EuroForGen – International Dissemination Conference
June 23, 2016
Venice, ITALY

Challenges in Forensic Genetics

John M. Butler, Ph.D.

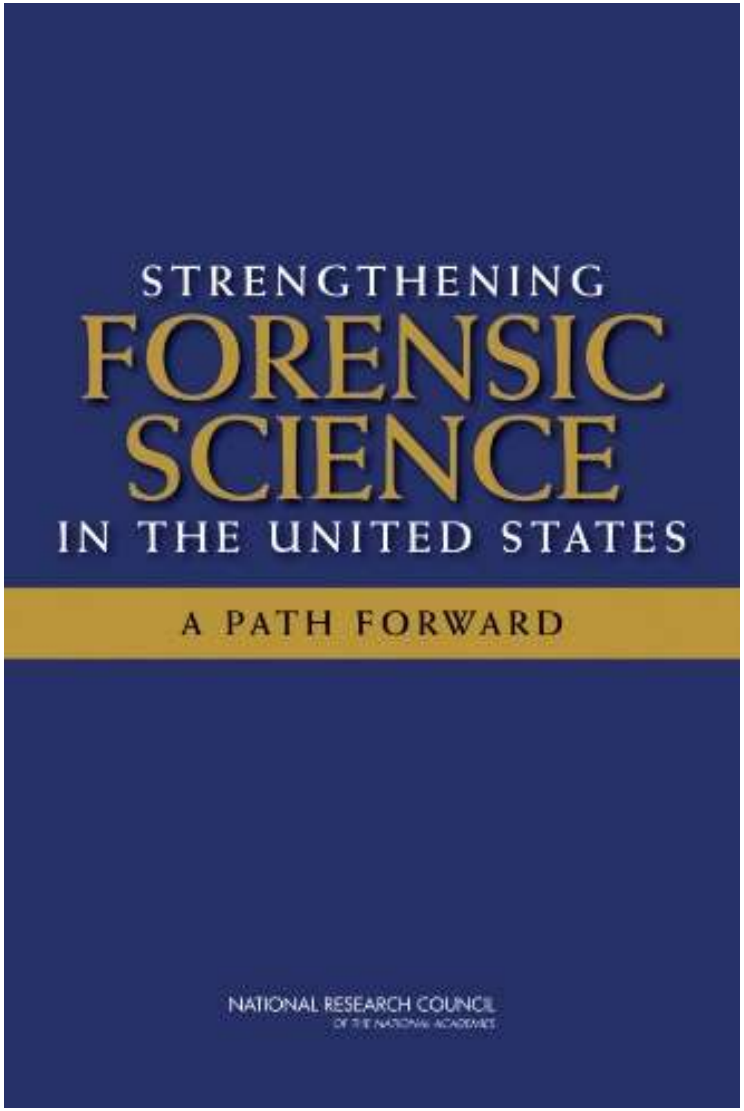
NIST Fellow & Special Assistant to the Director for Forensic Science

U.S. National Institute of Standards and Technology

Landmark Report Gives DNA Testing a Pass

The U.S. National Research Council of the National Academies issued a major report on forensic science in Feb. 2009.

*"**With the exception of nuclear DNA analysis**, no forensic method has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between evidence and a specific individual or source." (p. 41)*



p. 100 mentions limitations with DNA mixtures

International conference

The hidden side of DNA profiles. Artifacts, errors and uncertain evidence

Auditorium, Università Cattolica del Sacro Cuore

Rome, 27-28 April, 2012

UNIVERSITÀ CATTOLICA del Sacro Cuore

David Balding: “Low-template DNA cases are coming to court with limited abilities for sound interpretation. ... There are dangers with LTDNA but we know how to handle and manage them. Unfortunately, proper management is not a universal practice.”

Peter Schneider: “If you cannot explain your evidence to someone that is not from the field (like a judge) – and you need a lot of technical excuses to report something – then the result is not good. You should leave it on your desk and not take it to court. This is a very common sense approach to this problem.”

Reviewing the Past Helps Us Understand Potential Future Directions

PHILOSOPHICAL TRANSACTIONS B

rstb.royalsocietypublishing.org

Opinion piece

Cite this article: Butler JM. 2015 The future of forensic DNA analysis. *Phil. Trans. R. Soc. B* **370**: 20140252.
http://dx.doi.org/10.1098/rstb.2014.0252

Accepted: 26 February 2015

One contribution of 15 to a discussion meeting issue 'The paradigm shift for UK forensic science'.

The future of forensic DNA analysis

John M. Butler

National Institute of Standards and Technology, Gaithersburg, MD, USA

The author's thoughts and opinions on where the field of forensic DNA testing is headed for the next decade are provided in the context of where the field has come over the past 30 years. Similar to the Olympic motto of 'faster, higher, stronger', forensic DNA protocols can be expected to become more rapid and sensitive and provide stronger investigative potential. New short tandem repeat (STR) loci have expanded the core set of genetic markers used for human identification in Europe and the USA. Rapid DNA testing is on the verge of enabling new applications. Next-generation sequencing has the potential to provide greater depth of coverage for information on STR alleles. Familial DNA searching has expanded capabilities of DNA databases d where it is allowed. Challenges and opportunities that will impact the future of forensic DNA are explored including the need for education and training to improve interpretation of complex DNA profiles.

1. Introduction

Butler, J.M. (2015) The future of forensic DNA analysis. *Phil. Trans. R. Soc.* B 370: 20140252

Stages of Forensic DNA Progression

Stages	Time Frame	Description
Exploration	1985 - 1995	Beginnings, different methods tried (RFLP and early PCR)
Stabilization	1995 - 2005	Standardization to STRs, selection of core loci, implementation of Quality Assurance Standards
Growth	2005 - 2015	Rapid growth of DNA databases, extended applications pursued
Sophistication	*2015 to 2025 and beyond*	*Expanding tools available,* ***confronting privacy concerns***

Table 1 from J.M. Butler (2015) The future of forensic DNA analysis. *Phil. Trans. R. Soc. B* 370: 20140252

Critical Challenges Faced Today

- **Success of DNA testing** → significant growth in sample submissions → sample backlogs
  - Laboratory automation and expert system data review
  - Restrictive case acceptance policies to avoid law enforcement investigator 'swab-athons' at crime scenes

- **Greater detection sensitivity** → more complex DNA mixtures and low-template DNA with 'touch' evidence
  - Probabilistic genotyping to cope with increase in data interpretation uncertainty
  - Use of a complexity threshold to avoid "skating on thin ice"

Butler, J.M. (2015) The future of forensic DNA analysis. *Phil. Trans. R. Soc.* B 370: 20140252

Going Beyond the Core Competencies of Forensic DNA Testing…

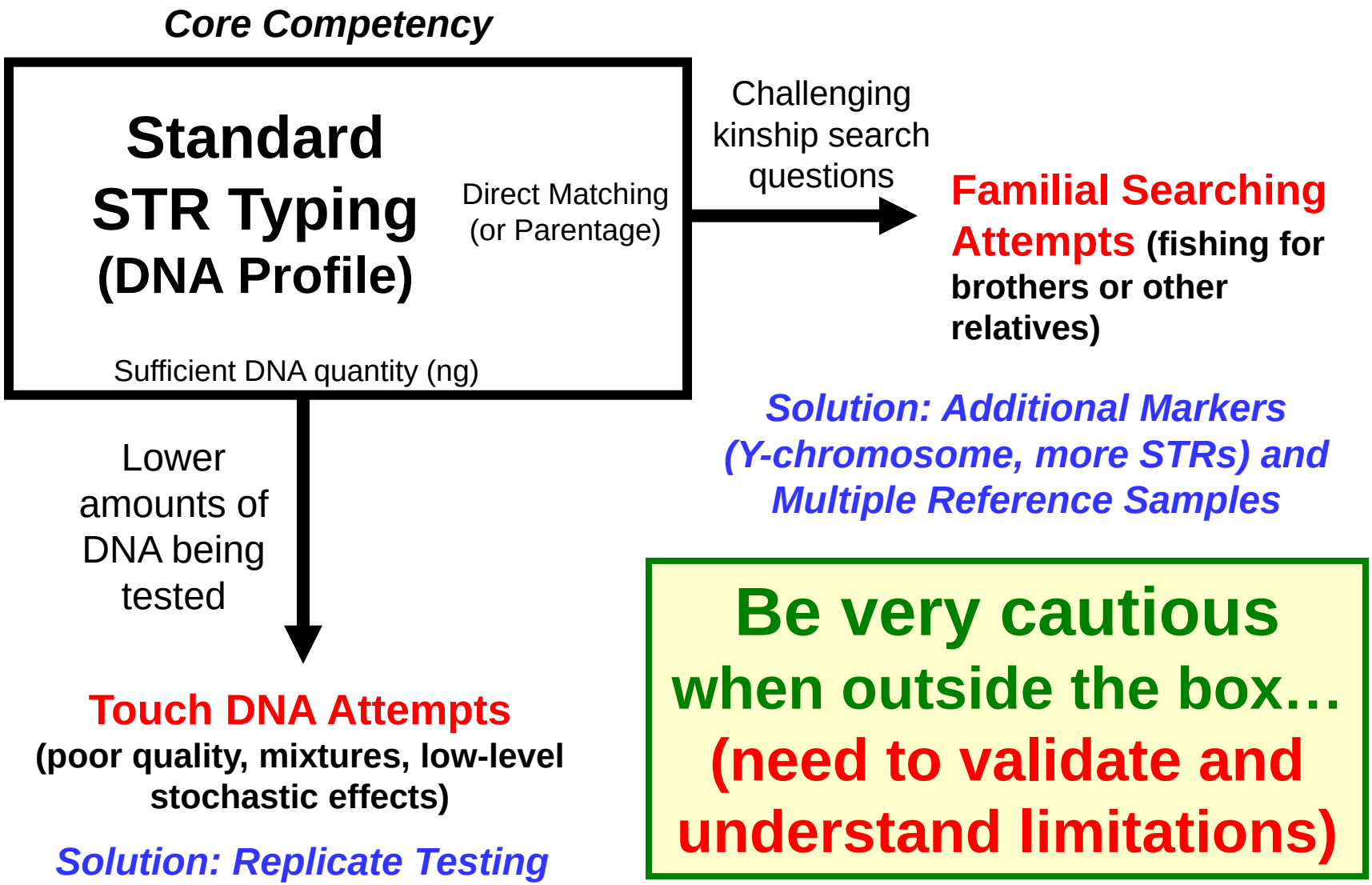


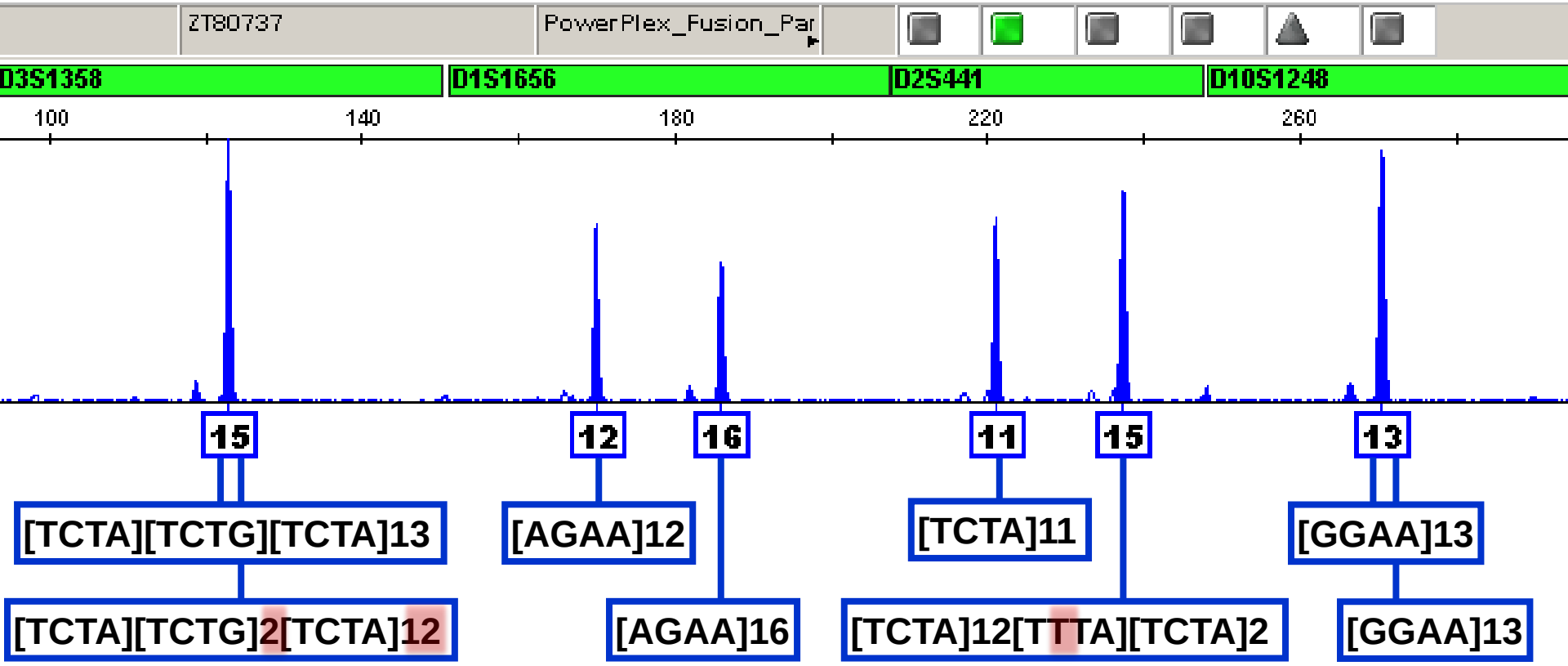
Figure 18.3 in Butler (2010) *Fundamentals of Forensic DNA Typing*

Current Trends in Forensic DNA

- ***Faster** results:* Rapid DNA capabilities and new sample-to-answer integrated instruments
- ***Higher** sensitivity:* New assays lowering the limits of detection, which makes interpretation more challenging
- ***Higher** information content:* Next-generation sequencing (NGS) for more markers & STR allele information
- ***Stronger** conclusions:* Mixture interpretation with probabilistic genotyping models

Butler, J.M. (2015) The future of forensic DNA analysis. *Phil. Trans. R. Soc.* B 370: 20140252

Forensic STR Sequence Diversity



Sequence-Based Heterozygote: A locus that appears homozygous in length-based measurements (such as CE), but is heterozygous by sequence

Next Generation Sequencing (NGS)/ Massively Parallel Sequencing (MPS)

- **Higher information content** with sequence data
  - Expanded number of STR loci and other genetic markers such as SNPs and InDels
    - New markers may enable additional applications (e.g., biogeographical ancestry and phenotypic prediction)
  - **Deeper depth of information on STR alleles**
    - For example, eight different sequence versions of D12S391 alleles among 197 samples examined (Gelardi et al. 2014)

- **Significant challenges with BIG data**
  - STR allele nomenclature issues (ISFG DNA Commission - Parson et al. 2016)
  - Data storage (do you retain terabytes of data?)
  - Data analysis time will increase…
  - Privacy concerns with additional genomic information

True Sample Components

Sample Processing

DNA Data Obtained

Potential STR alleles

12 13 14 15 16 17 18 19

Total DNA amplified

4x

Genotype

13,17

female

13

17

1x

13 14

male

Mixture Ratio of Components

Potential Allele Overlap & Stacking

Number of Contributors
(sample components)

Validation
establishes variation and limits in the processes involved

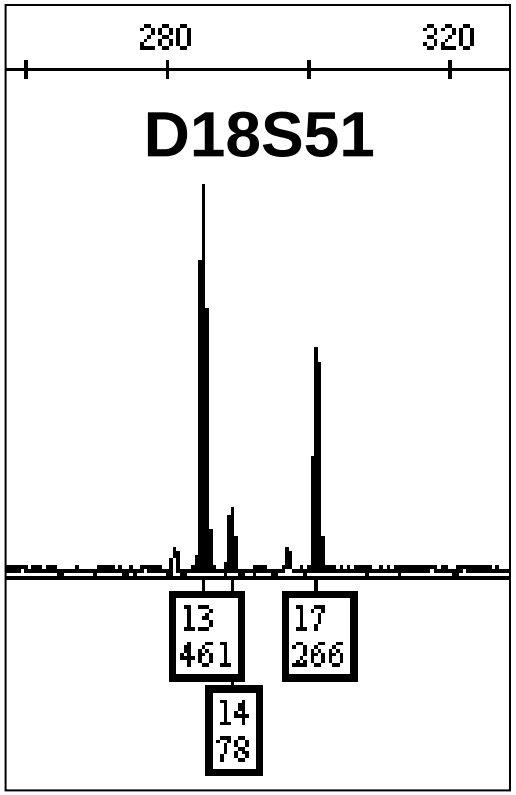
Extraction

PCR

CE Injection

CE Detection

portion of a CE electropherogram



Infer possible genotypes & determine sample components

From available data

Goal of Interpretation

Figure 6.2 in Butler (2015) *Advanced Topics in Forensic DNA Typing: Interpretation*

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

More Sensitive Assays and Instruments

- **Superb sensitivity is available** with DNA amplification using the polymerase chain reaction and laser-induced fluorescence detection with capillary electrophoresis

- Since 2007 (beginning with the release of the MiniFiler STR kit), **improved buffers and enzymes** have been used to boost DNA sensitivities in all STR kits
  - In 2010 the ABI 3500 Genetic Analyzer was released with 4X signal over the previous ABI 3100 and ABI 310 instruments
  - Energy-transfer dyes are used with some of the STR kits
  - Some labs increase the sensitivity dial with additional PCR cycles

- **So what is wrong with have improved sensitivity?**

Improved Sensitivity is a Two-Edged Sword

"As sensitivity of DNA typing improves, laboratories' abilities to examine smaller samples increases. This improved sensitivity is a two-edged sword. **With greater capabilities comes greater responsibilities to report meaningful results.** Given the possibility of DNA contamination and secondary or even tertiary transfer in some instances, **does the presence of a single cell (or even a few cells) in an evidentiary sample truly have meaning?**…"

Butler, J.M. (2015) *Advanced Topics in Forensic DNA Typing: Interpretation* (Elsevier Academic Press: San Diego), p. 458

Ian Evett and Colleagues' **Case Assessment and Interpretation:**

Hierarchies of Propositions

TABLE 16.2 Hierarchical Levels of Propositions Originally Developed by the UK Forensi 1998a, 1998b, Evett et al. 2000a, 2000b, Gill 2001)

Hierarchy Levels		Propositions	Decision Maker
Level III	Offense	Supplies the probability that a suspect has committed a criminal offense	Responsibility of the jury or judge
Level II	Activity	Informs regarding the kinds of activities which may have produced the forensic evidence	Jury or possibly scientist if given adequate case circumstances
Level I	Source	Addresses the source of the sample	Scientist
Sub-level I	Sub-source	With low amounts of DNA, the scientist may not be able to infer how the DNA arrived at the site where the DNA sample was collected	Scientist

Butler, J.M. (2015) *Advanced Topics in Forensic DNA Typing: Interpretation* (Elsevier Academic Press: San Diego), p. 458

More Touch Evidence Samples

https://www.ncjrs.gov/pdffiles1/nij/grants/222318.pdf

The DNA Field Experiment:
Cost-Effectiveness Analysis of the Use of DNA in the Investigation of High-Volume Crimes

John K. Roman
Shannon Reid
Jay Reid
Aaron Chalfin
William Adams
Carly Knight

Expanded DNA testing for burglary cases

NIJ April 2008 Research Report

http://www.nij.gov/journals/261/pages/dna-solves-property-crimes.aspx

DNA Solves Property Crimes (But Are We Ready for That?)
by Nancy Ritter

NIJ Journal October 2008 (vol. 261, pp. 2-12)

- **More poor-quality samples are being submitted**
  - Samples with <100 pg of DNA submitted in Belgium:
    19% (2004) → 45% (2008)
    (Michel 2009 FSIGSS 2:542-543)
- AAFS 2014 presentations showed poor success rates
  - NYC (**A110**): **only 10% of >9,500 touch evidence swabs from 2007 to 2011 produced usable DNA results**
  - Allegheny County (**A114**): examined touch DNA items processed from 2008 to 2013 across different evidence types (e.g., 6 of 56 car door handles yielded "resolvable profiles")

New Options Exist for Statistical Analysis

- Increase in approaches to try and cope with potential allele dropout → number of **probabilistic genotyping** methods have grown since Balding & Buckleton 2009 article

- Many possible choices for **probabilistic genotyping software** with commercial interests at stake

Balding, D.J. & Buckleton, J. (2009) Interpreting low template DNA profiles. *Forensic Sci. Int. Genet.* 4(1):1-10.

Gill P, Whitaker J, Flaxman C, Brown N, Buckleton J. (2000) An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA. *Forensic Sci. Int.* 112(1):17-40.

TABLE 13.1 **Probabilistic Genotyping Software Programs (as of March 2014)**

Program Name	Type	Creator(s)	Availability
LRmix	Discrete (semi-continuous)	Hinda Haned & Peter Gill	Open-source https://sites.google.com/site/forensicdnastatistics/PCR-simulation/lrmix
Lab Retriever	Discrete (semi-continuous)	Developed by David Balding and maintained by Norah Rudin and colleagues	Open-source http://www.scieg.org/lab_retriever.html
likeLTD	Discrete (semi-continuous)	David Balding	Open-source https://sites.google.com/site/baldingstatisticalgenetics/software/likeltd-r-forensic-dna-r-code
FST	Discrete (semi-continuous)	Adele Mitchell	Proprietary to the NYC OCME Forensic Biology Laboratory
Armed Xpert	Discrete (semi-continuous)	Developed by USACIL and maintained and improved by NicheVision	Commercial product http://www.armedxpert.com/
TrueAllele	Fully-continuous	Mark Perlin	Commercial product http://www.cybgen.com/
STRmix	Fully-continuous	Duncan Taylor, Jo-Anne Bright, John Buckleton	Commercial product http://strmix.esr.cri.nz/
DNA View Mixture Solution	Fully-continuous	Charles Brenner	Commercial product http://dna-view.com/

Discrete (semi-continuous) methods use only the allele information in conjunction with probabilities of drop-out and drop-in.
Fully-continuous methods use peak height data and other parameters in addition to the allele information.

Butler, J.M. (2015) *Advanced Topics in Forensic DNA Typing: Interpretation* (Elsevier Academic Press: San Diego), p. 341

Probabilistic Genotyping via Modeling Simulations

Mathematical Modeling of the Data

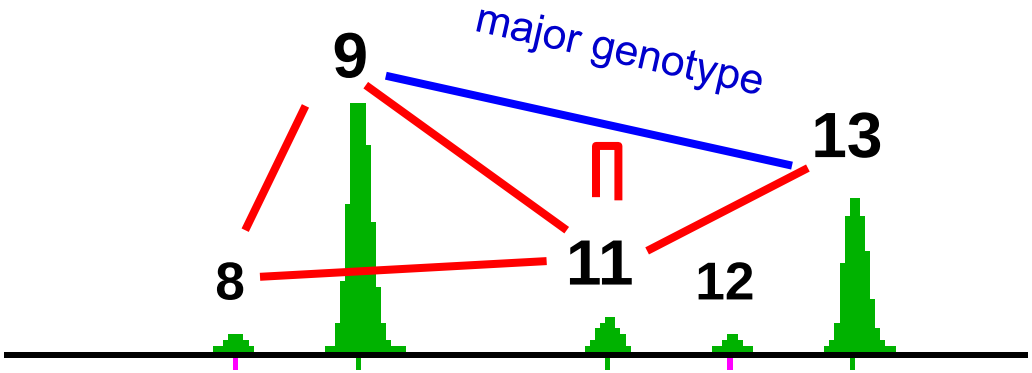
PHR, mix ratio, stutter, etc…

Typically thousands of simulations are performed

(MCMC)

Probable ***Genotypes*** to explain the mixture

D16S539



Minor Contributor Possible Genotypes	Probability
9,11	76%
11,11	15%
11,13	2%
8,11	2%
8,9	<1%
…	<1%

- Quantitative computer interpretation using numerous Markov Chain Monte Carlo (MCMC) simulations
- Models peak uncertainty and infers possible genotypes
- Results are presented as the Combined LR

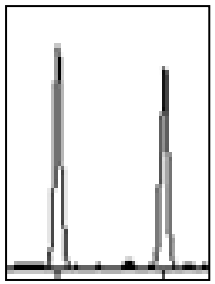
Math Analogy to DNA Evidence

$2 + 2 = 4$

$2x^2 + x = 10$

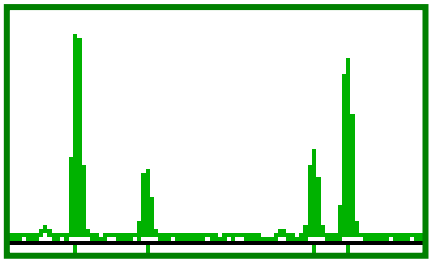
$$\int_{x=0}^{\infty} f(x)\,dx$$

Basic Arithmetic



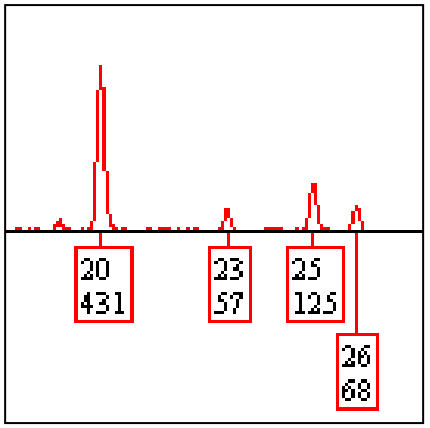
Single-Source DNA Profile
(DNA databasing)

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

Many laboratories are not prepared
to cope with complex mixtures

- Have **appropriate validation studies** been performed to inform proper interpretation protocols? (curriculum & classroom instruction)

- Are **appropriately challenging proficiency tests** being given? (graded homework assignments)

- **Would we want to go into a calculus exam only having studied algebra and having completed homework assignments involving basic arithmetic?**

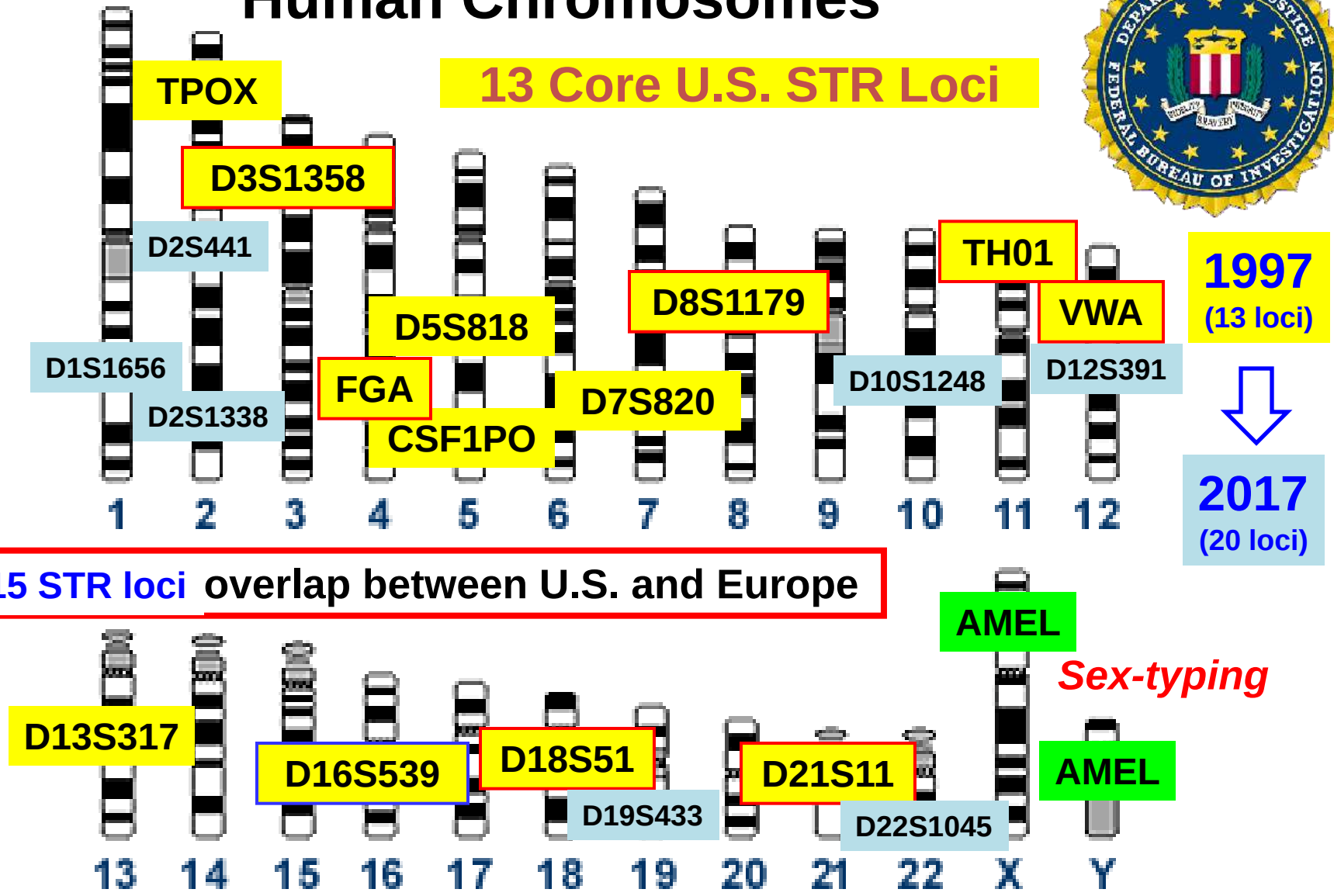
Why are we where we are today?

- The incredible success of DNA has lead to more sensitive methods and more samples being provided which has led to more complex mixtures (we are pushing the envelope)
  - Lower template DNA profiles have more uncertainty associated with them in terms of allele peak height variation

- Statistical interpretation techniques have not kept pace with the methodology improvements
  - Much of the forensic DNA community is effectively using a 1992 statistical tool (CPI) on 21st century data

Position of Forensic STR Markers on Human Chromosomes

Core STR Loci for the United States



Thoughts on Potential Improvements

Know the literature

Know the question being asked

Know the limits of what you can do

Steps in Forensic DNA Analysis

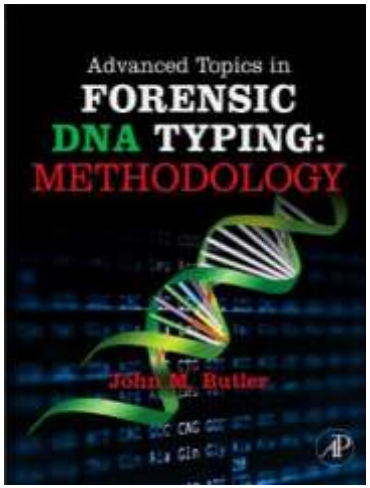
Understanding Results Obtained & Sharing Them

Gathering the Data

Interpretation

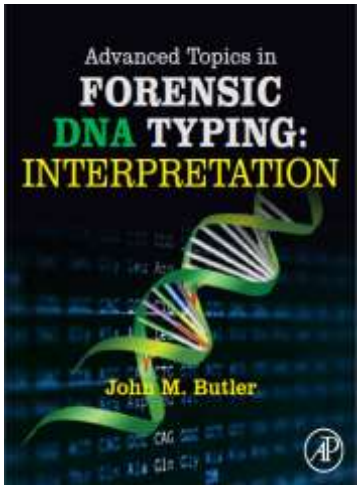
>1300 pages of information with >5000 references cited in these two books

Advanced Topics: Methodology



August 2011

Advanced Topics: Interpretation



October 2014

Know What Question You Are Trying to Answer

David Balding
University of Melbourne
Professor of Mathematics and Statistics

"…**Focus on the relevant question**. Many misleading statistical approaches [turn] out to be providing valid answers to the wrong questions."

– David Balding, Interpreting DNA evidence: can probability theory help? In J.L. Gastwirth (ed.) *Statistical Science in the Courtroom* (pp. 51-70) New York: Springer, 2000

Different Calculations Answer Different Questions

Method used	Questions being answered
Profile probability (random match probability, RMP)	What is the rarity of a specific DNA profile given the alleles observed? **What is the chance that a particular profile exists** in a population based on allele frequencies?
Match probability	Given that a particular profile has been seen (in the crime scene evidence and in the suspect), **what is the chance of it occurring again?**
Database match probability	How often would a DNA profile match the relevant forensic sample in a database of size *N*?

Adapted from Table 11.7, J.M. Butler (2015) ***Advanced Topics in Forensic DNA: Interpretation*** (Elsevier Academic Press)

Ian Evett on Interpretation

"The crucial element that the scientist brings to any case is the *interpretation* of those observations. This is the heart of forensic science: it is where the scientist adds value to the process."

Evett, I.W., et al. (2000). The impact of the principles of evidence interpretation on the structure and content of statements. *Science & Justice, 40,* 233-239.

Know the Limits of What You Can Do

- I have advocated for development of a "complexity (or uncertainty) threshold" with DNA evidence interpretation

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

Information from Chapter 7 of my New Book

Advanced Topics in Forensic DNA Typing: Interpretation

CHAPTER

7

Low-Level DNA and Complex Mixtures

"The limits of each DNA typing procedure should be understood, especially when the DNA sample is small, is a mixture of DNA from multiple sources, or is contaminated with interfering substances."

NRC I, 1992, p. 8

"For the complex DNA profile, there is no predominant or overarching standard interpretation method."

Peter Gill (Gill et al. 2012, report to the UK Forensic Science Regulator, p. 18)

"The limits of each DNA typing procedure should be understood, especially when the DNA sample is small, is a mixture of DNA from multiple sources…" (NRC I, 1992, p. 8)

Butler, J.M. (2015) *Advanced Topics in Forensic DNA Typing: Interpretation* (Elsevier Academic Press: San Diego), pp. 159-182

Perhaps We Should Slow Down with Some of the DNA Mixtures That We (Scientists and Lawyers) Are Taking On…

Poor Quality Conditions

Large Numbers of Contributors

The Future of Forensic DNA

is Similar to the Olympic Motto of

"Swifter, Higher, Stronger"

Resources **Training** **Action**

Acknowledgment and Disclaimers

I quote from my recent book entitled “Advanced Topics in Forensic DNA Typing: Interpretation” (Elsevier, 2015). I do not receive any royalties for this book. Completing this book was part of my job at NIST.

Although I chaired the SWGDAM Mixture Committee that produced the 2010 STR Interpretation Guidelines**, I cannot speak for or on behalf of the Scientific Working Group on DNA Analysis Methods.**

I have been fortunate to have had discussions with numerous scientists on interpretation issues including Mike Coble, Bruce Heidebrecht, Robin Cotton, Charlotte Word, Catherine Grgicak, Peter Gill, Ian Evett …

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

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

National Commission on Forensic Science (NCFS):
www.justice.gov/ncfs

Organization of Scientific Area Committees (OSAC):
www.nist.gov/forensics/osac/index.cfm

NIST FORENSIC SCIENCES

www.nist.gov/forensics

+1-301-975-4049 john.butler@nist.gov

Forensic Conference Organized by NIST

Planning has started for a second Symposium
Date: July 24-28, 2017 (Tentative)
Location: Washington DC
Sponsors that have been approached
DoD, FBI, NIST

http://www.nist.gov/director/international_forensics_home.cfm

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

U.S. initiatives to strengthen forensic science & international standards in forensic DNA

John M. Butler*

National Institute of Standards and Technology, Gaithersburg, MD, USA

- This review article covers recent U.S. activities to strengthen forensic science including the formation of the National Commission on Forensic Science and the Organization of Scientific Area Committees
- DNA documentary standards and guidelines from organizations around the world are also included

Butler, J.M. (2015) U.S. initiatives to strengthen forensic science & international standards in forensic DNA. *FSI Genetics* (volume 18, pp. 4-20)