Thoughts on Mixture Interpretation Issues

Facing the Forensic DNA Community

(not just Texas!)

John M. Butler, Ph.D.

National Institute of Standards and Technology

DNA Mixture Interpretation Panel Meeting (Dallas, TX, September 18, 2015)

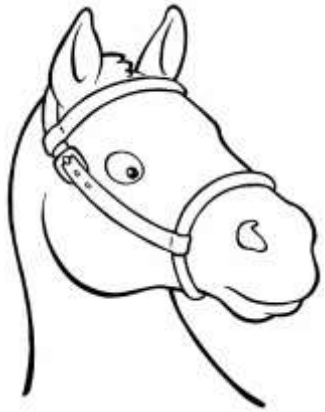
TX FSC August 21, 2015 Notification Letter

1. **Allele frequency corrections**
   - **Minor changes in values used in statistical calculations**
   - **Will depend on the alleles present in the specific case**
   - For example with FBI Caucasians, D8S1179 allele 11 (0.0587), allele 13 (0.3393), allele 16 (0.0128) there were no changes to allele frequency values

2. **DNA mixture interpretation changes**
   - Application of a new stochastic threshold in conjunction with a CPI calculation may lead to a removal of loci from consideration
   - With being more “conservative” and considering less information across a DNA mixture profile, statistical results **may** drop dramatically or even **go from an inclusion to inconclusive**

Illustration of Issues Involved

FBI Allele Frequency Corrections
(like changing the color of a horse)



http://www.malvorlagengratis.net/uploads/pferde-11.jpg

https://s-media-cache-ak0.pinimg.com/736x/5e/a2/d0/5ea2d0780fc4d639b6de094032693d7b.jpg

Changing DNA Mixture Interpretation Protocols
(like changing the transportation vehicle)

http://atlanticautotint.com/wp-content/uploads/2011/09/truck-fender-flares-dodge.jpg

CPI Approach to DNA Mixtures

Requires new skills & thinking

Probabilistic Genotyping

The same data can be perceived differently

what we “see” (interpret) depends on our “prescription”

(perspective, training, model used to evaluate information, etc.)

A colleague’s comments: You realize someone was reading a document two years ago with the wrong prescription glasses, so you give them new glasses today based on what they should have had two years ago. Their prescription has changed in that intervening time so you haven’t fixed the problem. **You need to assess the current status before you take any corrective action.**

Different Thresholds Used with CE Data

Example values
(empirically determined based on own internal validation)

Peak real, can be used for CPI

150 RFUs — **Stochastic Threshold (ST)**

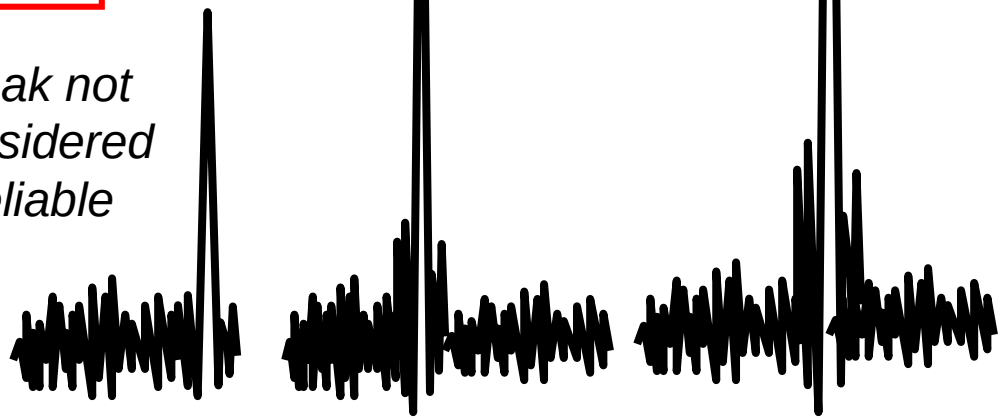
(Match Interpretation/ Dropout/Reporting)

Peak real, but not used for CPI

50 RFUs — **Analytical Threshold (AT)**

(Reporting/Noise Limit-of-Detection)

Peak not considered reliable



Noise

Stochastic thresholds are not a new idea

- Cetus 1992 article (Walsh et al. *PCR Methods Appl.* 1: 241-250)
  - "Preferential amplification due to stochastic fluctuation can occur when amplifying very low amounts of target DNA molecules. … This problem can be avoided by adjusting the cycle number such that approximately 20 or more copies of target DNA are required to give a typing result for that PCR system." **[this is why STR kit cycle numbers are usually set to 28 cycles by manufacturers in order to limit detection of full profiles to ~125 pg]**

- FBI 1995 PM and DQα validation (Budowle et al. *JFS* 40: 45-54)
  - "The S dot from the PM typing strip can be used to evaluate whether or not stochastic effects should be considered" **[the "S" stands for "stochastic"]**

- FBI 2001 article (Moretti et al. *JFS* 46: 647-660)
  - "When few copies of the DNA template are present, stochastic amplification may occur, … [see next slide for further quote]

Quote from Moretti et al. 2001 JFS 46: 647-660

Validation of short tandem repeats (STRs) for forensic usage: performance testing of fluorescent multiplex STR systems and analysis of authentic and simulated forensic samples

"When few copies of the DNA template are present, stochastic amplification may occur, resulting in either a substantial imbalance of two alleles at a given heterozygous locus or allelic dropout. Therefore, the amount of DNA used in the PCR can have an impact on stochastic effects. The reverse dot blot systems (AmpliType PM and DQA1+PM, Applied Biosystems) include a means (e.g., the "S" or "C" dots) of evaluating whether a DNA template used in the PCR is above the level at which stochastic effects may impact on the relative yield of two alleles at a given heterozygous locus. Similarly, peak heights can serve as the equivalent of a stochastic control for STR typing. **The quality control measure for an effective stochastic interpretation threshold should be developed based on a minimum peak height value.** This minimum threshold should be determined in-house because of variation in DNA quantitation efficiency and sensitivity of detection of analytical instruments. **Peaks with heights below the threshold should be interpreted with caution.** Finally, because of the possibility of stochastic effects on amplification when analyzing low copy number DNA templates, caution should be used in modifying the thermocycling parameters (e.g., using additional cycles) and electrophoretic conditions (e.g., increasing the injection time during capillary electrophoresis) to enhance product intensity."

This FBI validation article was written in 1999 (submitted to the *Journal of Forensic Sciences* on July 29, 1999 but not published until May 2001)

SWGDAM "Retroactive" Statement

QUESTION: Within many of the SWGDAM guidelines the statement is made that these guidelines are not intended to be used retroactively. What is the intent of this "retroactive" statement?

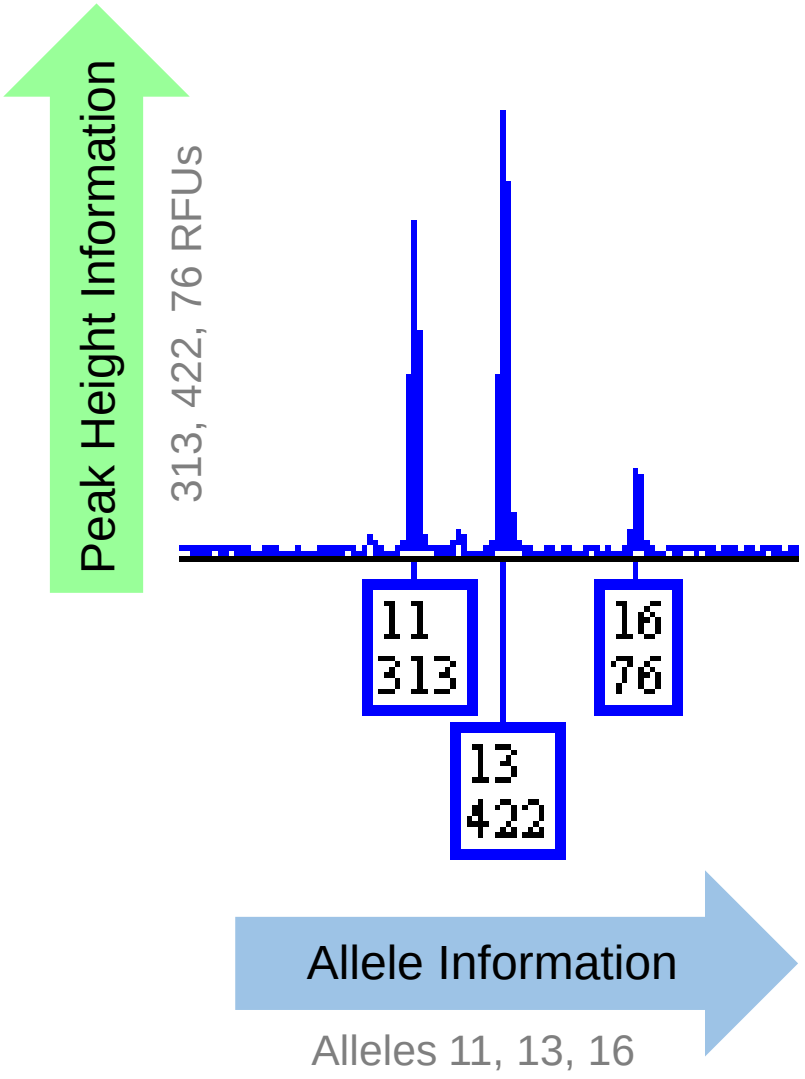
SWGDAM Response: SWGDAM includes a "retroactive" statement with the intent that the revised guidance be applied prospectively and not retroactively. **With the underlying assumption that work (validation, training, analysis, interpretation) performed prior to the issuance of the revisions <u>was appropriate and scientifically valid</u>**, revision of the applicable guidelines is not intended to invalidate or call into question the previous work.

http://www.swgdam.org/#!faq/cqu4

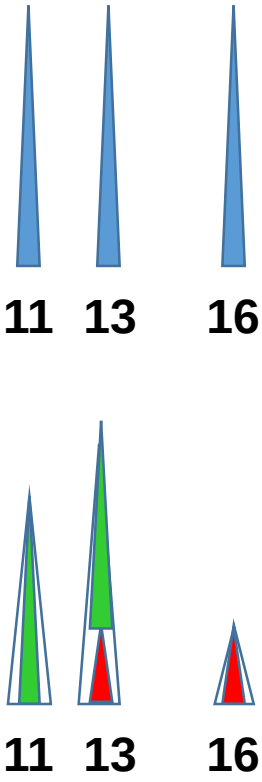
FREQUENTLY ASKED QUESTIONS

On occasion, SWGDAM will use this page to post responses to frequently asked questions from the forensic DNA community or other interested parties for thepurposes of general information. The intent of this page is not for it to be acomprehensive list of answers to all of the inquires SWGDAM receives, but rather a collection of those inquires that SWGDAM recognizes to be of interest to a broad spectrum of forensic DNA science practitioners and/or consumers.

Mixture Data Observed at D8S1179



With CPI statistics, peak height information is ignored (calculations would be the same if all peaks were of equal height). Because **all genotype combinations are considered equally probable**, information from the profile is not used optimally

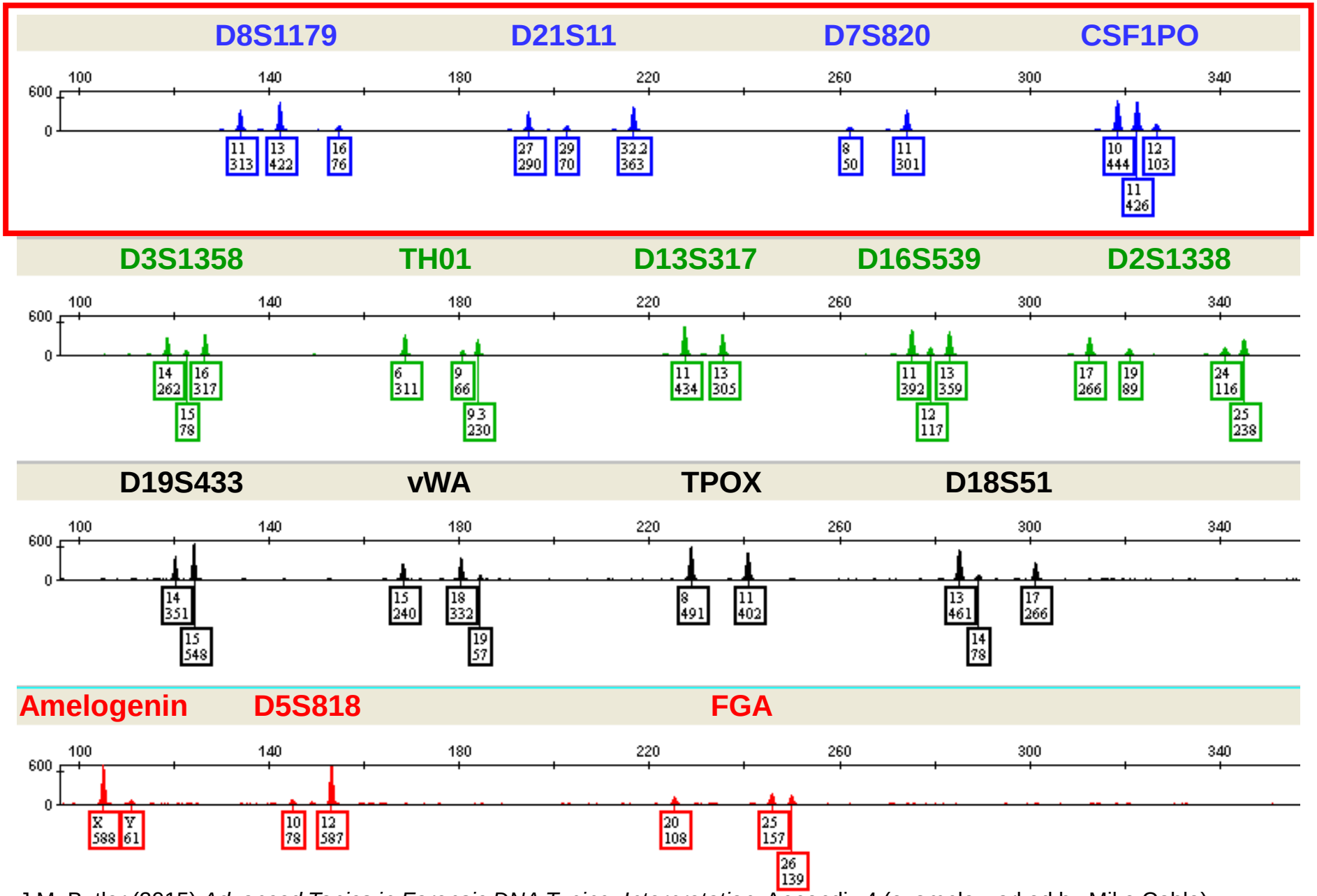


Most logical combination

11,13 major

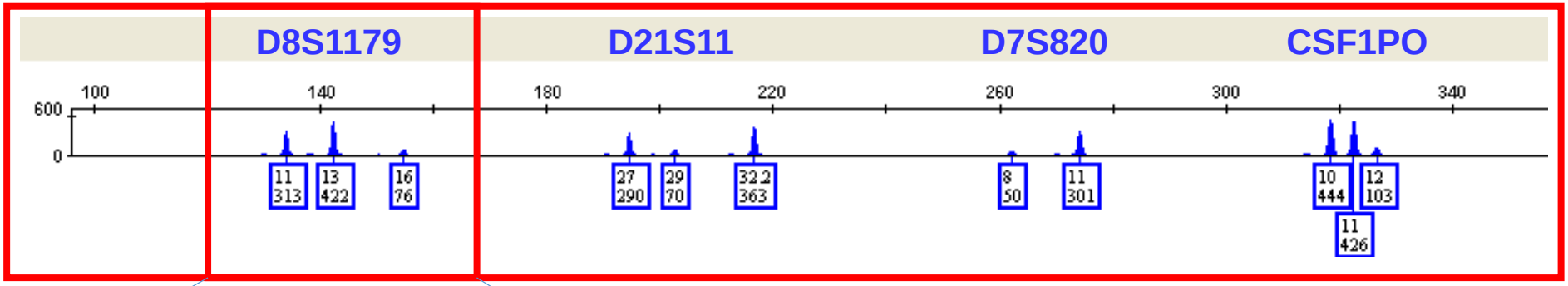
13,16 minor

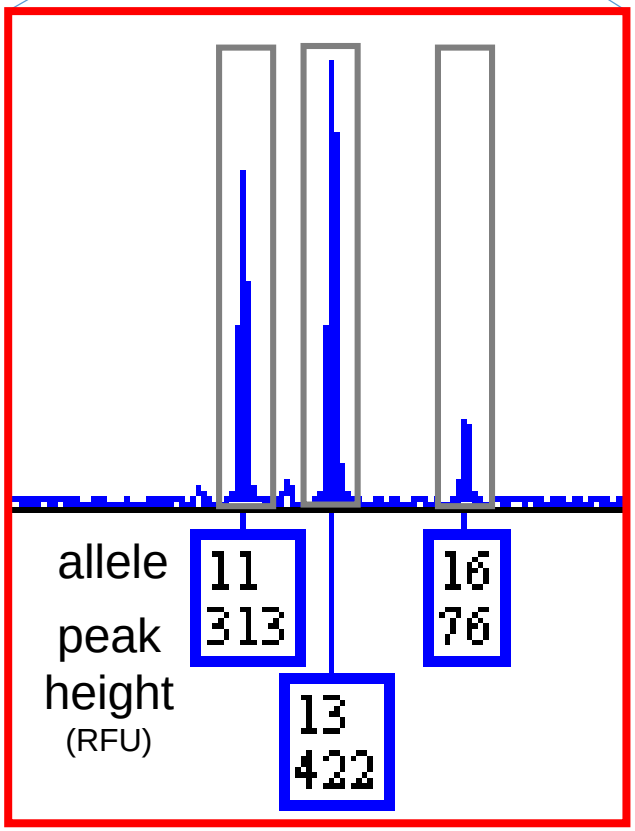
DNA Mixture Example



J.M. Butler (2015) *Advanced Topics in Forensic DNA Typing: Interpretation*, Appendix 4 (example worked by Mike Coble)

Just the top row of the Identifiler DNA mixture profile





3 alleles present

11 13 16

6 possible genotype combinations

(without considering peak heights)

11,11 or **13,13** or **16,16**

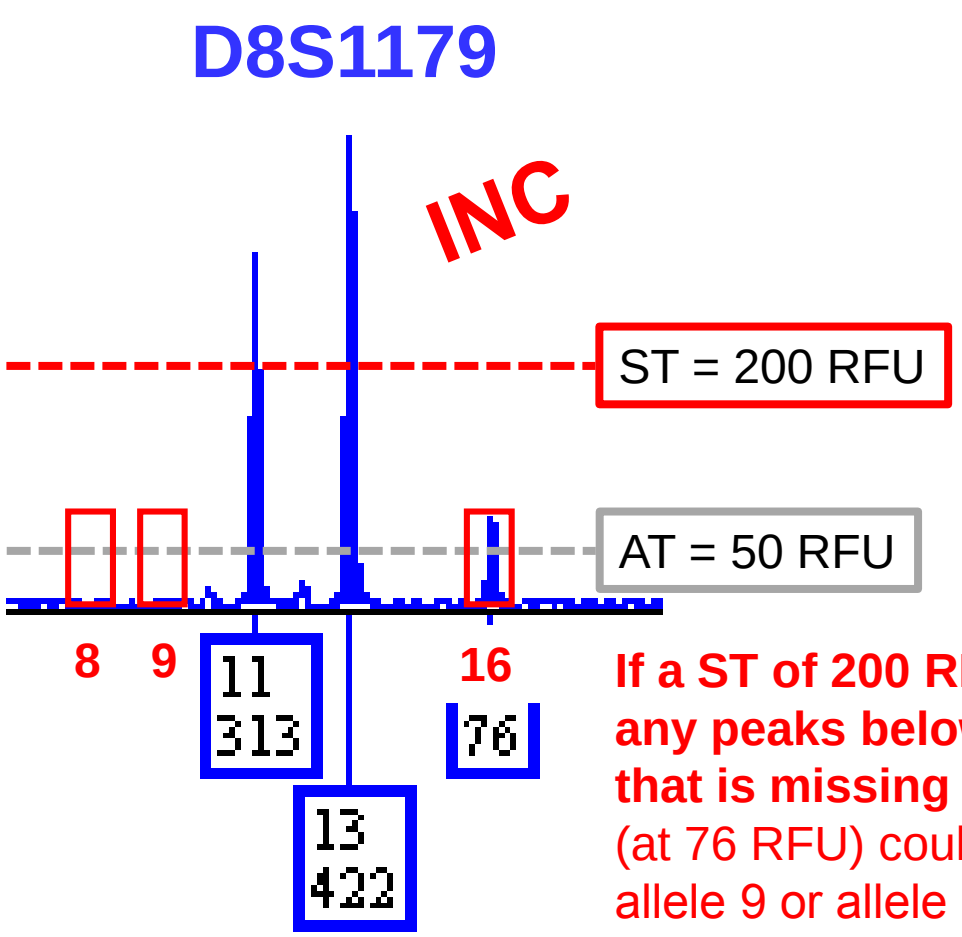
11,13 or **13,16** or

11,16

Applying a ST of 200 RFU when the allele 16 peak is below this value, leads to a CPI statistic of 1 in 1 for this locus. **Essentially this locus then becomes "inconclusive" (INC) – of no value in either including or excluding a suspect…**

3 alleles present

11 13 16



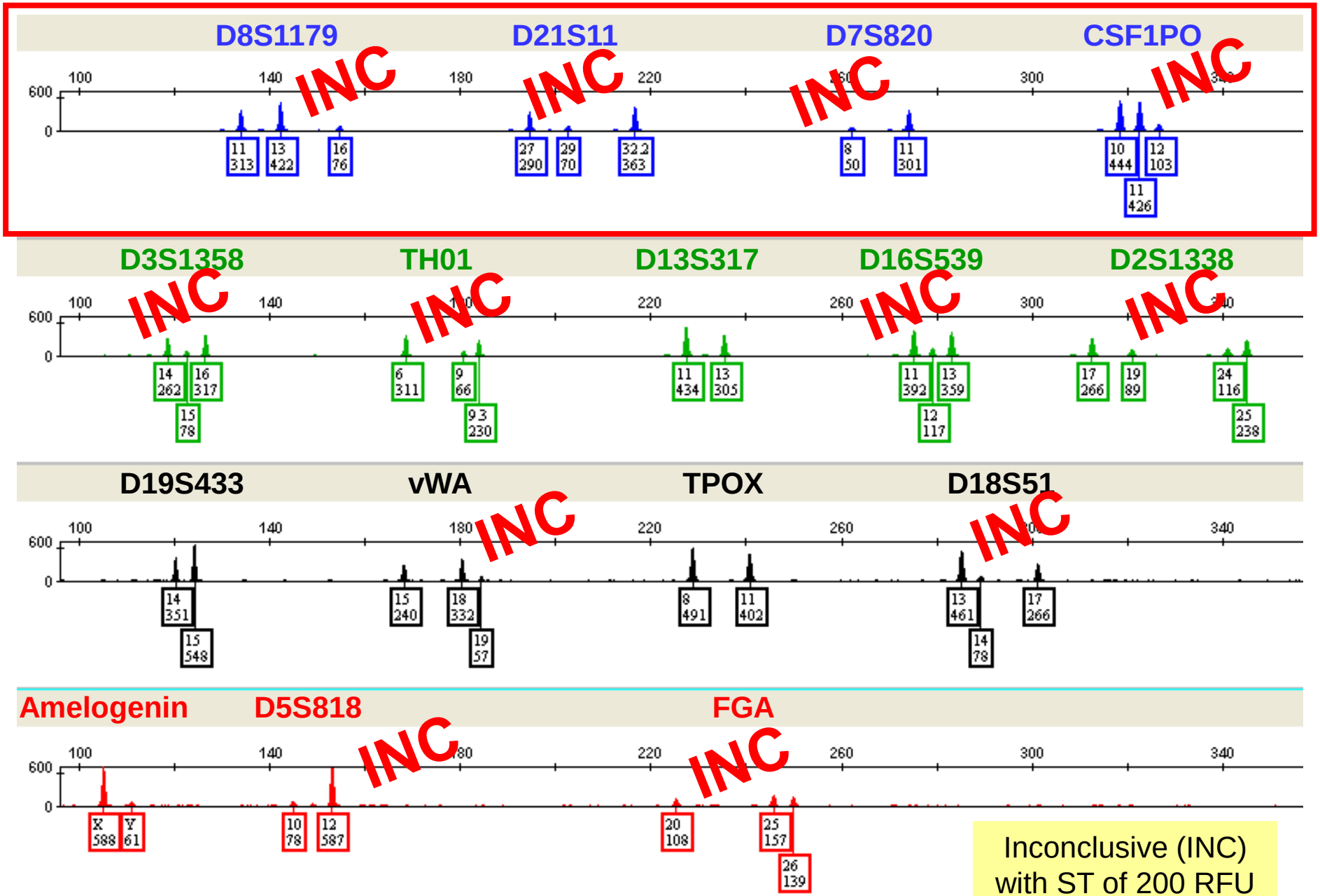
6 possible genotype combinations
(without considering peak heights)

11,11 or **13,13** or **16,16**
11,13 or **13,16** or
11,16 or **16, anything**

If a ST of 200 RFU is applied, then we assume that any peaks below ST could be paired with an allele that is missing due to allele drop-out. Thus, allele 16 (at 76 RFU) could be paired with a missing allele 8 or allele 9 or allele 14 or allele 15 or … Combining all of these possibilities leads to **an inclusion probability of 1 (i.e., anyone could be in the mixture at that locus)**.

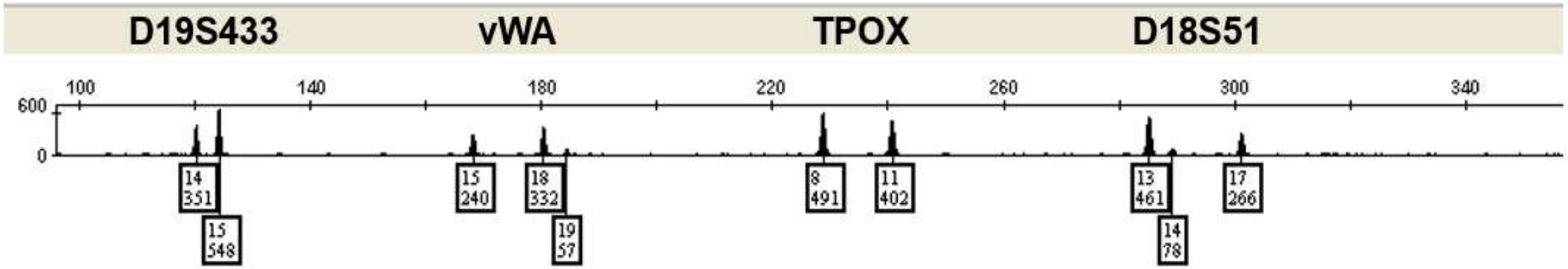
ST: stochastic threshold
AT: analytical threshold

DNA Mixture Example



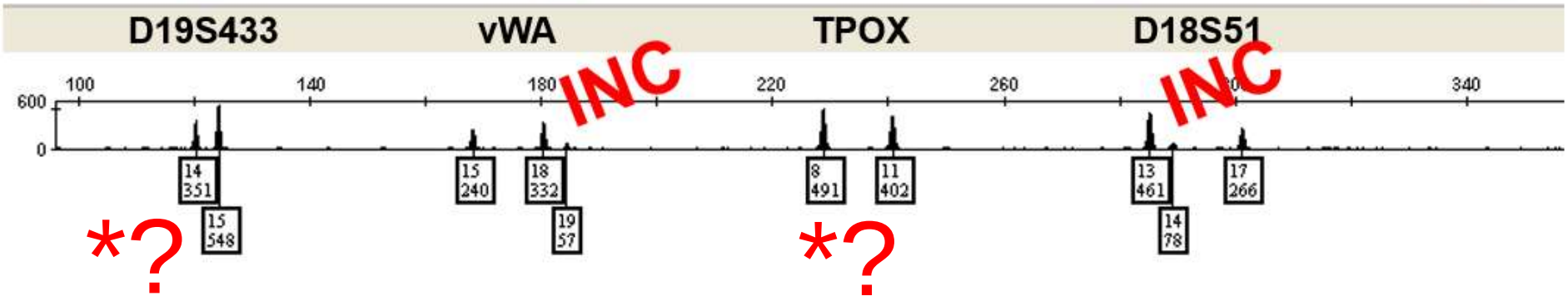
Inconclusive (INC) with ST of 200 RFU

Thresholds and Frequencies (CPI)

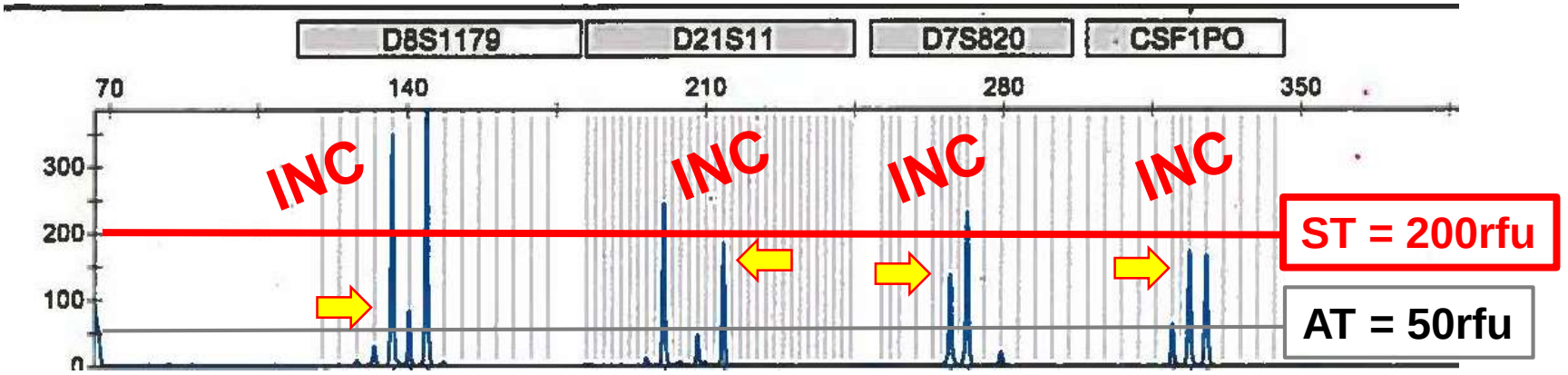


9 markers (17 alleles) affected by FBI allele frequency changes

	"Old" FBI Frequencies	"Revised" FBI Frequencies
CPI (AT = 50rfu; no ST)	1 in 1.40×10^9	1 in 1.39×10^9
CPI (AT = 50rfu; ST = 200rfu)*	1 in 38.6	1 in 37.7



Todd Bille (ATF Lab) Case Example



Blue channel of the Identifiler STR profile

	Interpretation Approach Used	Profile Probability (in FBI Caucasians)
Similar to 1999-2015 TX DPS method	CPI (AT=50rfu; no ST)	1 in 710,000,000
New TX DPS method	CPI (with ST=200rfu)	1 in 2.5
	mRMP (with ST=200rfu)	1 in 710
Future TX DPS method?	Probabilistic genotyping (with STRmix)	**2.91×10^{17}**

Comparison of CPI (with ST), mRMP, and two probabilistic genotyping approaches using 50 2-person mixtures

Electrophoresis 2014, *35*, 3125–3133

Todd W. Bille[1]
Steven M. Weitz[1]
Michael D. Coble[2]
John Buckleton[3]
Jo-Anne Bright[3]

[1]Bureau of Alcohol, Tobacco, Firearms and Explosives, Beltsville, MD, USA
[2]National Institute of Standards and Technology, Gaithersburg, MD, USA
[3]Institute of Environmental Science and Research Limited, Auckland, New Zealand

Received March 6, 2014
Revised August 2, 2014
Accepted August 12, 2014

Research Article

Comparison of the performance of different models for the interpretation of low level mixed DNA profiles

DNA analyses from forensic casework samples commonly result in complex DNA profiles. Often, these profiles consist of multiple contributors and display multiple stochastic events such as peak height imbalance, allelic or locus drop-out, allelic drop-in, and excessive or indistinguishable stutter. This increased complexity has established a need for more sophisticated methods of DNA mixture interpretation. This study compares the effectiveness of statistical models in the interpretation of artificially created low template two person mixed DNA profiles at varying proportions and template quantities. Two binary models (combined probability of inclusion and random match probability), a semicontinuous (Lab retriever), and continuous model (STRmix™) were compared. Generally, as the sophistication of the models increases, the power of discrimination increases. Differences in discrimination often correlate to each model's ability to use observed data effectively. Binary models require static thresholds resulting in unused data and outliers that may lead to difficult or incorrect interpretation. Semicontinuous and continuous models eliminate the stochastic threshold, however Lab Retriever does not account for stochastic events beyond drop-out and drop-in leading to possible less effective use of the data. STRmix™ incorporates all stochastic events listed above into the calculation making the most effective use of the observed data.

T.W. Bille et al. (2014) Comparison of the performance of different models for the interpretation of low level mixed DNA profiles. *Electrophoresis* 35: 3125-3133

American Academy of Forensic Sciences

Jurisprudence Section

Orlando, FL

February 20, 2015

ORLANDO 2015

http://www.cstl.nist.gov/strbase/pub_pres/Butler-DNA-interpretation-AAFS2015.pdf

Why DNA Interpretation Has Become More Challenging in Recent Years

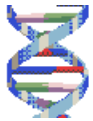
John M. Butler, Ph.D.

NIST Fellow & Special Assistant to the Director for Forensic Science

National Institute of Standards and Technology

Gaithersburg, Maryland

NIST FORENSIC SCIENCES



NIST

5 Reasons that DNA Results Are Becoming More Challenging to Interpret

1. **More sensitive DNA test results**
2. **More touch evidence samples** that are poor-quality, low-template, complex mixtures
3. **More options exist** for statistical approaches involving probabilistic genotyping software
4. **Many laboratories are not prepared** to cope with complex mixtures
5. **More loci being added** because of the large number of samples in DNA databases

http://www.cstl.nist.gov/strbase/pub_pres/Butler-DNA-interpretation-AAFS2015.pdf

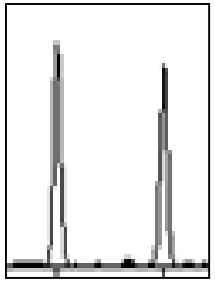
Math Analogy to DNA Evidence

$2 + 2 = 4$

$2x^2 + x = 10$

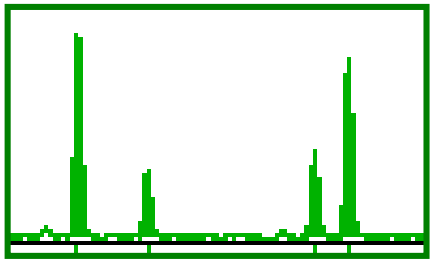
$$\int_{x=0}^{\infty} f(x)\,dx$$

Basic Arithmetic



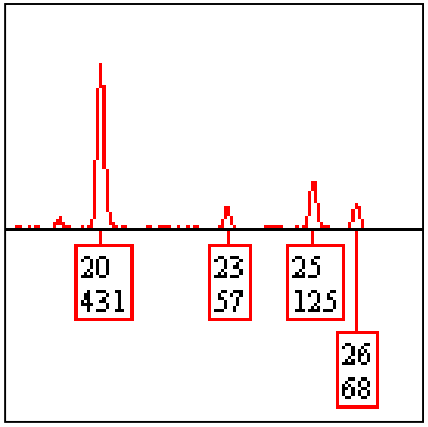
Single-Source DNA Profile
(DNA databasing)

Algebra



Sexual Assault Evidence
(2-person mixture with high-levels of DNA)

Calculus



Touch Evidence
(>2-person, low-level, complex mixtures perhaps involving relatives)

http://www.cstl.nist.gov/strbase/pub_pres/Butler-DNA-interpretation-AAFS2015.pdf

Options, Questions, and Challenges
(the challenge of wading into a moving stream of ongoing cases)

1. Do nothing and hope that past cases where CPI was inappropriately applied are okay
   - Not an option if you are interested in the best forensic science

2. Review old cases
   - Back to what date? 2008? 1999?
   - Potentially thousands of cases… cost, how to handle relative to current cases?

   a) Review CPI data with a stochastic threshold (ST)
      - **What ST value should be used?** ST is impacted by PCR conditions, CE injection time, sample desalting
      - Many low level DNA cases will go from an inclusion to inconclusive because no loci qualify with peaks below ST – impact on legal cases where statistical value of the DNA evidence essentially goes to zero

   b) Wait and get probabilistic genotyping (PG) method(s) online and then use PG to evaluate old cases
      - How long will it take to get PG methods validated and online?
      - PG requires method-specific calibration of allele drop-out and other parameters; **what values should be used for old data?** Some low level DNA mixture cases may still be inconclusive

Thank you for your attention

Acknowledgements: Recent discussions on these issues with Mike Coble (NIST), John Buckleton (ESR), Bruce Budowle (UNTHSC), and Todd Bille (ATF)

Contact Information

Michael D. Coble, Ph.D.
Research Biologist, NIST Applied Genetics Group
michael.coble@nist.gov
301-975-4330

John M. Butler, Ph.D.
NIST Fellow & Special Assistant to the Director for Forensic Science
Vice-Chair, National Commission on Forensic Science
john.butler@nist.gov
301-975-4049