Risks and Errors in DNA Identification

John M. Butler, Ph.D.
NIST Fellow & Special Assistant to the Director for Forensic Science
U.S. National Institute of Standards and Technology
Communication Across the Criminal Justice System is Important

Law Enforcement

Judicial

Laboratory

What Every Law Enforcement Officer Should Know About DNA Evidence

https://forensic.training.nij.gov/
Login

Security Upgrades

You should already have noticed and agreed to a new disclaim made to make our registration process and courses more sec...

- Your password must now be at least 8 characters and inc letters, numbers, and special characters (e.g., ~!@#$%
- You will be asked to reset your password every 90 days.
- You may not reuse any of your past 6 passwords.

Please Login

Username

Password

Login

Don't have an account? Register Now

Forgot your username? Retrieve your username now.
Butler Books on Forensic DNA Typing
DNA Capabilities to Aid Forensic Investigations

1. The **ability to identify the perpetrator**
2. Weight-of-evidence based on established genetic principles and statistics (Hardy-Weinberg 1908)
3. Established characteristics of genetic inheritance enables close **biological relatives** to be used for reference points using kinship associations
4. Superb **sensitivity** with PCR amplification (opens the possibility for contamination)
5. Well-established **quality assurance measures**
6. New **technology development** aided by genomics

Successful interpretation of DNA (Q-to-K comparison) depends on quality of the crime scene evidence (Q) and availability of suitable reference samples (K)
Concerns have been Raised over Potential for DNA Contamination

Previous articles by Peter Gill on this topic:


Discusses the Amanda Knox case DNA results
Forensic DNA Testing in the United States

- We have ~200 public (state and local government) laboratories performing forensic DNA analysis
  - Two large private companies (Bode Cellmark and Sorenson Forensics) and a few smaller ones perform forensic DNA analysis

- Almost 15 million DNA profiles in the national DNA database (NDIS: National DNA Index System) run by the FBI Laboratory
  - Since 1998, the U.S. has included 13 core STR (short tandem repeat) markers; starting in 2017, this number will increase to 20 required STR loci

- Laboratories have many different protocols and in some cases, submitting the same sample to two different laboratories could result in two different results
  - Efforts are underway to improve standardization in the field
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”

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
PCAST Report Comments on Forensic DNA

- Supports appropriate use of single-source and simple mixture DNA analysis
- Expresses reservations with complex DNA mixtures (≥3 contributors)

PCAST Co-Chairs

Eric Lander  John Holdren

Report to the President: Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods

Executive Office of the President: President’s Council of Advisors on Science and Technology

September 2016
Recent Forensic Problems in the News

**Washington DC Crime Lab** problems with **DNA Mixture Interpretation**

Director of D.C.'s embattled DNA lab resigns after suspension of testing

Max M. Houck had been the director since the lab opened in 2012. Auditors found major problems there.

Keith L. Alexander and Julie Zauzmer | Local | Apr 30, 2016

District could spend nearly $1 million for outside lab to test DNA evidence

The District is scrambling to find an alternative after the D.C. lab was ordered to cease DNA testing.

Keith L. Alexander | Crime | Apr 29, 2015

**Texas DNA Mixture Case Review**

August 2015

http://www.fsc.texas.gov/texas-dna-mixture-interpretation-case-review

April 2015


**Broward County Florida DNA Lab**

July 2016

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.”
New Book by Law Professor Erin Murphy

INSIDE THE CELL
THE DARK SIDE OF FORENSIC DNA
ERIN E. MURPHY

(Nation Books, Oct 2015)

400 pages

PART I
WHEN A MATCH ISN’T A MATCH: HOW DNA TESTING GOES WRONG

1. The Basics: DNA Typing for Dummies 3
2. The Dirty Business of Crime: The Challenge of Forensic Samples 18
3. Phantom Suspects: The Prevalence of DNA Transfer 29
4. Contamination, Mistake, and Outright Fraud 48
5. Single Cell Samples: Low Copy Number DNA 74

PART II
MISUNDERSTOOD MATCHES: HOW DNA STATISTICS MISLEAD

6. Brave New World of Probabilities 85
7. Dangers of the Database: Cold Hits and Coincidental Matches 106

PART IV
BUILDING A BETTER DNA POLICY

16. The Road to Reform: Efficiency, Accountability, Accuracy, Privacy, and Equality 265

Conclusion 311
Acknowledgments 313
Glossary 315
Appendix: The Mechanics of DNA Testing 321
Notes 335
Index 373
“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)
Updated Guidelines to Help with DNA Mixture Interpretation

Scientific Working Group on DNA Analysis Methods Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories

Current draft available for review is 90 pages long

http://www.swgdam.org/
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

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?…”
More Touch Evidence Samples

- More poor-quality samples are being submitted
  - Samples with <100 pg of DNA submitted in Belgium:
      (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”)

NIJ April 2008 Research Report

NIJ Journal October 2008 (vol. 261, pp. 2-12)
New Options Exist for Statistical Analysis

• Increase in approaches to try and cope with potential allele dropout \(\rightarrow\) number of **probabilistic genotyping** methods have grown since Balding & Buckleton 2009 article

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


Single-Source Sample vs Mixture Results

Possible combinations at D3S1358 include:
- 14, 17 with 16,16
- 14,14 with 16,17
- 14,16 with 17,17

Maternal and paternal allele are both 16 so the signal is twice as high

Multiple possible combinations could have given rise to the mixture observed here
Probabilistic Genotyping via Modeling Simulations

Mathematical Modeling of the Data

Typically thousands of simulations are performed (MCMC)

Probable Genotypes to explain the mixture

PHR, mix ratio, stutter, etc…

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

<table>
<thead>
<tr>
<th>Minor Contributor Possible Genotypes</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>9,11</td>
<td>76%</td>
</tr>
<tr>
<td>11,11</td>
<td>15%</td>
</tr>
<tr>
<td>11,13</td>
<td>2%</td>
</tr>
<tr>
<td>8,11</td>
<td>2%</td>
</tr>
<tr>
<td>8,9</td>
<td>&lt;1%</td>
</tr>
<tr>
<td>...</td>
<td>&lt;1%</td>
</tr>
</tbody>
</table>
Math Analogy to DNA Evidence

2 + 2 = 4

Basic Arithmetic

2x^2 + x = 10

Algebra

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

Calculus

Single-Source DNA Profile (DNA databasing)

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

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

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?
Netherlands Forensic Institute (NFI)
Article on Forensic DNA Error Rates

Error rates in forensic DNA analysis:
Definition, numbers, impact and communication

Ate Kloosterman a,b,c,*, Marjan Sjerps b,d, Astrid Quak a

a Department of Human Biological Traces (HBS), Netherlands Forensic Institute, P.O. Box 24044, 2490 AA The Hague, The Netherlands
b Department of Science, Interdisciplinary Research, Statistics and Knowledge Management (WTK), Netherlands Forensic Institute, P.O. Box 24044, 2490 AA The Hague, The Netherlands
c Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands
d Korteweg-de Vries Institute for Mathematics, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands

Kloosterman et al. (2014) Error rates in forensic DNA analysis: definition, numbers, impact and communication. FSI Genetics 12: 77-85
**Reported DNA Error Rates**

<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>Tests</th>
<th>Errors</th>
<th>1 in</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plebani &amp; Carraro [33]</td>
<td>1997 (3 mo.)</td>
<td>40,490</td>
<td>189</td>
<td>214</td>
<td>0.47%</td>
</tr>
<tr>
<td>Carraro &amp; Plebani [36]</td>
<td>2007 (3 mo.)</td>
<td>51,746</td>
<td>160</td>
<td>323</td>
<td>0.31%</td>
</tr>
<tr>
<td>Stahl et al. [34]</td>
<td>1998 (3 yr.)</td>
<td>676,564</td>
<td>4,135</td>
<td>164</td>
<td>0.61%</td>
</tr>
<tr>
<td>Hofgärtner &amp; Tait [35]</td>
<td>1999 (1 yr.)</td>
<td>88,394</td>
<td>293</td>
<td>302</td>
<td>0.33%</td>
</tr>
</tbody>
</table>

**# notifications**

<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>Tests</th>
<th>Errors</th>
<th>1 in</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>NFI DNA casework</td>
<td>2008</td>
<td>66,391</td>
<td>328</td>
<td>202</td>
<td>0.49%</td>
</tr>
<tr>
<td>NFI DNA casework</td>
<td>2009</td>
<td>82,896</td>
<td>329</td>
<td>252</td>
<td>0.40%</td>
</tr>
<tr>
<td>NFI DNA casework</td>
<td>2010</td>
<td>89,977</td>
<td>435</td>
<td>207</td>
<td>0.48%</td>
</tr>
<tr>
<td>NFI DNA casework</td>
<td>2011</td>
<td>100,407</td>
<td>526</td>
<td>191</td>
<td>0.52%</td>
</tr>
<tr>
<td>NFI DNA casework</td>
<td>2012</td>
<td>132,456</td>
<td>572</td>
<td>232</td>
<td>0.43%</td>
</tr>
<tr>
<td>NIST Identifier JFS 2003</td>
<td>2003</td>
<td>11,200</td>
<td>7</td>
<td>1600</td>
<td>0.06%</td>
</tr>
<tr>
<td>FBI errata JFS 2015 population data</td>
<td>1999</td>
<td>30,550 alleles</td>
<td>51</td>
<td>599</td>
<td>0.17%</td>
</tr>
</tbody>
</table>

Even with single-source, pristine samples, the error-rate is not zero!
Not all quality issue notifications (aka “errors”) are equal

Table 3
Types of quality issue notifications (QINs) at the NFI in the years 2008–2012. In 2011 it was decided to no longer incorporate the type c QIN: opportunities for improvement (n = 2 in 2011 and n = 10 in 2012) in the yearly totals of this overview.

<table>
<thead>
<tr>
<th>Category</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. External origin</td>
<td>23</td>
<td>10</td>
<td>23</td>
<td>54</td>
<td>100</td>
</tr>
<tr>
<td>b. External contamination</td>
<td>3</td>
<td>0</td>
<td>5</td>
<td>24</td>
<td>22</td>
</tr>
<tr>
<td>c. Room for improvement</td>
<td>11</td>
<td>6</td>
<td>3</td>
<td>(2)</td>
<td>(10)</td>
</tr>
<tr>
<td>d. Positive response</td>
<td>19</td>
<td>9</td>
<td>11</td>
<td>6</td>
<td>17</td>
</tr>
<tr>
<td>e. Clerical (no adverse outcome)</td>
<td>29</td>
<td>25</td>
<td>92</td>
<td>77</td>
<td>82</td>
</tr>
<tr>
<td>f. Not related to case work</td>
<td>13</td>
<td>9</td>
<td>20</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>g. Other (NFI related)</td>
<td>230</td>
<td>270</td>
<td>281</td>
<td>355</td>
<td>346</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>328</td>
<td>329</td>
<td>435</td>
<td>526</td>
<td>572</td>
</tr>
</tbody>
</table>
# Checks and Controls on Forensic DNA Results

<table>
<thead>
<tr>
<th>Community</th>
<th>FBI DNA Advisory Board’s Quality Assurance Standards (also interlaboratory studies)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laboratory</td>
<td>ASCLD/LAB, ANAB, A2LA Audits and Accreditation</td>
</tr>
<tr>
<td>Analyst</td>
<td>Proficiency Tests &amp; Continuing Education</td>
</tr>
</tbody>
</table>
| Method/Instrument | Validation of Analytical Performance  
(with aid of *traceable reference materials*) |
| Protocol  | Standard Operating Procedure is followed                                          |
| Data Sets | Allelic ladders, positive and negative amplification controls, and reagent blanks are used |
| Individual Sample | Internal size standard present in every sample                                  |
| Interpretation of Result | Second review by qualified analyst/supervisor |
| Court Presentation of Evidence | Defense attorneys and experts with power of discovery requests |
Planning has started for a second Symposium
Date: **July 24-28, 2017**
Location: Gaithersburg, MD
Sponsors that have been approached
DoD, FBI, NIST

http://www.nist.gov/director/international_forensics_home.cfm
National Institute of Standards and Technology

- Science agency **part of the U.S. Department of Commerce**
- Started in 1901 as the **National Bureau of Standards**
- Name changed in 1988 to the **National Institute of Standards and Technology (NIST)**
- Forensic science research activities dating back to 1920s
- Partnership since 2013 with U.S. Department of Justice to create the National Commission on Forensic Science (NCFS) and the Organization of Scientific Area Committees (OSAC)

- Primary campus in Gaithersburg, Maryland (near Washington, D.C.)
- >3,400 employees and >3,700 associates
- Supplies >1300 reference materials
- Defines official time for the U.S.