system requirements between the CE instrument software, interpretation or stats software, and computer operating system.

Create your own software



- MDSP created our own Excel spreadsheets for ULR stats that can incorporate both dropout and IFS.
- Created our own Excel workbook for mixture deconvolution.

Transition period

- Most labs are not allowed to shut down in order to have time to learn new procedures.
- Analysts have to learn new procedures while issuing reports under current policies.
- This transition period can be very frustrating.

Transition period

- Hold regular meetings to discuss known mixtures and/or interesting casework mixtures.
- Learn from each other.
- Ask "Why?"

Greg Matheson on Forensic Science Philosophy

- The CAC News – 2nd Quarter 2012 – p. 6
  "Generalist vs. Specialist: a Philosophical Approach"
  http://www.cacnews.org/news/2ndq12.pdf
- If you want to be a technician, performing tests on requests, then just focus on the policies and procedures of your laboratory. If you want to be a scientist and a professional, learn the policies and procedures, but go much further and learn the philosophy of your profession. **Understand the importance of why things are done** the way they are done, the scientific method, the viewpoint of the critiques, the issues of bias and the importance of ethics.

Slide created by John Butler

Writing an SOP

- Involve the analysts in the SOP review process to gather feedback before implementation.
- Review other labs' protocols and report writing guidelines
  - STRBase as a resource

DNA Technical Leader from another lab:

"Thanks for doing this workshop. It will help me make a perfect SOP." …

"It will help me make a better SOP."

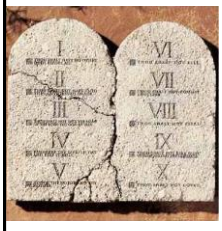
Writing an SOP

- The interpretation of results in casework is a matter of professional judgment and expertise.
- Not every situation can or should be covered by a pre-set rule.
- However, it is important that the laboratory develops and adheres to minimum criteria for interpretation of analytical results.
- These criteria are based on validation studies, literature references, and casework experience and are developed with maximum input from analysts.

Updating an SOP

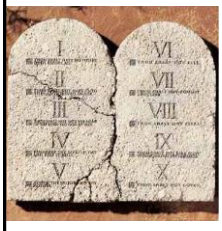
- Alleles in stutter positions ("N-1" repeat positions) with a ratio (RFU of the "N-1" peak divided by the RFU of the "N" peak) equal to or below the following stutter guidelines will be designated as stutter ***and no conclusions will be drawn from these stutter peaks.***
- Peaks in stutter positions ("N-1" repeat positions) with a ratio (RFU of the "N-1" peak divided by the RFU of the "N" peak) equal to or below the following stutter guidelines will be designated as stutter, ***or may be designated as "indistinguishable from stutter" in the case of mixtures based upon the criteria in Sections ….***

Reviewing a case



- The interpretation of results in casework is a matter of professional judgment and expertise.
- As long as expert opinion is a part of interpretation there will be some amount of differences between analysts.
- The goal of the rules in the SOP should be to minimize interpretation differences between analysts.
- Reviewers need to be aware of what is wrong and what is professional judgment.

Reviewing a case



- Reviewers need to be aware of what is wrong and what is professional judgment.

SOP states that a minimum of 5 loci needed to be able to declare a match.

Analyst declares a match using only 3 loci.

SOP violation.

Reviewing a case

- Reviewers need to be aware of what is wrong and what is professional judgment.

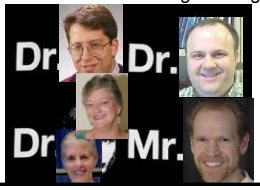
SOP states that both mixture deconvolution with RMP stat and likelihood ratio are appropriate to use for a given mixture.

While analyst could declare an RMP match with a stat of 1 in a billion, they decide to use LR with a stat of 10 million to 1.

Professional judgment.

Thank you

- All the members of SWGDAM who have helped further my knowledge of mixture interpretation.
- Maryland State Police for allowing me time to participate in a variety of workshops.
- NIST for hosting this event.
- Dr. Butler for inviting me to participate.
- To Howard Wolowitz, who proves that a guy with a Master's degree can be included along with a group of PhD's!



NIST National Institute of Standards and Technology • U.S. Department of Commerce

FORENSICS@NIST

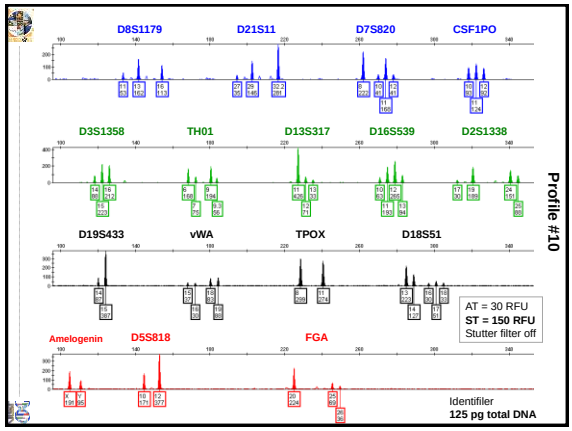
DNA Mixture Interpretation Webcast
April 12, 2013

http://www.nist.gov/oles/forensics/dna-analyst-training-on-mixture-interpretation.cfm

http://www.cstl.nist.gov/strbase/mixture.htm

Low Template DNA Challenges
and Validation Suggestions

John M. Butler
National Institute of Standards and Technology



Clayton et al. (1998)
ISFG (2006) Rec. #4

Step #1 Identify the Presence of a Mixture

Step #2 Designate Allele Peaks

Step #3 Identify the Number of Potential Contributors

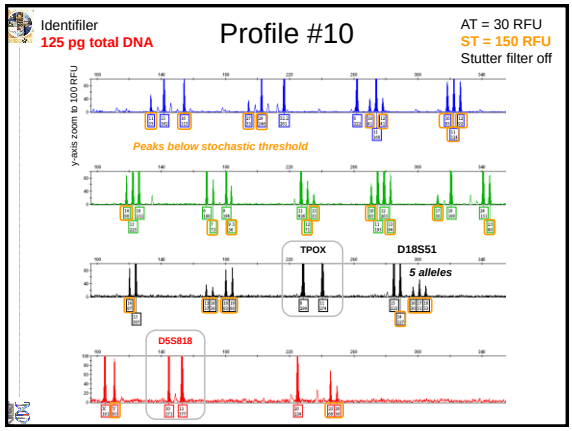
Step #4 Estimate the Relative Ratio of Contributors

Step #5 Consider All Possible Genotype Combinations

Step #6 Compare Reference Samples

Impact of Results with Low Level DNA

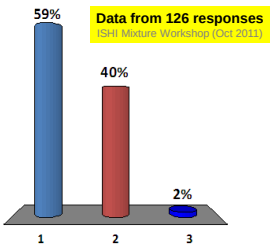
When amplifying low amounts of DNA (e.g., 125 pg), allele dropout is a likely possibility leading to **higher uncertainty** in the potential number of contributors and in the possible genotype combinations



Previous Response to This Question

Would you do a CPE/CPI statistic on TPOX and D5S818 because all alleles are above the stochastic threshold?

1. Yes
2. No
3. I don't work in a lab

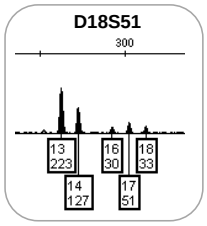


Data from 126 responses
ISHI Mixture Workshop (Oct 2011)

What Can We Say about this Result?

- Low level DNA (only amplified 125 pg total DNA)
  - likely to exhibit stochastic effects and have allele dropout
- Mixture of at least 3 contributors
  - Based on detection of 5 alleles at D18S51
  - If at equal amounts, ~40 pg of each contributor (if not equal, then less for the minor contributors); **we expect allele dropout**
- At least one of the contributors is male
  - Based on presence of Y allele at amelogenin
- Statistics if using CPI/CPE
  - Would appear that we can only use TPOX and D5S818 results with a stochastic threshold of 150 RFU *(will explore this further)*
- **Due to potential of excessive allele dropout, we are unable to perform any meaningful Q-K comparisons**

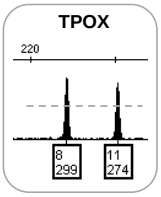
Uncertainty in the Potential Number of Contributors with this Result



5 alleles observed

- Several of the peaks are barely above the analytical threshold of 30 RFU
  - In fact, with an analytical threshold of 50 RFU or even 35 RFU, there would only be three detected alleles at D18S51
- Stochastic effects could result in a high degree of stutter off of the 17 allele making alleles 16 and 18 potential stutter products
- No other loci have >4 alleles detected

All Detected Alleles Are Above the Stochastic Threshold – **Or Are They?**



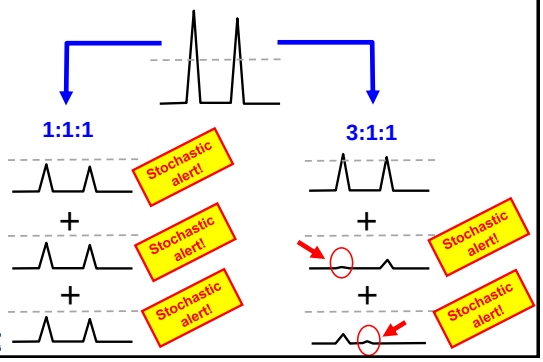
Stochastic threshold = 150 RFU

Does this result guarantee no allele drop-out?

We have assumed three contributors. If result is from an equal contribution of 3 individuals…

Then some alleles from individual contributors would be below the stochastic threshold and we could not assume that all alleles are being observed!

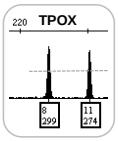
Assuming Three Contributors…
Some Possible Contributions to This Result



All Loci Are Not Created Equal
when it comes to mixture interpretation

- In the case of less polymorphic loci, such as TPOX, there are fewer alleles and these occur at higher frequency. Thus, there is a greater chance of allele sharing (peak height stacking) in mixtures.
- **Higher locus heterozygosity is advantageous for mixture interpretation** – we would expect to see more alleles (within and between contributors) and thus have a better chance of estimating the true number of contributors to the mixture

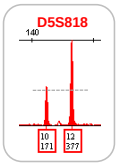
Even if you did attempt to calculate a CPI/CPE statistic using loci with all observed alleles above the stochastic threshold on this result…



TPOX Allele Frequencies (NIST Caucasian, Butler et al. 2003)
8 = 0.53
11 = 0.24
CPI = $(0.53 + 0.24)^2$ = 0.59 or **59%**

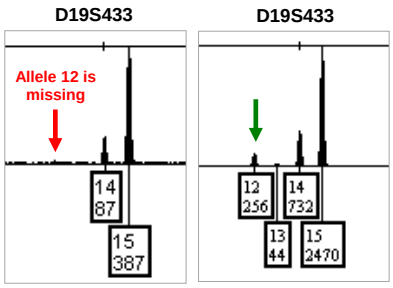
Combine loci = 0.59 x 0.18 = 0.11 or **11%**

Approximately 1 in every 9 Caucasians could be included in this mixture



D5S818 Allele Frequencies (NIST Caucasian, Butler et al. 2003)
10 = 0.05
12 = 0.38
CPI = $(0.05 + 0.38)^2$ = 0.18 or **18%**

Impact of Amplifying More DNA



125 pg total DNA amplified

500 pg total DNA amplified

True Contributors
3 contributors with a 2:1:1 mixture

15,15 (2x)
14,15 (1x)
12,14 (1x)

How should you handle the suspect comparison(s) with this case result?

- **No suspect comparisons should be made** as **the mixture result has too much uncertainty** with stochastic effects that may not account for all alleles being detected
- **It would be best to declare the mixture result "inconclusive"**
  - Report wording could include an additional phrase to emphasize that low signal makes this result **inadequate for ANY comparisons to potential reference sample(s)** using currently available techniques

How not to handle this result

- "To heck with the analytical and stochastic thresholds", **I am just going to see if the suspect profile(s) can fit into the mixture allele pattern observed** – and then if an allele is not present in the evidentiary sample try to explain it with possible allele dropout due to stochastic effects
- This is what Bill Thompson calls "painting the target around the arrow (matching profile)…"

Thompson, W.C. (2009) Painting the target around the matching profile: the Texas sharpshooter fallacy in forensic DNA interpretation. *Law, Probability and Risk* 8: 257-276

Value of Using a Profile Interpretation Worksheet

Example worksheet available at **http://www.cstl.nist.gov/strbase/mixture.htm**

PROFILE INTERPRETATION WORKSHEET
IDENTIFILER

PROFILE NAME: *Case Example #3*

ANALYST: *John Butler*

DATE: *11 October 2010*

MIXTURE: ■ yes □ no □ unsure

Analytical threshold: *30 RFU*

Stutter % used: *0% (filter turned-off)*

Stochastic threshold: *150 RFU*

Peak height ratio: *60%*

Comments: *low level DNA (125 pg)*

Allele and Locus Assessments

ID LOCUS	Alleles called	Alleles above Stochastic Threshold	Stutter or other peaks to consider	Possible allele dropout ? Y/N	Stochastic issues? (e.g., elevated stutter, PHR imbalance, drop-in, etc.) Y/N	Degradation / inhibition (obvious)? Y/N	If mixture, restricted genotypes can be used? Y/N	Can this locus be interpreted ? Y/N	Additional Comments
D8S1179	11,13,16	13	Maybe	Y	Y	N	N	N	

Make decisions on the evidentiary sample and document them prior to looking at the known(s) for comparison purposes

CPE/CPI (RMNE) Limitations

- A CPE/CPI approach assumes that all alleles are present (i.e., cannot handle allele drop-out)
- Thus, statistical analysis of low-level DNA CANNOT be correctly performed with a CPE/CPI approach because some alleles may be missing
- Charles Brenner in his AAFS 2011 talk addressed this issue
- Research is on-going to develop allele drop-out models and software to enable appropriate calculations

What to do with low level DNA mixtures?

- **German Stain Commission "Category C"** (Schneider et al. 2006, 2009)
  - Cannot perform stats because stochastic effects make it uncertain that all alleles are accounted for
- **ISFG Recommendations #8 & #9** (Gill et al. 2006)
  - Stochastic effects limit usefulness
- *Fundamentals of Forensic DNA Typing* (2010) Butler 3rd edition (volume 1), chapter 18
  - Don't go "outside the box" without supporting validation

ISFG Recommendations on Mixture Interpretation

http://www.isfg.org/Publication;Gill2006

1. The likelihood ratio (LR) is the preferred statistical method for mixtures over RMNE
2. Scientists should be trained in and use LRs
3. Methods to calculate LRs of mixtures are cited
4. Follow Clayton et al. (1998) guidelines when deducing component genotypes
5. Prosecution determines H_p and defense determines H_d and multiple propositions may be evaluated
6. When minor alleles are the same size as stutters of major alleles, then they are indistinguishable
7. Allele dropout to explain evidence can only be used with low signal data
8. No statistical interpretation should be performed on alleles below threshold
9. Stochastic effects limit usefulness of heterozygote balance and mixture proportion estimates with low level DNA

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

A Complexity/Uncertainty Threshold

New Scientist article (August 2010)

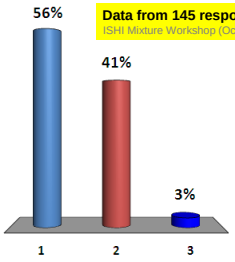
- **How DNA evidence creates victims of chance**
  - 18 August 2010 by Linda Geddes
- From the last paragraph:
  - **In really complex cases, analysts need to be able to draw a line** and say "This is just too complex, I can't make the call on it," says Butler. "Part of the challenge now, is that every lab has that line set at a different place. But the honest thing to do as a scientist is to say: **I'm not going to try to get something that won't be reliable**."

http://www.newscientist.com/article/mg20727743.300-how-dna-evidence-creates-victims-of-chance.html

Results from a Previous Training Workshop

Has your laboratory implemented a "stop testing" approach with complex and/or low-level DNA mixtures?

1. Yes
2. No
3. I don't work in a lab



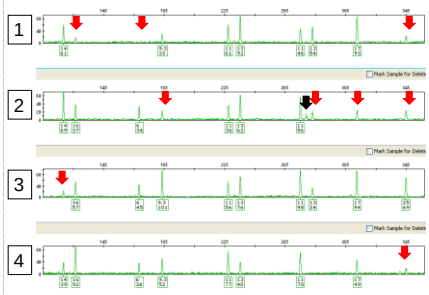
Data from 145 responses
ISHI Mixture Workshop (Oct 2011)

What "Stochastic" Means…

- Variability and allele dropout can occur anywhere in a DNA profile with low template DNA amounts…
- Peak height variability means that expected peak height ratios for paired alleles in heterozygotes quickly breaks down making mixture interpretation more challenging
- Confidence can be increased through replicate testing – but this requires splitting an already limited sample into smaller amounts

Stochastic Variation Observed

Same DNA – Amplified in Quadruplicate



Red arrows indicate allele drop-out (signal below analytical threshold)

Some observations

- in replicate #1 (top panel), lower size alleles drop-out (red arrows) more than larger size alleles
- variation exists between replicates: #3 and #4 had only a single missing allele while #2 is missing four alleles
- stutter peak (black arrow) in replicate #2 is almost as high as the second allele

Summary

- Do not blindly use a stochastic threshold with complex mixtures as assumptions regarding the number of contributors can impact interpretation
- Going back to try and get a better sample from the evidence (if available) is wiser than spending a lot of time trying to work with a poor quality DNA result

Future of Complex, Low-level Mixtures

- **If you want to work in this area, you need supporting validation data** (collecting a few results at high DNA levels and extrapolating to greater complexity and smaller amounts of DNA will not be sufficient)
- Recent efforts are focused on **modeling uncertainty through probabilistic genotype approaches**
- Will require software to perform all of the calculations
- See articles included in STRBase mixture section literature listing: http://www.cstl.nist.gov/strbase/mixture.htm

December 2012 Issue of *FSI Genetics*

Contents lists available at SciVerse ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Editorial

Focus issue—Analysis and biostatistical interpretation of complex and low template DNA samples

Contents lists available at SciVerse ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

DNA commission of the International Society of Forensic Genetics: Recommendations on the evaluation of STR typing results that may include drop-out and/or drop-in using probabilistic methods

P. Gill [a,b,*], L. Gusmão [c], H. Haned [d], W.R. Mayr [e], N. Morling [f], W. Parson [g], L. Prieto [h], M. Prinz [i], H. Schneider [j], P.M. Schneider [k], B.S. Weir [l]

Some of the articles present in this issue…

Contents lists available at SciVerse ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Exploratory data analysis for the interpretation of low template DNA mixtures

H. Haned [a,*], K. Slooten [a,b], P. Gill [c,d]

Contents lists available at SciVerse ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Validation of a DNA mixture statistics tool incorporating allelic drop-out and drop-in

Adele A. Mitchell *, Jeannie Tamariz, Kathleen O'Connell, Nubia Ducasse, Zoran Budimlija, Mechthild Prinz, Theresa Caragine

A Statistical Modeling Approach

Kelly, H., et al. (2012). The interpretation of low level DNA mixtures. *Forensic Science International: Genetics, 6(2),* 191-197

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

The interpretation of low level DNA mixtures

Hannah Kelly [a,*], Jo-Anne Bright [a], James Curran [b], John Buckleton [a]

Development of statistical models that account for the possibility of allele drop-out

A Simulation Approach

Forensic Science International: Genetics 5 (2011) 525–531

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Estimating drop-out probabilities in forensic DNA samples: A simulation approach to evaluate different models

H. Haned [a,*], T. Egeland [b], D. Pontier [a], L. Pène [c], P. Gill [b,d]

A Logistic Regression Model

Forensic Science International: Genetics 3 (2009) 222–226

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Estimating the probability of allelic drop-out of STR alleles in forensic genetics

Torben Tvedebrink [a,*], Poul Svante Eriksen [a,1], Helle Smidt Mogensen [b,2], Niels Morling [b,3]

Statistical model for degraded DNA samples and adjusted probabilities for allelic drop-out

Torben Tvedebrink [a,*], Poul Svante Eriksen [a,1], Helle Smidt Mogensen [b,2], Niels Morling [b,3]

Allelic drop-out probabilities estimated by logistic regression—Further considerations and practical implementation

Torben Tvedebrink [a,*], Poul Svante Eriksen [a,1], Maria Asplund [b,2], Helle Smidt Mogensen [b,3], Niels Morling [b,4]

A Logistic Regression Model

ARTICLE IN PRESS

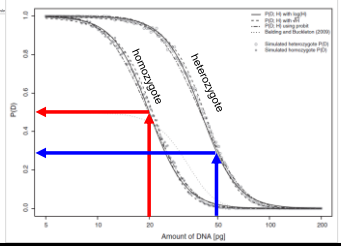
Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Allelic drop-out probabilities estimated by logistic regression—Further considerations and practical implementation

At **20 pg**, approximately **50% of homozygote** alleles will have dropped out

At **50 pg**, approximately **30% of heterozygote** alleles will have dropped out



Validation Analogy

- Validation studies can be compared to efforts involved in learning to drive a car properly
- My 16-year old daughter recently obtained her driving permit and is learning how to drive
- Age thresholds must be passed before someone can be considered for a driving permit and license
- The ultimate success of obtaining a driver's license and staying accident-free is based on training and preparation

Acquiring a Maryland Driver's License

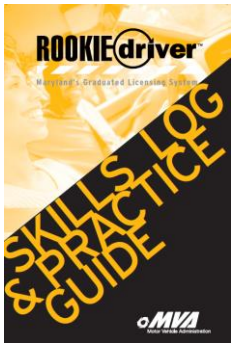
- A knowledge test must first be passed to be eligible

Allele peaks must first be observed to be interpreted…

- Three Stages for Rookie Drivers:
  1) Learner's Permit
     - Minimum age: 15 years 9 months old
     - Drives only with a qualified supervising driver
     - Must complete 60 hours of supervised driving experience
  2) Provisional License
     - Minimum age: 16 years 6 months old
  3) Full Driver's License
     - Minimum age: 18 years old

http://www.mva.maryland.gov/Driver-Services/RookieDriver/bgeneral.htm

Requirements for New Maryland Drivers



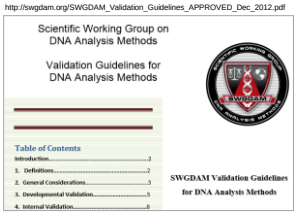
http://www.mva.maryland.gov/Resources/RD-006.pdf

New motor vehicle drivers (under 25 years old) must have:

- **60 hours of supervised driving experience** of which **10 hours** must be done **at nighttime**
- Must hold their **learner's permit for a minimum of 9 months**

New SWGDAM Validation Guidelines (2012)

http://swgdam.org/SWGDAM_Validation_Guidelines_APPROVED_Dec_2012.pdf


Scientific Working Group on DNA Analysis Methods

Validation Guidelines for DNA Analysis Methods

Table of Contents

SWGDAM Validation Guidelines for DNA Analysis Methods

Available on SWGDAM website: www.swgdam.org

"Each laboratory seeking to evaluate a new system must determine **which validation studies are relevant to the methodology**, in the context of its application, and determine **the number of samples required to satisfy each study**."

Internal Validation Data Should Drive Laboratory Interpretation Guidelines

SWGDAM Validation Guidelines – Approved December 2012

2.2.2.2 Quality assurance parameters and interpretation guidelines shall be derived from internal validation studies. For example, lower template DNA may cause extreme heterozygote imbalance; as such, empirical heterozygote peak-height ratio data could be used to formulate mixture interpretation guidelines and determine the appropriate ratio by which two peaks are determined to be heterozygotes. In addition to establishing an analytical threshold, results from sensitivity studies could be used to determine the extent and parameters of quality control tests that reagents require prior to their being used in actual casework.

Appropriate Samples Need to Be Evaluated During Validation Studies

3.6 **Case-type samples:** The ability to obtain reliable results should be evaluated using samples that are representative of those typically encountered by the testing laboratory. Where appropriate, consistency of typing results should be demonstrated by comparing results from the previous procedures to those obtained using the new procedure.

3.8 **Mixture studies:** The ability to obtain reliable results from mixed-source samples should be determined. These studies will assist the laboratory to establish guidelines for mixture interpretation, which may include determination of the number of contributors to the mixture, determination of the major and minor contributor profiles, and contributor ratios or proportions.

Important Things to Keep in Mind When Conducting Validation Studies

- Validation should establish the limits of a technique – thus **test in appropriate ranges**
  - PHR (Hb) variation tested at 1 ng will not apply to <100 pg data due to inherent stochastic variation with lower levels of DNA template
- **Replicate testing of the same DNA template**, especially at low levels, helps establish limits of reproducibility
- **Use known DNA samples** so reliability of genotypes and full profiles can be assessed
  - In the case of mixtures, plan specific ratios to evaluate
- **Test multiple DNA templates** as the quantitation of a single sample may not be what you think it is…

Experiment – Do Not Extrapolate

- It is not possible to fully apply concepts from single-source or 2-person mixtures like PHRs to more complex mixtures due to allele stacking possibilities
- If three person mixtures are being encountered regularly in your laboratory, then three person validation studies should be performed with known samples
  - Results of the validation study should be used to shape interpretation protocols
  - Establish the limits of reliable performance and stay within them (i.e., keep your car on the road)

Evaluate Reliability After Establishing Interpretation Guidelines

- Following validation experiments and establishment of specific parameters in the lab SOPs, challenge the new interpretation protocol with known samples to see if reliable results are obtained
  - For example, if the heterozygote peak height ratio has been set at 60%, then test multiple 2-person and 3-person mixtures with known genotypes and determine if reliable profiles can be deduced
  - **If an interpretation SOP does not work with known samples, how can it be expected to work reliably with casework samples?**

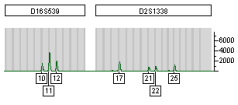
From Maryland Rookie Driver Information

- "…Recording each driving and practice experience is an easy way to track the progress of the new driver. **Each practice experience should be planned and present challenges for the new driver. Simply having the new drivers drive around the neighborhood will not prepare them for the time when they have a license and are driving without a supervisor**. Take the time to make your new driver the best possible driver they can be."

http://www.mva.maryland.gov/Resources/RD-006.pdf

Validation Studies Should Correspond to Needed Levels of DNA Interpretation

Easy drive around the neighborhood



- **Are your laboratory validation studies like a simple "drive around the neighborhood" of DNA testing?**
  - If the mixture portion of your validation studies involved mixing 9947A and 9948 in five different mixture ratios (e.g., 1:9, 1:3, 1:1, 3:1, & 9:1), then perhaps you should explore some more difficult scenarios as real-world casework is more complicated!

DNA Validation Should Prepare for Casework Situations to **Help Understand Limitations and to Develop Interpretation Protocols**

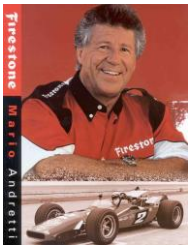
- **"Each practice experience should be planned and present challenges for the new driver…"** (Maryland Rookie Driver information)

Coping with >2 contributors

Under pressure with a "speed" case

Knowledge Obtained from Validation Studies Should Shape Interpretation SOPs and Benefit the Quality of Future Work

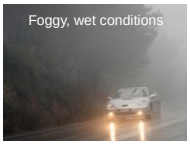
- "**Take the time** to make your new driver **the best possible driver they can be**…" (Maryland Rookie Driver information)



Want to avoid accidents!

There are times when you should slow down or perhaps not drive at all…

Poor Quality Conditions



Large Numbers of Contributors

NIST National Institute of Standards and Technology • U.S. Department of Commerce

FORENSICS@NIST

DNA Mixture Interpretation Webcast
April 12, 2013

http://www.nist.gov/oles/forensics/dna-analyst-training-on-mixture-interpretation.cfm

http://www.cstl.nist.gov/strbase/mixture.htm

Mixtures Go to Court

Robin Cotton

BOSTON UNIVERSITY

Testimony

- Why is it hard
  - Mixtures and the related scientific questions can be complicated
  - Court testimony can be challenging in many circumstances
- What makes it easier
  - Understanding your role
  - Scientific knowledge
  - Preparation

Review of Roles: the Prosecutor

- Is a representative of the government having justice as the main interest
- Must prosecute within the bounds of the law
- Ensure that the government's evidence is probative and reliable
- Has a duty to provide to the defense any exculpatory material

ABA **Standard 3-3.3 Relations With Expert Witnesses**

- A prosecutor who engages an expert for an opinion should respect the independence of the expert and should not seek to dictate the formation of the expert's opinion on the subject.
- To the extent necessary, the prosecutor should explain to the expert his or her role in the trial as an impartial expert called to aid the fact finders and the manner in which the examination of witnesses is conducted.

http://www.americanbar.org/publications/criminal_justice_section_archive/crimjust_standards_pfunc_blk.html#3.3

Review of Roles: the Defense Attorney

- Be a zealous advocate of the client within the bounds of the law
- Insures that the defendant's rights are protected
  - Interpose the defendant's constitutional rights against overreaching by the government
  - Duty to obtain all relevant and material discovery and disclosure of exculpatory information
  - Expose through cross examination the weaknesses of the testimony of government witnesses

Standard 4- 4.4 Relations With Expert Witnesses

- Defense counsel who engages an expert for an opinion should respect the independence of the expert and should not seek to dictate the formation of the expert's opinion on the subject.
- To the extent necessary, defense counsel should explain to the expert his or her role in the trial as an impartial witness called to aid the fact finders and the manner in which the examination of witnesses is conducted.

http://www.americanbar.org/publications/criminal_justice_section_archive/crimjust_standards_dfunc_blk.html#4.4

This means:

- Attorneys have an obligation to facilitate your testimony which will provide, among other things, your unbiased expert opinion.
- You are not on anyone's side or part of the prosecution or defense "team".
- The trial outcome is not your responsibility.

Our Role as Expert Witnesses is Different from that of Other Participants

- The expert witness:
  - **As a neutral participant** - presents objective opinions based on sound **Scientific Principles** correctly applied to question before the court.
  - has special knowledge or skill gained by education, training or experience which is beyond that of an ordinary person in a field applicable to the case before the court
  - is allowed to give opinion evidence based on the expertise of the witness

What is different about testimony related to a mixture? IT'S HARDER!

- The results are likely to be more complicated than for a single source profile
- You may need to explain one or more of the following
  - How you *know* a profile is a mixture
  - Why you cannot be certain of the number of contributors
  - How are you able to deduce the profile of a second contributor by assuming the presence of a known person
  - Why is the inclusion not an identification
  - Why are some results inconclusive
  - What is the Combined Probability of Inclusion
  - What is a likelihood ratio
  - What is a threshold: analytical, stochastic
  - What is a major contributor
  - What is an indistinguishable mixture
  - What does "polymorphism" mean

The solution: BE PREPARED!!

- Good preparation is essential for good testimony
- Both:
  - Your preparation
  - Preparation with the attorney who will present your direct testimony

Consider the following question and possible answers:

Question: How do you *know* the profile contains a mixture?

Correct answers:

1. There are more than two alleles per locus
2. Many peak height ratios are < 50%
3. Peak heights at amelogenin indicate a mixture

Do these work as expert witness answers?

-or-

Question: Please explain allele drop out?

Answer: Well…..(long pause)

How do you bridge the gap between what you know and what you can say to answer this question that is *understandable* to a juror?

or Can you explain DNA mixtures to a 5th grader?

Maybe you can't explain DNA mixtures to a 5th grader, but you can explain them to a 10th grader!!

the GAP is bridged by:

*A very careful translation which you can construct, and practice, for **any question** you may be think will be difficult to explain.*

1. Define what is the ***minimum*** number of concepts that are needed to answer the question
   - Make the list and be ruthless in removing unnecessary information
2. In what order would you present these concepts to make the most sense
   - Order the list
3. What is the simplest translation from how you would explain these concepts to a laboratory colleague to how would you say them to a 10th Grader?
4. Write out the explanation in plain English

Question: Please explain allele drop-out?
Answer: Well…..(long pause)

- Even though our methods are sensitive it is possible to have less DNA obtained from a sample that you really need. When this happens the PCR reaction may, by chance, make fewer copies of one allele at a locus that the other. This results in the signal from one allele being less than the signal from the other allele and sometimes signal from an allele becomes so low that it is not detected. This loss of signal is called allele drop-out.

In summary: construct the following

- What would you say scientifically?
- What parts of the description are *essential* to understanding?
- Eliminate the unnecessary concepts
- Substitute common words for scientific terms
- Practice and practice again! (with a live audience)

Preparation with Attorney

- Discuss case results, statistics, discovery with attorney
- Explain the results and conclusions
- Be sure that the attorney understands what you will and will not be saying about the conclusions
  - Does your testimony fit with what the attorney thought you were going to say?
- Explain limitations of your testimony
  - Your areas of expertise
  - Limitation of the data, report, conclusions

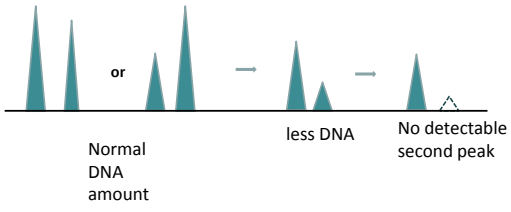
Preparation with Attorney

- Explain all issues and problem areas, related to the case, lab or yourself
  - Typos, strike outs, other small boo boos
  - Any testing irregularities with controls, contamination etc.
- NO SURPRISES-Attorneys do not like surprises
- Consider what may be asked in cross exam questions and plan for re-direct

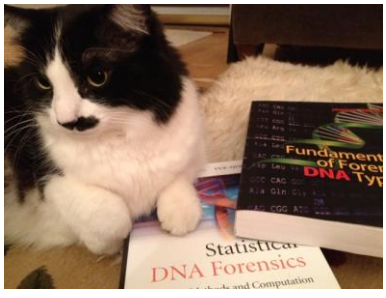
Preparation with Attorney: Materials and Exhibits

- Has the attorney prepared any charts or other visual aids? *(These may be more creative than you anticipated)*
- Is the information on these items accurate?
- Let the attorney know if you need paper and easel. You may want to teach something
- Consider whether drawing a diagram would help with your explanation of drop out?

Would this help the jury to visualize allele drop out?



Your Preparation-Plan a nice outfit and study hard



Your Preparation

- Review case carefully with the goal of deciding: How can the information in case be best presented?
  - Do a complete new technical review
  - Review SOP, validation data or any other documents
  - Outline complicated information
  - Critically review the case data and report(s)
    - What issues do you find?
    - What would you address or challenge if consulting for opposing counsel?

In Court

- Be honest in all answers no matter how difficult or uncomfortable this may be
  - You may be aware that the honest answer assists the case of the opposing attorney
- Treat all parties with respect all the time
  - Demeanor and tone is the same regardless of who asks a question
- You are the face of your organization during testimony

Get Comfortable with "Uncertainty"

- There will be some degree of uncertainty in
  - The number and ratio of contributors
  - Whether all alleles are present
  - The genotypes of the contributors
  - The strength of the conclusion
- Explain why it is not possible to know the TRUE answer
  - Admit other possibilities exist and state/quantitate likelihood
  - Exceptions become important when more probable

Use precise language in reports and in testimony-

- Be clear what you know about the number of contributors
  - Validate a properly defined analytical threshold
  - While "two or more contributors" includes the possibility of three or more contributors
  - Be precise and state if the number of alleles indicates "three or more contributors"

Use precise language in reports and in testimony-

- What constitutes a DNA profile
  - One peak
  - Two peaks at one locus
  - Peaks at more than one locus
- If you do not have a complete profile specify how many loci have data or refer to the table
- Do not refer to one peak as "the DNA profile obtained from the bloodstain…."
- If you have results at 6 loci you can say that

Statistics

- Be able to clearly state the question that is being answered with the statistic for the evidence
- Consider other relevant statistics which could be applied using a different method or different assumptions

Statistics

- Focus on the "commonness" or "rareness" of the profile
- Use likelihood ratios
- Clearly state that the numbers presented are "appoximate" and the true number would fall in some range around this estimate (based on population samples and Hardy-Weinberg assumptions)

Inconclusive

- Inability/failure to include or exclude
- Why were the results deemed uninterruptable or inconclusive? No DNA or
- Too little DNA
  - Cannot determine genotypes
  - Have a partial profile, alleles below stochastic threshold, missing alleles?
  - Too many contributors
  - QC problem, contamination,
  - Cannot do CPI
  - Cannot determine major/minor genotypes

Use precise language in reports and in testimony-especially with inconclusive results

- In weak or inconclusive result where genotypes cannot be unambiguously determined and the best statistical method is use of a likelihood ratio
- Do not use imprecise language such as
  - "His alleles are here
  - "the alleles come back to him"
- These types of statements made by a witness or an attorney are misleading

The need to use non-scientific terms does not mean you can be “loose” when stating results.

- Get out of the witness box and teach when you have to
- Be clear about how much data you have from a sample
- Results at 4 loci are not the same as results at 15
- Everyone can become a better witness

If you hear a Mistake, CORRECT IT!!

- If you realize you misspoke
- Attorney misstates your testimony in any way
- Attorney misstates your conclusion
- Attorney misrepresents the data or meaning of the statistic

You have a new SOP and an old report, what to do?

- Issue an amended report
- Science does not stand still and few people expect it too
- Your knowledge has increased and therefore your opinion has changed
- The new report will reflect the new opinion
- If reports are not affected by SOP changes then no action is needed

Clear Communications: the ethical and professionally responsible forensic scientist…

- Presents accurate and complete data in reports, testimony, publications and oral presentations

ASCLD/LAB Guiding Principles of Professional Responsibility for Crime Laboratories and Forensic Scientists; http://www.ascld-lab.org/about_us/guidingprinciples.html

Clear Communications: the ethical and professionally responsible forensic scientist…

- Testify to results obtained and conclusions reached only when they have confidence that the opinions are based on good scientific principles and methods. Opinions are to be stated so as to be clear in their meaning. Wording should not be such that inferences may be drawn which are not valid, or that slant the opinion to a particular direction.

ASCLD/LAB Guiding Principles of Professional Responsibility for Crime Laboratories and Forensic Scientists; http://www.ascld-lab.org/about_us/guidingprinciples.html

Clear Communications: the ethical and professionally responsible forensic scientist…

- Attempt to qualify their responses while testifying when asked a question with the requirement that a simple “yes” or “no” answer be given, if answering “yes” or “no” would be misleading to the jury.

ASCLD/LAB Guiding Principles of Professional Responsibility for Crime Laboratories and Forensic Scientists; http://www.ascld-lab.org/about_us/guidingprinciples.html

In a recent publication in: Behavioral Sciences and the Law (2010): The Witness Credibility Scale: an Outcome Measure for Expert Witness Research by S.L. Brodsky, et al.

- These 4 features of the expert witness, taken together, explain approximately 70% of the variance in ratings of the expert from the 264 test participants.

Characteristic	% Variance explained
Confident	50%
Likable	9%
Trustworthy	7%
Knowledgeable	5%

Brodsky, S.L., Griffin, M. P., Cramer, R.J. 2010 The Witness Credibility Scale: an Outcome Measure for Expert Witness Research, Behavioral Sciences and the Law, **28:** 892-907

Confidence in yourself and effective testimony comes from:

- What you know
  - Molecular biology, genetics, statistics applied to evaluate or provide weight to the data
  - Scientific literature
  - Validation data
  - Case results and conclusions
- Training and experience
- Your ability to communicate your answers effectively (i.e., in understandable language).

Confidence and effective testimony do *NOT* come from:

- Your SOP
- Your Technical Leader
- Your QA system
- Other lab policy
- You lab accreditation

- The jury can only see *you*. These other people or entities are not present for them to evaluate.

What is the effect of answering a question by referring to the SOP, technical leader, lab policy, etc.?

- Have you demonstrated true familiarity with the topic?
- Have you demonstrated you know the underlying answer?
- Do you sound well informed?

- The answer is likely to be NO to each of these questions

And finally; In Court

- Honesty is the only absolute requirement

  – Any other thing that goes wrong is repairable

" The right to search for the truth implies also a duty; one must not conceal any part of what one has recognized to be true."

Albert Einstein
1879-1953

NIST National Institute of Standards and Technology • U.S. Department of Commerce

DNA Mixture Interpretation Webcast
April 12, 2013

http://www.nist.gov/oles/forensics/dna-analyst-training-on-mixture-interpretation.cfm

http://www.cstl.nist.gov/strbase/mixture.htm

Lessons Learned, Recent Literature, and Future Directions

John M. Butler
National Institute of Standards and Technology

Comments on Mixture Training We Have Conducted The Past Three Years

- Trying to help analysts better understand the SWGDAM 2010 Interpretation Guidelines
  - It is important to note that **the 2010 SWGDAM Guidelines were written primarily for 2-person mixtures situations**
- However, **many labs are doing or attempting more complex mixtures often without appropriate underlying validation support** or consideration of complicating factors
- **The information content in our workshops has continued to evolve to include the latest published articles…**

Greg Matheson on Forensic Science Philosophy

The CAC News – 2nd Quarter 2012 – p. 6
"Generalist vs. Specialist: a Philosophical Approach"
http://www.cacnews.org/news/2ndq12.pdf

- If you want to be a technician, performing tests on requests, then just focus on the policies and procedures of your laboratory. If you want to be a scientist and a professional, learn the policies and procedures, but go much further and learn the philosophy of your profession. **Understand the importance of why things are done** the way they are done, the scientific method, the viewpoint of the critiques, the issues of bias and the importance of ethics.

My Goals in This Presentation

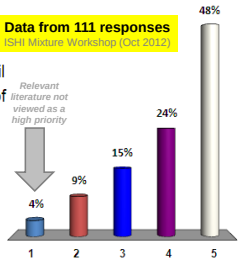
- Valuable mixture literature and how to obtain it
- Important lessons & common misunderstandings
- Thoughts on where we need to go as a community to improve mixture interpretation

2012 Response at ISHI Workshop

Which of the topics below would be your first choice for additional training?

1. Relevant literature
2. How to validate thresholds in more detail
3. Reporting and the use of assumptions
4. **Interpretation of low level mixtures**
5. **Likelihood ratios and other statistical approaches**

~75% want more information on these topics



Mixture Literature *you should be reading…*

See DNA Mixtures Reference List on STRBase mixture section

http://www.cstl.nist.gov/strbase/mixture.htm

Quality Assurance Standard Requirement for Literature Review

Quality Assurance Standards for Forensic DNA Testing Laboratories
(effective September 1, 2011)

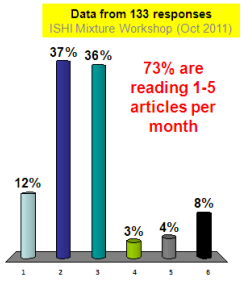
5.1.3.2. The laboratory shall have a program approved by the technical leader for the **annual review of scientific literature** that documents **the analysts' ongoing reading of scientific literature**. **The laboratory shall maintain or have physical or electronic access to a collection of current books, reviewed journals, or other literature applicable to DNA analysis**.

http://www.fbi.gov/about-us/lab/codis/qas-standards-for-forensic-dna-testing-laboratories-effective-9-1-2011

2011 Response at ISHI Workshop

How many DNA-related articles would you estimate that you read in a typical month?

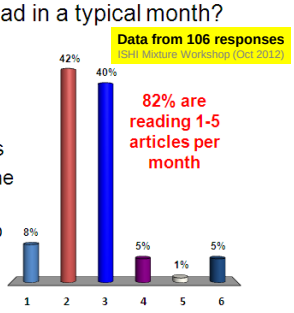
1. None
2. 1 article
3. 2 to 5 articles
4. More than 5 articles
5. None, I only read the abstracts
6. I don't make time to read!



2012 Response at ISHI Workshop

How many DNA-related articles would you estimate that you read in a typical month?

1. None
2. 1 article
3. 2 to 5 articles
4. More than 5 articles
5. None, I only read the abstracts
6. I don't make time to read!



Importance of Reading the Literature

How can you keep up and improve?

- Develop a culture in your laboratory to read the literature and share information with one another
- Obtain access to appropriate journals
  - Join AAFS and/or ISFG
  - Develop a relationship with a local university in order to get access to the latest journal articles
- Read, Think, and Implement Improvements!

Useful Articles on DNA Mixture Interpretation

- **Buckleton, J.S. and Curran, J.M. (2008) A discussion of the merits of random man not excluded and likelihood ratios. *Forensic Sci. Int. Genet.* 2: 343-348.**
- Budowle, B., *et al.* (2009) Mixture interpretation: defining the relevant features for guidelines for the assessment of mixed DNA profiles in forensic casework. *J. Forensic Sci.* 54: 810-821.
- Clayton, T.M., *et al.* (1998) Analysis and interpretation of mixed forensic stains using DNA STR profiling. *Forensic Sci. Int.* 91: 55-70.
- **Gill, P., *et al.* (2006) DNA commission of the International Society of Forensic Genetics: Recommendations on the interpretation of mixtures. *Forensic Sci. Int.* 160: 90-101.**
- Gill, P., *et al.* (2008) National recommendations of the technical UK DNA working group on mixture interpretation for the NDNAD and for court going purposes. *FSI Genetics* 2(1): 76–82.
- Schneider, P.M., *et al.* (2009) The German Stain Commission: recommendations for the interpretation of mixed stains. *Int. J. Legal Med.* 123: 1-5.

Read to Maintain a Big Picture View!

If you are not following the recent literature, you would have missed:

- Software applications & implementation
- Impact of allele dropout on stats
- Studies on number of contributors

- The literature is changing very fast
  - Read more than *Journal of Forensic Sciences* to stay caught up
- **Analysts need time to read and ask critical questions**

Number of Articles Published
on DNA and DNA Mixtures

http://www.ncbi.nlm.nih.gov/pubmed

Journal Name	"DNA"	"DNA mixtures"	"DNA mixtures" in 2012
Forensic Sci. Int. / FSI Genetics	1484	68	15
J. Forensic Sci.	1196	45	2
Int. J. Legal Med.	659	39	5
Croatian Med. J.	155	12	4
Science & Justice	73	5	0

PubMed.gov search conducted September 14, 2012 using "DNA" or "DNA mixtures" and journal name with and without "and 2012"

STRBase DNA Mixtures Reference List

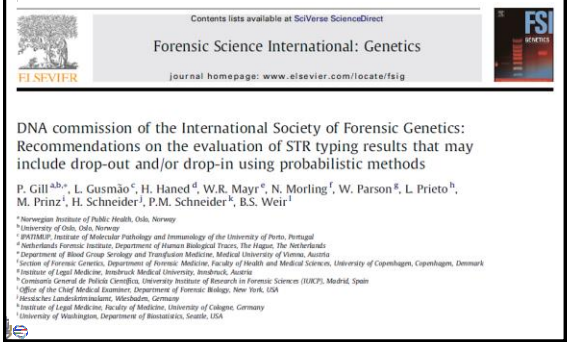
Topic category	# References
Mixture Principles & Recommendations	13
Setting Thresholds	11
Stutter Products & Peak Height Ratios	19
Stochastic Effects & Allele Dropout	18
Estimating the Number of Contributors	15
Mixture Ratios	9
Statistical Approaches	23
Low Template DNA Mixtures	8
Separating Cells to Avoid Mixtures	3
Software (plus 12 websites)	7
Probabilistic Genotyping Approach	11
General Information on Mixtures	7
TOTAL	144

7/8 in the past year; mostly in *FSI Genetics*

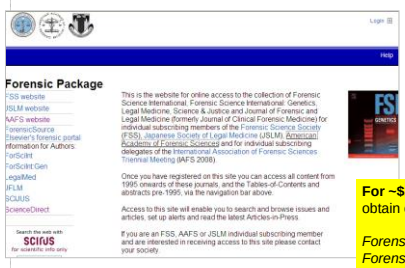
Will be regularly updated on **http://www.cstl.nist.gov/strbase/mixture.htm**

Recent articles on mixtures not found in JFS…

December 2012 Issue of *FSI Genetics*
is on **DNA Interpretation Challenges and Solutions**



Elsevier Journal Package
Available with AAFS Membership



http://www.sciencedirect.com/forpac

For ~$100 per year, you obtain electronic access to:

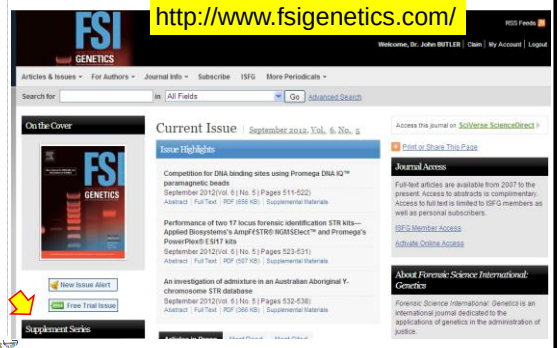
Forensic Sci Int: Genetics
Forensic Sci Int
Science & Justice
Legal Medicine
Forensic & Legal Medicine

Join ISFG and Receive FSI Genetics

http://www.isfg.org/Membership

Abstracts are Freely Available on Website

http://www.fsigenetics.com/



FSI Genetics Supplement Series Articles are Freely Available

Articles (2-3 pages each) covering presentations given at the ISFG meetings every two years

http://www.fsigeneticssup.com

2011: 281 articles
2009: 253 articles
2007: 272 articles

Know the Literature

- Sometimes articles may not be all that they claim to be – evaluate them critically
- Stay informed in order to be a good scientist
- **M**ixtures **U**sing *SOUND* **S**tatistics, **I**nterpretation, and **C**onclusions involves knowing the literature (past and present)

Mixtures **U**sing *SOUND* **S**tatistics, **I**nterpretation, & **C**onclusions

2012

Important Lessons

- People think they understand the basics of interpretation better than they actually do – this is what leads to observed variation in interpreting mixtures, which is typically due to using different subsets of the data and/or different assumptions
- Increased complexity of mixtures (with more allele sharing) leads to **higher uncertainty**, which leads to lack of confidence in potential contributor genotypes
- Worked examples are beneficial in training (participants need to work through the examples themselves)
- There is value in using a profile interpretation worksheet to document assumptions and decisions made

Value of Using a Profile Interpretation Worksheet

Example worksheet available at **http://www.cstl.nist.gov/strbase/mixture.htm**

PROFILE INTERPRETATION WORKSHEET
IDENTIFILER

PROFILE NAME: *Case Example #3*
ANALYST: *John Butler*
DATE: *11 October 2010*
MIXTURE: ■ yes □ no □ unsure

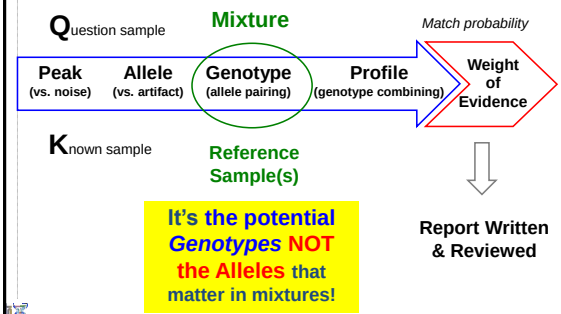
Analytical threshold: *30 RFU*
Stutter % used: *0% (filter turned-off)*
Stochastic threshold: *150 RFU*
Peak height ratio: *60%*
Comments: *low level DNA (125 pg)*

Allele and Locus Assessments

ID LOCUS	Alleles called	Alleles above Stochastic Threshold	Stutter or other peaks to consider	Possible allele dropout ? Y/N	Stochastic issues? (e.g., elevated stutter, PHR imbalance, drop-in, etc.) Y/N	Degradation Inhibition (obvious)? Y/N	If mixture, restricted genotypes can be used? Y/N	Can this locus be interpreted ? Y/N	Additional Comments
D8S1179	11,13,16	13	Maybe	Y	Y	N	N	N	

Make decisions on the evidentiary sample and document them prior to looking at the known(s) for comparison purposes

Steps in DNA Interpretation



It's the potential *Genotypes* NOT the Alleles that matter in mixtures!

Common Misunderstandings

- Using CPI stats is conservative to the defendant
  - The numerical stat is low but by throwing out information the ability to EXCLUDE innocent people is reduced
  - With PopStats, a single peak is calculated as p^2 (not 2p)
- Using CPI stats means that the potential number of contributors is not important
  - Higher numbers of contributors dilutes out the amount of DNA for each contributor which leads to more stochastic effects and the possibility of allele dropout (more uncertainty)
  - The CPI stat cannot handle allele dropout!

Handling Complex Mixtures

- Stochastic thresholds are necessary in combination with CPI statistics
  - but a stochastic threshold may not hold much meaning for >2 person mixtures (due to potential allele sharing)
- Most labs are not adequately equipped to cope with complex mixtures
  - Extrapolating validation studies from simple mixtures will not be enough to create appropriate interpretation SOPs

David Balding (UK professor of statistical genetics): "LTDNA cases are coming to court **with limited abilities for sound interpretation**." (Rome, April 2012 meeting)

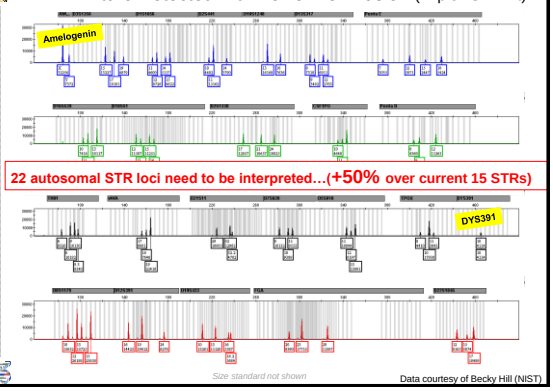
Thoughts on Where We Need to Go (1)

- Away from CPI and towards likelihood ratio approaches
  - As noted in the Gill et al. (2006) ISFG DNA Commission recommendation #2
- This will require software to perform the calculations
  - This software will need to be validated
  - Peter Gill and others are pushing freeware solutions
- Still will require analysts to understand what is going on in the computer calculations!
  - Will require more significant engagement in mixture training

Thoughts on Where We Need to Go (2)

- Validation studies need to support interpretation SOPs and software packages
- The U.S. will be moving to more STR loci in the near future (from 13 to ~20 core STRs)
  - Using additional loci with better powers of discrimination will improve detection of mixtures
  - **But more loci means more interpretation time!**

DNA Mixture Detected with PowerPlex Fusion (24plex STR kit)



Webcast Format for Training

- **With cuts in federal budgets, webcasts or webinars may become more appealing in the future to reduce costs in providing training**
- Please let us know about any technical difficulties that you may have faced so that we can improve future webcasts
- We welcome suggestions for additional content or topics to cover in future webcast training events
- Please contact John Paul Jones at 301-975-2782 or john.jones@nist.gov

Posting of Video from this Event

- Following transcription of this webcast (this process takes about a month), **we plan to post videos of each presentation on a publicly-available NIST website**
- All those who registered for the webcast (onsite or online) will receive email notification of this website URL
- A link to the webcast video website will also be available from the STRBase mixture website to enable future viewing or downloading of video or presentation materials
- Due to costs of maintaining large video files on NIST servers, **webcast videos may only be available for a limited time** (we are planning on at least six months)

Concern for Potential Misuse of Webcast Presentations

- We remind current and future viewers that presentations reflect the presenters' opinions at the time they were given on April 12, 2013
- Please do not take any specific comments of the webcast presenters out of context in order to advance either scientific or legal arguments
- Science advances with new discoveries and therefore scientific opinions may change over time given exposure to new ideas or techniques

Acknowledgments

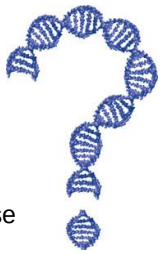
- **Fellow presenters: Mike, Robin, Bruce, and Charlotte**
- **NIST A/V and webcast staff**
- Past and present funding to the NIST Applied Genetics Group from the **National Institute of Justice** (NIJ) and the **NIST Law Enforcement Standards Office** (OLES)
- NIJ Forensic Science Training Development and Delivery Program Grant (2008-DN-BX-K158) to the Biomedical Forensic Science Program at **Boston University** School of Medicine
- **Catherine Grgicak** (BU Program) for mixture data
- **Maryland State Police Forensic Sciences Division** for permitting involvement of Bruce Heidebrecht
- **John Paul Jones** from NIST OLES for organizing and coordinating this event

Thank you for your attention

Contact Information

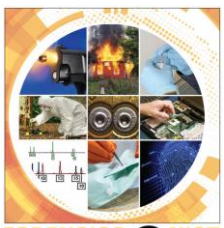
John M. Butler
NIST Fellow
john.butler@nist.gov
301-975-4049

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



Additional DNA mixture information available at:
http://www.cstl.nist.gov/strbase/mixture.htm

NIST FORENSIC SCIENCES



FORENSICS @ NIST

NIST Applied Genetics

BOSTON UNIVERSITY