CONTINUING LEGAL EDUCATION

WINTER 2016

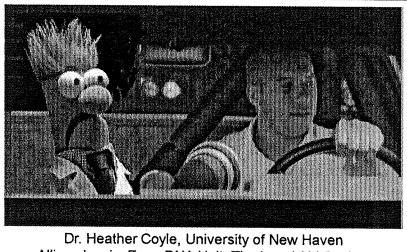
FEBRUARY 4, 2016

DNA - FST UPDATE

HEATHER COYLE, PH.D., ALLISON LEWIS, ESQ., AND KYLE WATTERS, ESQ.

Sponsored by:
Appellate Division, First Department
Assigned Counsel Plan for the First Department

FST Update: Science in the Fast Lane



Dr. Heather Coyle, University of New Haven
Allison Lewis, Esq., DNA Unit, The Legal Aid Society
Kyle Watters, Esq.

1

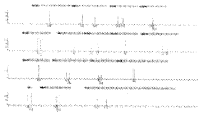
Today's CLE

- The basics of DNA and FST
- Current state of FST challenges
- Potential evidentiary challenges going forward
- Cross exam ideas for FST and discussion

DNA and FST Basics

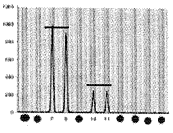
A classic DNA profile…

SINGLE SOURCE, DNA PROFILE (EG BUCCAL SWAB, SINGLE SOURCE EVIDENCE):



Max two alleles (peaks) at each locus

…OR MIXTURE WITH DEDUCIBLE PROFILE AT EACH LOCUS; I.E., ONE PROFILE STANDS OUT IN THE MIXTURE:



RFUs (or peak height) separates the two contributors

DNA & FST BASICS 4

Where you can see one distinct profile = RMP

Random Match Probability = the probability that if you grabbed a random unrelated person in a given population, they would match the profile. Based on known allele frequencies.

Your report looks like:

RESULTS AND CONCLUSIONS:

STR DNA testing using the AmpFlSTR® Identifiler® PCR Amplification Kit was performed on the DNA sample from [redacted]. A DNA profile was determined and matches the DNA profile of Male Donor A.

This DNA profile of Male Donor A is expected to be found in approximately:

15 locus result (inner thigh swab, sperm cell fraction)
1 in greater than 6.80 trillion people

DNA & FST BASICS 5

RMP

- It's the classic statistical calculation for a single source profile (max two alleles at any one locus). How rare is that profile on its own… Considered the gold standard stat.
- Allows them to upload that evidence profile into database (CODIS).
- NOT probability your client committed the crime.
- Also 6.80 trillion is a base number. The actual number is probably higher. They cap it for testifying purposes.

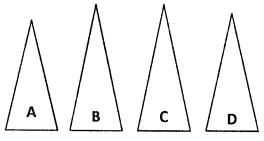
DNA & FST BASICS 6

The statistic (an aside)

- You will hear them say "you would need a thousand planet Earths to see that profile again..."
- That is just not true. It is simply a product of a statistical calculation. There could theoretically be a person on the jury that has the same profile. Or there could be a dozen in the world. That is unknowable. Rather, these are statistical flourishes.
- "Statistics is a battle between mathematics and reality." Anon, OCME

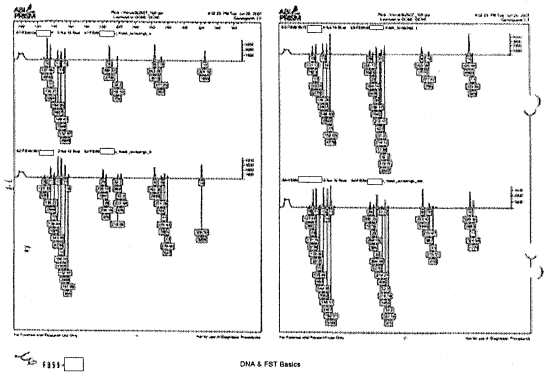
☐ LIES
☐ DAMNED LIES
☑ STATISTICS

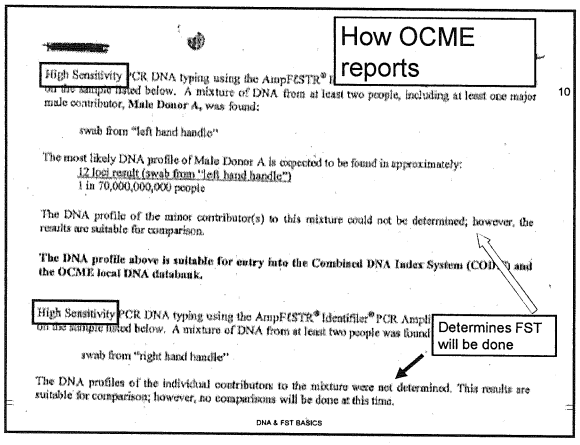
FST is used for complex mixtures



Does A go with B?
With C?
With D? Or is it a homozygote?

a complex mixture is a "soup of alleles" – you can't tell genotypes (individual DNA profiles)




How OCME reports

High Sensitivity PCR DNA typing using the AmpFℓSTR® I
on the sample listed below. A mixture of DNA from at least two people, including at least one major male contributor, **Male Donor A**, was found:

swab from "left hand handle"

The most likely DNA profile of Male Donor A is expected to be found in approximately:
12 loci result (swab from "left hand handle")
1 in 70,000,000,000 people

The DNA profile of the minor contributor(s) to this mixture could not be determined; however, the results are suitable for comparison.

The DNA profile above is suitable for entry into the Combined DNA Index System (COD[illegible]) and the OCME local DNA databank.

High Sensitivity PCR DNA typing using the AmpFℓSTR® Identifiler® PCR Ampli
on the sample listed below. A mixture of DNA from at least two people was found

Determines FST will be done

swab from "right hand handle"

The DNA profiles of the individual contributors to the mixture were not determined. This results are suitable for comparison; however, no comparisons will be done at this time.

DNA & FST BASICS

10

So what is FST?

A statistical calculation which compares two likelihoods:

The mixture occurred WITH Δ as contributor
(prosecution theory)

v.

The mixture occurred WITHOUT Δ as contributor
(defense theory)

They can include another known contributor (the cw, a cop, the bouncer that took the gun off your guy)... BUT THEY DECIDE WHO!!

DNA & FST BASICS 11

The math of FST

12

$$\frac{\Delta + 2 \text{ unknown unrelated individuals}^*}{3 \text{ unknown unrelated individuals}}$$

Prosecutor's Hypothesis

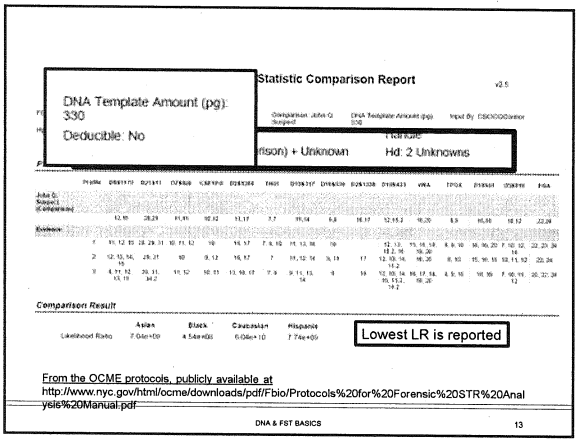
Defense Hypothesis

LR >1= supports prosecution hypothesis

LR <1 =supports defense hypothesis

*FST is not validated to be used with 4 contributors

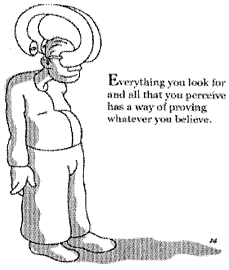
DNA & FST BASICS



From the OCME protocols, publicly available at http://www.nyc.gov/html/ocme/downloads/pdf/Fbio/Protocols%20for%20Forensic%20STR%20Analysis%20Manual.pdf

The Numerator of LR Models: Presuming the Defendant's Guilt

- Prosecutor's hypothesis begins with assumption that the suspect's DNA was included in the mixture;
- Versus RMP, which looks at the rarity of each allele across the profile without looking at a particular suspect.



Forensic Statistical Tool Results

15

The DNA profile of [ ] as compared to the results in the following case:

FBI Number	Complaint Number	Entity	Case Report ID
[ ]	[ ]	PSNY	CRT [ ]2

Based on a comparison of the DNA profile of [ ] the mixture found on the sample listed below, he is included as a possible contributor. Therefore, a likelihood ratio was calculated.

swab of "slide/slide groove"

This DNA mixture is approximately **1.41 times more probable** if the sample originated from [ ] and one unknown, unrelated person than if it originated from two unknown, unrelated persons. **Therefore, there is limited support that [ ] and one unknown, unrelated person contributed to this mixture, rather than two unknown, unrelated persons.**

FST is a stat about the MIXTURE, not the Δ

NOT a statistic of inclusion...

But rather a likelihood comparing two possibly hypotheses.

Keep an eye out for [judges, ADAs, criminalist, pretty much anyone working with this stuff] to MISSTATE what the FST means.

The criminalist should give you this pretty easily. This is NOT a statistic of how likely it is to be him...

Often stated incorrectly...

program allows OCME to calculate likelihood ratios. Likelihood ratios are the formula by where calculations are made that determine how likely or probable an individual is a contributor to a DNA sample from multiple contributors or how unlikely or improbable that an individual is not a contributor to that particular multiple contributor DNA sample. In order

ANON.

Consider People v. Wright

The Court of Appeals found IAC for failure to object to π's overstatement of what the DNA evidence was able to tell the jury.

"Here, defense counsel failed to object when the prosecutor misrepresented the scientific import of the DNA evidence, suggested that the evidence directly linked defendant to the murder although it did not, and made statements that contradicted the expert testimony about the limitations of YSTR DNA analysis. Given the significance of the DNA evidence, defense counsel's silence is inexplicable, and under the circumstances of this case, his inaction was error."

25 NY3d 769, 780.

And this was AFTER defense counsel made these issues clear on cross.

Current State of FST Admissibility Challenges

Frye Standard of Admissibility

method must have gained general acceptance in the relevant scientific community

Daubert Standard of Admissibility: *"Gatekeepers, not armed guards!"*

- Called Frye's general acceptance test "uncompromising" and contrary to liberal thrust of federal rules
- Factors for court to consider:
  - General Acceptance
  - Error Rate
  - Is theory testable using the scientific method
  - Existence & maintenance of standards controlling technique's operation
  - Subjected to Peer Review

FST is accepted: People v. Rodriguez

"[T]hat FST rests firmly upon two pillars, Polymerase Chain [sic] Reaction-Short Tandem Repeat (PCR-STR) DNA analysis and the likelihood ratio or LR. Both have long been generally accepted by forensic scientists as reliable."

People v. Rodriguez, Ind.5471/2009, (Sup. Ct. N.Y.Co. 2014) (Carruthers, J.) at 8.

OCME methods ruled inadmissible after Frye hearing in People v. Collins/Peaks

15 N.Y.S.3d 564 (Kings Co. Sup. Ct. 2105)

LOW COPY NUMBER/ HIGH SENSITIVITY DNA TESTING	FORENSIC STATISTICAL TOOL (FST)
▪Testing method on small amounts of DNA	▪Statistical method
▪Reserved for violent felonies with a victim	▪Software program for stat in complex mixtures
	▪Used only when analyst cannot exclude person from mixture

FST is *not* accepted: People v. Collins

49 Misc.3d 59515 N.Y.S.3d 564 (Sup.Ct. Kings Co. July 2, 2015) (Dwyer, J.)

....many courts have stopped there, finding that there is nothing novel about the FST and thus that there is no basis for a Frye attack upon it. But that does no justice to the actual positions pressed by the defense. The key advance in programs like the FST is that they factor into the Bayesian calculations the likelihood that alleles have appeared or failed to appear as a result of stochastic effects. The defense contends that the manner in which the drop-in and drop-out rates are assessed at each locus is not generally accepted in the DNA community.

See also People v. Abney, (Sup.Ct. Kings Co. 2015) (Riviezzo, J.)

What is the proper scope of analysis for the court?

N.Y. EXECUTIVE LAW § 995 (McKinneys 2014) defines ***DNA testing methodology:***

"methods and procedures used to extract and analyze DNA material, as well as the methods, procedures, assumptions, and studies used to draw statistical inferences from the test results."

FST Not Admitted

Hearing held & decision rendered:

People v. Collins, 15 N.Y.S.3d 564.

Hearings ordered in:

People v. Abney
- No ruling because DA withdrew evidence

Hearings held in:

United States v. Rashawn Smalls, 14-CR-00414 (BMC)
- No ruling because AUSAs withdrew evidence

FST Admissible

People v. Debraux, 2015 WL 5778744 (Sup.Ct. NY Co. 2015) (Kahn, J.)

People v. Rodriguez, Ind. 5471/2009 (Sup.Ct. NY Co. 2013) (Carruthers, J.)

People v. Belle, 47 Misc.3d 1218(A) (Sup. Ct. Bx.Co 2015) (Fabrizio, J.)

People v. Garcia, 39 Misc.3d 482 (Sup.Ct.Bx. Co. 2013) (Iacovetta, J.)

People v. Styles, 40 Misc.3d 1205(A)975 N.Y.S.2d 369 (Table) (Sup.Ct. Kings Co. 2013) (Donnelly, J.)

Numerous other unpublished trial court cases in NYC and one in NJ: **NJ v. Rochat,** (Superior Ct. NJ Bergen Co. 2015) (unpublished).

Let the jury figure it out...

US v. Morgan

53 F.Supp.3d (SDNY 2014)

"The publications and testimony the Government has presented demonstrate persuasively that, despite dissenting voices, a sufficient scientific underpinning for OCME's LCN methodology exists. It is appropriate, and indeed important, to consider the more controversial aspects of OCME's LCN methodology – but the

proper forum for the debate is before a jury, and it is for the jury to decide which of two conflicting experts' testimony to credit, and how much weight to give the evidence it accepts."

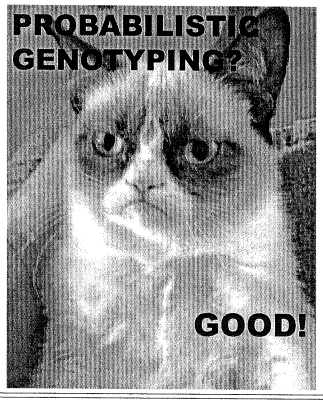
November 7, 2014 – Holding

"To have a technique that is so controversial that the community of scientists who are experts in the field can't agree on it and then to throw it in front of a lay jury and expect them to be able to make sense of it, is just the opposite of what the Frye standard is all about. And so ultimately I had to think that it was inconsistent with Frye to give it to a jury when so many experts in the field don't think that it is appropriate."

United States v. Smalls

14-CR-00414 (BMC) (EDNY, Cogan, J.)

- Number of contributors argument
- OCME's conclusion that the gun sample contained "at least 3" contributors was not supported by OCME's own published work on estimating # of contributors
- FST not validated to be run on 4 contributors
- Affidavits from Drs. Dan Krane and Ranajit Chakraborty
- AUSAs withdrew evidence once looked like defense would prevail after Daubert hearing on papers



LR Approaches Worldwide

LoComatioN (Peter Gill)

STRMix (Kelly et al)*

likeLTD (Balding)

Lab Retriever (Lohmueller)

TrueAllele (Perlin)

Forensim/LRmix (Haned)

DNAMixtures

Forensic Statistical Tool (OCME)

* OCME has indicated they will move to STRMix

TrueAllele

http://www.cybgen.com/information/admissibility/page.shtml

Number of admissibility decisions, including:

Com. v. Foley, 38 A.3d 882 (PA Superior Ct. 2012)

- No Frye hearing required
- Not novel; instead, trial ct found a "refined application of the product rule"
- Found "no legitimate dispute" over TrueAllele methodology

People v. Wakefield, 47 Misc.3d 8509 N.Y.S.3d 540 (N.Y.Sup.Ct. Schen.Co. 2015)

- Hearing held
- Ct found that the "various scientific principles used by Cybergenetics TrueAllele Casework have been long ago accepted and endorsed by the scientific community"
- Validated and peer reviewed
- "a plethora of evidence in favor of Cybergenetics TrueAllele Casework, *and there is no significant evidence to the contrary.*"

Other Potential Evidentiary Challenges to FST

"As applied" Frye challenge?

- ✓Number of contributors?
- ✓Used in situations involved related people?
- ✓Low level (low copy number, high sensitivity) DNA? Very degraded?

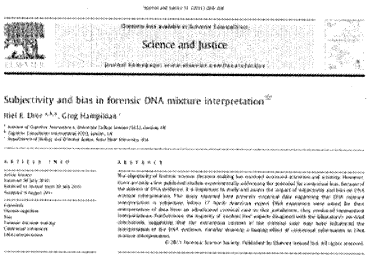
FST only for 2 and 3 person mixtures

...check the number of contributors

As explained above all evidence profiles are finalized without consideration of any potential known contributor sample. Samples must be categorized as single source or mixtures and, for mixtures, the number of contributors must be estimated. Characteristics of a mixture can be used to determine whether to treat the mixture as two, three, or four people [24,33–36]. Yet, due to allele sharing amongst related and unrelated individuals, there will always be a level of uncertainty to this determination [33]. Using only the maximum number of alleles observed at any locus to estimate the number of contributors to a mixture will often lead to an underestimate [33,34]. Nevertheless, using the minimum number of contributors typically results in the lowest possible LR, the LR that most favors the defendant [34]. Tools to better estimate the number of contributors are being explored [24,35,36]. FST is currently online for analysis of two- or three-person mixtures. Validation of four-person models is currently in progress.

See attached article; Validation of a DNA mixture statistics tool incorporating allelic drop-out and drop-in, A. Mitchell, et al, FSI, 2012.

Number of contributors not easy to calculate & subjective (think Theresa Carragine)



Science and Justice

Subjectivity and bias in forensic DNA mixture interpretation☆

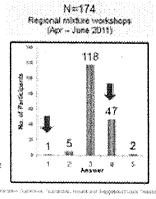
Itiel E. Dror a,b,*, Greg Hampikian c

Even the US gov't agrees:

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

If you asked 10 analysts in your laboratory to interpret a complex mixture you would get:

1. 100% consensus
2. 100% non-consensus
3. **Mostly consensus**
4. **A large range of answers**
5. We don't interpret complex mixtures



There is a recognition of the variation that exists with how analysts interpret complex mixtures.

The Denominator of LR Models: Is it *Really* the "Defense" Hypothesis?

39

Norwegian Institute of Public Health

EUROFORGEN

Numbers of contributors

- There is no need to anchor the number of contributors to be the same under Hp and Hd – they will often be different
- There will be differences between prosecution and defence hypotheses that courts will wish to explore. Software will facilitate the exploration

Adding an extra contributor only to the denominator

Accepted Manuscript

Title: The effect of varying the number of contributors on likelihood ratios for complex DNA mixtures

Author: Corina C.G. Benschop Hinda Haned Loes Jeurissen Peter D. Gill Titia Sijen

"That **we obtained lower LRs more often when an extra contributor was considered under Hd** [the defense hypothesis] is to be expected, as a random person may fit more easily with more unknowns, and therefore increases the likelihood of Hd and lowers the LR." S. 3.2.4

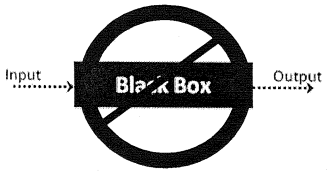
Try to get the source code...

LAS currently litigating this issue in a handful of cases.

One subpoena signed in the Bronx (Judge Clancy), but we'll keep you posted.

Eg, there are open source programs you can use online:

http://arxiv.org/pdf/1307.4956.pdf



Slate

future tense THE CITIZEN'S GUIDE TO THE FUTURE

Convicted by Code

By Rebecca Wexler

Open Source Necessary for Adequate Scientific Review

Open source refers to online or published access to source code; allows for alternative, true defense hypothesis to be tested.


Forensic Science International: Genetics

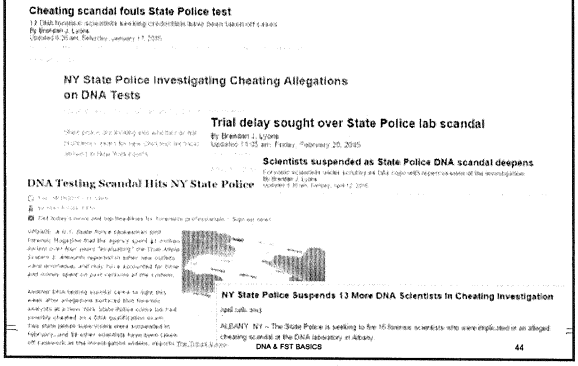
DNA commission of the International Society of Forensic Genetics: Recommendations on the evaluation of STR typing results that may include drop-out and/or drop-in using probabilistic methods

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

"Open-source is strongly encouraged since this solution offers unrestricted peer review and best assurance that methods are fit for purpose."

Gill, et. al. at p. 684.

Cheating on the black box


Cheating scandal fouls State Police test

12 DNA forensic scientists banking credentials have been taken off cases

By Brendan J. Lyons

Updated 6:20 am, Saturday, January 17, 2015

NY State Police Investigating Cheating Allegations on DNA Tests

Trial delay sought over State Police lab scandal

By Brendan J. Lyons

Updated 11:25 am, Friday, February 20, 2015

Scientists suspended as State Police DNA scandal deepens

For some scientists under scrutiny at DNA lab, plea deals require waiver of the investigation

By Brendan J. Lyons

Updated 5:30 pm, Tuesday, April 12, 2016

DNA Testing Scandal Hits NY State Police

NY State Police Suspends 13 More DNA Scientists In Cheating Investigation

April 14th, 2015

ALBANY, NY – The State Police is seeking to fire 15 forensic scientists who were implicated in an alleged cheating scandal at the DNA laboratory in Albany

STRmix Source Code Debacle

45

Queensland authorities confirm 'miscode' affects DNA evidence in criminal cases

DAVID MURRAY THE COURIER-MAIL MARCH 21, 2015 12:00AM

Relevance??

They report out an LR of 1.1??

Plenty of experts in the field would say that an LR under 10 or under 100 has no mathematical significance whatsoever...

Meanwhile:

Limited, moderate, strong or very strong support: These terms describe the strength or weakness of different ranges of a likelihood ratio (as shown in the table below). Examples of factors that affect the LR value include the amount of DNA tested, the type of mixture (for example, the number of contributors), instances when one or more of the individual's DNA alleles are not seen in the mixture, the presence of rare alleles in the mixture, and the presence of extra DNA alleles in the mixture.

Reported value	Qualitative interpretation
1	No conclusions
1 to 10	Limited support
10 to 100	Moderate support
100 to 1000	Strong support
Greater than 1000	Very strong support

Note, if the LR value is less than one, this means that the mixture is better explained if an unknown, unrelated person contributed to the mixture rather than the known person. This situation is reported as 1/LR and the qualitative terms from the table above are applied.

FST Cross Ideas

Number of contributors

How many people in the mixture?
Possibly more than 3?
Intelligent minds could disagree.
A la Carragine...
Maybe different than the Number OCME says?
Does that change the LR?

The Question Mark Guy

Who else fits into the mixture?
Would the judge have an LR above 1?

Q: "So can you tell the jury what percentage of the NYC population FST would include in this mixture?
A: "No."

- NOTE: OCME started to develop this and stopped.
- See also *Validation of a DNA mixture statistics tool incorporating allelic drop-out and drop in*, A. Mitchell, et al.

49

FST CROSS IDEAS

Suspect based statistic (v. RMP)

- Contrast FST with RMP
- You already know the answer you want
- Cognitive bias?
- Theme of defense?
  - No investigation
  - The gov't took the easiest answer and looked no deeper

DNA & FST BASICS 50

Suspect based cross – cont'd.

Look at the evidence profile – any place Δ missing?

- Drop-out = a true phenomenon but no way to prove it, so when your client not in the evidence profile, it's excused.
- Drop-in = a phenomenon that could explain any of your client's alleles, and no way to know it *wasn't* drop-in.

Make a nice big chart – highlight where Δ is NOT included in the mixture

CAUTION: Always consult an expert before attempting...

DNA & FST BASICS 51

Teach your criminalist well...

If you have an article or point you want to make, you may need to bring it up with the criminalist BEFORE your trial

- Pretrial with the criminalist (do this anyway)
- Ask them to read your materials
- Ask them to look into certain OCME protocols and/or validation studies they might not be up on
- [Remember to set this up on voir dire of expert too.]

Or don't...

- The criminalist may not understand how FST works.
- Ask your myriad questions about FST (validation, how drop in and drop out are calculated, the error rate of the quantitation (30%), no case specific false positive rate, number of contributors changing the LR, etc) and s/he won't be able to answer them.
- Burden on π? PBRD? [start in voir dire, empower your jury]

Other Cross Points

- Allele frequency – your guy has such a common profile, he could fit into any mixture!
- No one else uses this – OCME created FST and no one else has picked it up yet (also true for LCN).
- FST is not open source – we don't get to play with the code.
- FST (via OCME and π) decide Δ hypothesis – we don't get to change around the theory, insert other players, etc.

DNA Resources for Δ Bar

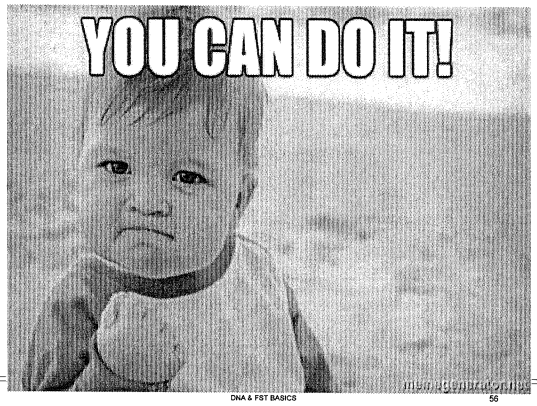
Dan Krane, Wright State University – tutorials online
- http://www.bioforensics.com/presentations/

The NIJ "DNA for the Defense Bar"
- https://www.ncjrs.gov/pdffiles1/nij/237975.pdf
- [note – this was created in June 2012, so very dated]

Christine Funk, "Dealing with DNA" – A Primer for Lawyers
- http://wispd.org/attachments/article/243/Defending%20Cases%20with%20Prosecution%20DNA%20Evidence.pdf

The Legal Aid Society, DNA Unit – DNAUnit@legal-aid.org

Butler's books



Thank You!

Jessica Goldthwaite
Legal Aid Society's DNA Unit
Michael Alperstein

Feel free to contact us at:
hcoyle@newhaven.edu
amlewis@legal-aid.org
kbwatters@aol.com

Touch DNA

Shared genetics in an alleged child abuse case – The effect of genetic relatives and extensive sampling

Case Scenario

- Juvenile victim
- Focus on pajamas and stain identification
- Accusation of forced oral sex and prosecuted based on saliva mixed with DNA of victim and defendant (family member) on sleeves (wrists)
- Confounding issue
  - Victim saliva on own pajamas due to young age might not be unusual
  - DNA and extensive sampling used to reconstruct (16 stains, shirt; 15 stains, pants = 31 total stains tested) by OCME Laboratory
  - Unexpected semen stain on hip area of the pajamas from a different DNA donor (family member)

Case Considerations

- Body fluid detection issue
  - Saliva – is it really saliva?
    - Amylase diffusion test for alpha-amylase
    - Indicates saliva but also could be from bacteria or vaginal secretions
    - Some detergents contain amylase too
    - Not 100% conclusive for saliva
    - DNA testing used to corroborate or connect source of stain to individual(s)

Table 1: Allele Values for Touch DNA samples

Locus	Victim	Uncle	Aunt	Mother	Father	Half-Brother	Shirt Stain 1	Shirt Stain 2	Pant Semen Stain 1B
D3S1358	16	16	16	16	16	15,16	16	16	15,16
D16S539	11,13	12,13	11,12	12,13	11,12	11,13	11,12,13	11,13	11,12,13
Sex	X,X	X,Y	X,X	X,X	X,X	X,Y	X,Y	X,X	X,Y
TH01	6,9	9.3	6,9	6,9.3	7,9	6,6	6,7,9	6,9	6,6,9,9.3
TPOX	8,11	11	8,9	8,11	8,11	8	8,10,11	8,11	8,8,11
CSF1PO	9,12	12	9,10	10,12	12,12	12	9,12	9,12	9,12
D7S820	8,9	8,11	8,12	8,11	12,12	8,10	8,9,12	8,9	8,9,10,11
VWA	14,17	17	14,18	14,17	17,18	17,17	14,17,18	14,17	11,14,17
FGA	22,24	24	22,24	24	22	21,24	22,24	22,24	21,22,24
D8S1179	12,15	13,13	13,15	12	14,15	12,14	12,14,15	12,15	12,14,15
D21S11	28,29	28,30	29,30	28,30	29	28	28,29	28,29	28,28
D18S51	17,18	17,19	14,17	14,18	17,20	12,18	17,18,20	17,18	12,17,18,23
D5S818	10	12,13	10,13	10,13	10,13	10,12	10,13	10	10,12,13
D13S317	11,12	11,13	8,11	11,13	8,12	11,13	8,11,12	11,12	8,11,12,13
D2S1338	17,20	20,23	17,20	20,20	17,20	20,23	17,20	17,20	17,20,21,23
D19S433	11,16.2	13,16.2	13,16.2	13,16.2	11,14	13,15.2	11,14,16.2	11,16.2	11,13,15.2,16.2

Z* refers to a missing allele or allele drop-out during processing of the DNA sample

Locus	Victim	Father	Shirt Stain 1	Shirt Stain 2
D3S1358	16	16	16	16
D16S539	11,13	11,12	11,12,13	11,13
Sex	X,X	X,X	X,Y	X,X
TH01	6,9	7,9	6,7,9	6,9
TPOX	8,11	8,10	8,10,11	8,11
CSF1PO	9,12	12,Z*	9,12	9,12
D7S820	8,9	12,Z*	8,9,12	8,9
VWA	14,17	17,18	14,17,18	14,17
FGA	22,24	22	22,24	22,24
D8S1179	12,15	14,15	12,14,15	12,15
D21S11	28,29	29	28,29	28,29
D18S51	17,18	17,20	17,18,20	17,18
D5S818	10	10,13	10,13	10
D13S317	11,12	8,12	8,11,12	11,12
D2S1338	17,20	17,20	17,20	17,20
D19S433	11,16.2	13,14	11,14,16.2	11,16.2

- Sampling Issues
  - Initially, if these are the only stains tested from wrist area of the pajama sleeves; allegation of forced oral sex appears consistent (Hp)
  - However, extensive sampling reveals a myriad of individuals (6 total) detectable either singly or as mixtures

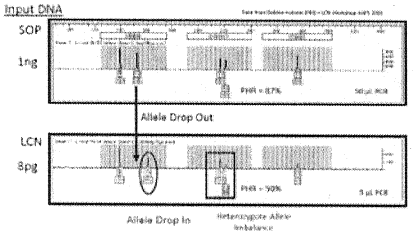
Locus	Victim	Half-Brother	Pant Semen Stain 1B
D3S1358	16	15,16	15,16
D16S539	11,13	11,13	11,12,13
Sex	X,X	X,Y	X,Y
TH01	6,9	6,6	6,6,9,9.3
TPOX	8,11	8	8,8,11
CSF1PO	9,12	12	9,12
D7S820	8,9	8,10	8,9,10,11
VWA	14,17	17,17	11,14,17
FGA	22,24	21,24	21,22,24
D8S1179	12,15	12,14	12,14,15
D21S11	28,29	28	28,29
D18S51	17,18	12,18	12,17,18,23
D5S818	10	10,12	10,12,13
D13S317	11,12	11,13	8,11,12,13
D2S1338	17,20	20,23	17,20,21,23
D19S433	11,16.2	13,15.2	11,13,15.2,16.2

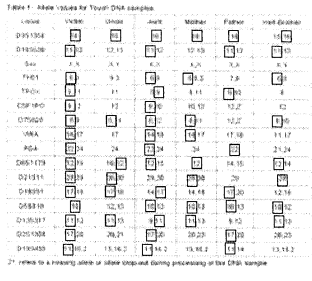
- Sampling Issues
  - Semen stain on hip of pajamas
  - Consistent with half-brother
  - Some additional alleles (7)
  - All explainable by other family members except 23 allele (drop in or unidentified individual)
    - Defense Hypothesis (Hd)
- * boxed alleles not from victim or half-brother; a third contributor or drop-in?
- ** allele counts indicate only 2 sources

FST and Genetic Relatives - Issues

- Probabilistic genotyping (PG)
- % shared alleles with the sample established mathematically
- Likelihood ratios (LR) based on detected DNA and mathematical probability that DNA is absent from the profile (drop-out) or introduced into the profile (drop-in/contamination)
- FST compares Hp to Hd for unrelated individuals due to high coincidental match rate between genetic relatives
- Coincidental matching should be considered even if LR values are high since it could be a false match by DNA with a high statistic for inclusion even for nonrelated individuals

Comparison of STR Kit Amplification SOP with LCN Using the Same DNA Donor



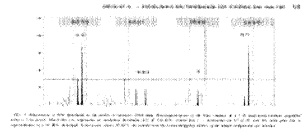


"Goldfish Theory"

- This case example with genetic relatives illustrates that too many individuals would be included with positive LR scores (likelihood ratios) to be able to distinguish one from another
- * boxed alleles indicate the first allele from the victim profile shared with other family members as coincidental match
- ** with FST estimated drop-in and drop-out rates, most or all would be anticipated to be included in the DNA sample
- *** too many fish in the pond

Requests for Electronic Data

- In New York City – not considered reports or documents
- Some success obtaining through subpoena or discovery
- Rationale for electronic data via data printouts (data below the 75 RFU)
- Data below the analytical threshold can clarify number of contributors



Bregu et al. 2013. JFS 58(1): 120-129.

Case Outcome

- Sampling of the clothing - overall extensive overview of stains (31)
- Detection of body fluids – ambiguity of saliva identification
- Detection of the touch DNA – large number of touch DNA contributors (6)
- Overall interpretation toward a theoretical reconstruction of events
- Jury believed that the single semen stain was the more likely assailant compared to the oral sex allegation (defendant)
- Case acquittal

Acknowledgements

- Attorney Alfred Dorfman

- Contact information
  - Heather Coyle
  - Associate Professor – University of New Haven
  - Identacode Consulting LLC - consultant
  - Tel. 914 438 8223
  - Email: Hcoyle@newhaven.edu

Touch DNA

Shared genetics in an alleged child abuse case – The effect of genetic relatives and extensive sampling

Case Scenario

- Juvenile victim
- Focus on pajamas and stain identification
- Accusation of forced oral sex and prosecuted based on saliva mixed with DNA of victim and defendant (family member) on sleeves (wrists)
- Confounding issue
  - Victim saliva on own pajamas due to young age might not be unusual
  - DNA and extensive sampling used to reconstruct (16 stains, shirt; 15 stains, pants = 31 total stains tested) by OCME Laboratory
  - Unexpected semen stain on hip area of the pajamas from a different DNA donor (family member)

Case Considerations

- Body fluid detection issue
  - Saliva – is it really saliva?
    - Amylase diffusion test for alpha-amylase
    - Indicates saliva but also could be from bacteria or vaginal secretions
    - Some detergents contain amylase too
    - Not 100% conclusive for saliva
    - DNA testing used to corroborate or connect source of stain to individual(s)

Table 1: Allele Values for Touch DNA samples.

Locus	Victim	Uncle	Aunt	Mother	Father	Half-Brother	Shirt Stain 1	Shirt Stain 2	Pant Semen Stain 15
D3S1358	16	16	16	16	16	15,16	16	16	15,16
D16S539	11,13	12,13	11,12	12,13	11,12	11,13	11,12,13	11,13	11,12,13
Sex	X,X	X,Y	X,X	X,X	X,X	X,Y	X,Y	X,X	X,Y
TH01	6,9	9.3	6,9	6,9.3	7,9	6,8	6,7,9	6,9	6,8,9,9.3
TPOX	9,11	11	8,9	8,11	9,10	8	9,10,11	9,11	8,9,11
CSF1PO	9,12	12	9,10	10,12	12,Z*	12	9,12	9,12	9,12
D7S820	8,9	8,11	8,12	8,11	12,Z*	8,10	8,9,12	8,9	8,9,10,11
VWA	14,17	17	14,18	14,17	17,18	11,17	14,17,18	14,17	11,14,17
FGA	22,24	24	22,24	24	22	21,24	22,24	22,24	21,22,24
D8S1179	12,15	10,12	12,15	12	14,15	12,14	12,14,15	12,15	12,14,15
D21S11	28,29	28,30	29,30	28,30	29	28	28,29	28,29	28,29
D18S51	17,18	17,18	14,17	14,18	17,20	12,18	17,18,20	17,18	12,17,18,23
D5S818	10	12,13	10,13	10,13	10,13	10,12	10,13	10	10,12,13
D13S317	11,12	11,13	9,11	11,13	9,12	11,13	9,11,12	11,12	9,11,12,13
D2S1338	17,20	20,21	17,20	20,23	17,20	20,23	17,20	17,20	17,20,21,23
D19S433	11,16.2	13,16.2	11,16.2	13,16.2	11,14	13,15.2	1,14,16.2	11,16.2	11,13,15.2,16.2

Z*, refers to a missing allele or allele drop-out during processing of the DNA sample

Locus	Victim	Father	Shirt Stain 1	Shirt Stain 2
D3S1358	16	16	16	16
D16S539	11,13	11,12	11,12,13	11,13
Sex	X,X	X,X	X,Y	X,X
TH01	6,9	7,9	6,7,9	6,9
TPOX	9,11	9,10	9,10,11	9,11
CSF1PO	9,12	12,Z*	9,12	9,12
D7S820	8,9	12,Z*	8,9,12	8,9
VWA	14,17	17,18	14,17,18	14,17
FGA	22,24	22	22,24	22,24
D8S1179	12,15	14,15	12,14,15	12,15
D21S11	28,29	29	28,29	28,29
D18S51	17,18	17,20	17,18,20	17,18
D5S818	10	10,13	10,13	10
D13S317	11,12	9,12	9,11,12	11,12
D2S1338	17,20	17,20	17,20	17,20
D19S433	11,16.2	11,14	1,14,16.2	11,16.2

- <u>Sampling Issues</u>
  - Initially, if these are the only stains tested from wrist area of the pajama sleeves; allegation of forced oral sex appears consistent (H_p)
  - However, extensive sampling reveals a myriad of individuals (6 total) detectable either singly or as mixtures

Locus	Victim	Half-Brother	Pant Semen Stain 15
D3S1358	16	15,16	15,16
D16S539	11,13	11,13	11,12,13
Sex	X,X	X,Y	X,Y
TH01	6,9	6,8	6,8,9,9.3
TPOX	9,11	8	8,9,11
CSF1PO	9,12	12	9,12
D7S820	8,9	8,10	8,9,10,11
VWA	14,17	11,17	11,14,17
FGA	22,24	21,24	21,22,24
D8S1179	12,15	12,14	12,14,15
D21S11	28,29	28	28,29
D18S51	17,18	12,18	12,17,18,23
D5S818	10	10,12	10,12,13
D13S317	11,12	11,13	9,11,12,13
D2S1338	17,20	20,23	17,20,21,23
D19S433	11,16.2	13,15.2	11,13,15.2,16.2

- <u>Sampling Issues</u>
  - Semen stain on hip of pajamas
  - Consistent with half-brother
  - Some additional alleles (7)
  - All explainable by other family members except 23 allele (drop in or unidentified individual)
    - Defense Hypothesis (H_d)
- * boxed alleles not from victim or half-brother; a third contributor or drop-in?
- ** allele counts indicate only 2 sources

FST and Genetic Relatives - Issues

- Probabilistic genotyping (PG)
- % shared alleles with the sample established mathematically
- Likelihood ratios (LR) based on detected DNA and mathematical probability that DNA is absent from the profile (drop-out) or introduced into the profile (drop-in/contamination)
- FST compares Hp to Hd for unrelated individuals due to high coincidental match rate between genetic relatives
- Coincidental matching should be considered even if LR values are high since it could be a false match by DNA with a high statistic for inclusion even for nonrelated individuals

Comparison of STR Kit Amplification SOP with LCN Using the Same DNA Donor

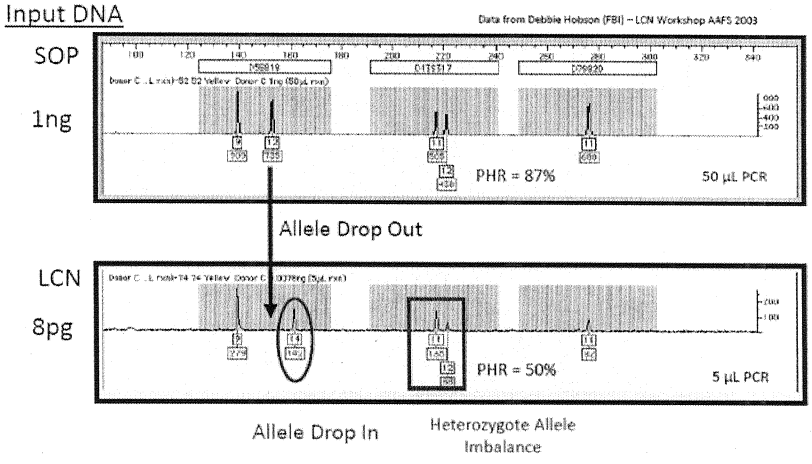


Table 1: Allele Values for Touch DNA samples.

Locus	Victim	Uncle	Aunt	Mother	Father	Half-Brother
D3S1358	16	16	16	16	16	15,16
D16S539	11,13	12,13	11,12	12,13	11,12	11,13
Sex	X,X	X,Y	X,X	X,X	X,X	X,Y
TH01	6,9	9.3	6,9	6,9.3	7,9	6,8
TPOX	9,11	11	8,9	8,11	9,10	8
CSF1PO	9,12	12	9,10	10,12	12,Z*	12
D7S820	8,9	8,11	8,12	8,11	12,Z*	8,10
VWA	14,17	17	14,18	14,17	17,18	11,17
FGA	22,24	24	22,24	24	22	21,24
D8S1179	12,15	10,12	12,15	12	14,15	12,14
D21S11	28,29	28,30	29,30	28,30	29	28
D18S51	17,18	17,18	14,17	14,18	17,20	12,18
D5S818	10	12,13	10,13	10,13	10,13	10,12
D13S317	11,12	11,13	9,11	11,13	9,12	11,13
D2S1338	17,20	20,21	17,20	20,23	17,20	20,23
D19S433	11,16.2	13,16.2	11,16.2	13,16.2	11,14	13,15.2

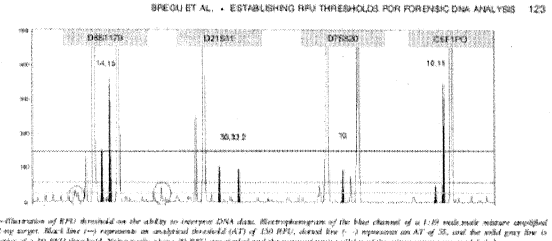
Z*, refers to a missing allele or allele drop-out during processing of the DNA sample

"Goldfish Theory"

- This case example with genetic relatives illustrates that too many individuals would be included with positive LR scores (likelihood ratios) to be able to distinguish one from another
- * boxed alleles indicate the first allele from the victim profile shared with other family members as coincidental match
- ** with FST estimated drop-in and drop-out rates, most or all would be anticipated to be included in the DNA sample
- *** too many fish in the pond

Requests for Electronic Data

- In New York City – not considered reports or documents
- Some success obtaining through subpoena or discovery
- Rationale for electronic data via data printouts (data below the 75 RFU)
- Data below the analytical threshold can clarify number of contributors



Bregu et al. 2013. JFS 58(1): 120-129.

Case Outcome

- Sampling of the clothing - overall extensive overview of stains (31)
- Detection of body fluids – ambiguity of saliva identification
- Detection of the touch DNA – large number of touch DNA contributors (6)
- Overall interpretation toward a theoretical reconstruction of events
- Jury believed that the single semen stain was the more likely assailant compared to the oral sex allegation (defendant)
- Case acquittal

Acknowledgements

- Attorney Alfred Dorfman

- Contact information
  - Heather Coyle
  - Associate Professor – University of New Haven
  - Identacode Consulting LLC - consultant
  - Tel. 914 438 8223
  - Email: Hcoyle@newhaven.edu

Forensic Science International: Genetics 6 (2012) 749–761

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

Department of Forensic Biology, Office of Chief Medical Examiner of The City of New York, 421 E 26th Street, New York, NY 10016, United States

ARTICLE INFO

Article history:
Received 25 May 2012
Received in revised form 8 August 2012
Accepted 9 August 2012

Keywords:
DNA mixture analysis
Likelihood ratio
Allelic drop-out
Allelic drop-in
Low template DNA

ABSTRACT

DNA mixture analysis is a current topic of discussion in the forensics literature. Of particular interest is how to approach mixtures where allelic drop-out and/or drop-in may have occurred. The Office of Chief Medical Examiner (OCME) of The City of New York has developed and validated the Forensic Statistical Tool (FST), a software tool for likelihood ratio analysis of forensic DNA samples, allowing for allelic drop-out and drop-in. FST can be used for single source samples and for mixtures of DNA from two or three contributors, with or without known contributors. Drop-out and drop-in probabilities were estimated empirically through analysis of over 2000 amplifications of more than 700 mixtures and single source samples. Drop-out rates used by FST are a function of the Identifiler® locus, the quantity of template DNA amplified, the number of amplification cycles, the number of contributors to the sample, and the approximate mixture ratio (either unequal or approximately equal). Drop-out rates were estimated separately for heterozygous and homozygous genotypes. Drop-in rates used by FST are a function of number of amplification cycles only.

FST was validated using 454 mock evidence samples generated from DNA mixtures and from items handled by one to four persons. For each sample, likelihood ratios (LRs) were computed for each true contributor and for each profile in a database of over 1200 non-contributors. A wide range of LRs for true contributors was obtained, as true contributors' alleles may be labeled at some or all of the tested loci. However, the LRs were consistent with OCME's qualitative assessments of the results. The second set of data was used to evaluate FST LR results when the test sample in the prosecution hypothesis of the LR is not a contributor to the mixture. With this validation, we demonstrate that LRs generated using FST are consistent with, but more informative than, OCME's qualitative sample assessments and that LRs for non-contributors are appropriately assigned.

1. Introduction

The development of increasingly sensitive forensic STR-based DNA testing techniques has expanded the application of DNA typing beyond biological fluids to a wide variety of evidence items such as touched weapons or touched clothes, and forensic laboratories now routinely obtain STR profiles from degraded DNA samples. Increased sensitivity in STR profiling may be accomplished by one of several methods, such as increased PCR cycle numbers [1–4], nested PCR [5,6] and purification of the PCR product [7]. With increased cycle number, full STR profiles can be obtained from 25 to 50 pg of DNA; partial profiles may be obtained from lower quantities of DNA [8,9]. However, stochastic events in early PCR cycles can result in lost alleles (drop-out) and increased sensitivity can produce extraneous alleles (drop-in) [2]. Due to a higher occurrence of allelic drop-out and drop-in with low template or degraded samples, relative to high template, robust samples, the DNA commission of the International Society of Forensic Genetics (ISFG) cautions that standard STR analysis methods should not be directly applied to low template samples [10]. At the Office of Chief Medical Examiner (OCME) of The City of New York, samples containing up to 100 pg of template DNA per amplification are considered low template (LT) DNA and are amplified in triplicate for 31 PCR cycles. Samples containing at least 100 pg of template DNA per amplification are considered high template (HT) DNA and are amplified once or in duplicate for 28 PCR cycles.

The standard statistic calculated when evidentiary and exemplar STR profiles are identical is the random match probability. This can be used for single source evidentiary profiles and for mixtures when individual contributors' profiles can be deduced. Two methods, Combined Probability of Inclusion (CPI), also known as

* Corresponding author.
E-mail address: AAMitchell@ocme.nyc.gov (A.A. Mitchell).

1872-4973/$ – see front matter © 2012 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.fsigen.2012.08.007

Random Man Not Excluded (RMNE), and likelihood ratio (LR), are commonly used to quantify the statistical weight of mixed DNA profiles when the profiles of individual contributors cannot be deconvoluted. The Scientific Working Group on DNA Analysis Methods (SWGDAM), which is organized by the FBI, published guidelines for interpretation of DNA mixtures in 2010 [11]. The guidelines specify that a statistic should accompany all positive associations between individuals and evidence sample mixtures and that either RMNE or LR may be used. The DNA commission of the ISFG recommends using the LR for such mixtures [10], as more available data are utilized and allelic drop-out and drop-in can be explicitly incorporated in the calculation. That said, RMNE does not require specification of the number of contributors to a mixture, whereas the LR does, and the RMNE calculation is more intuitive; therefore, RMNE can be easier to explain to a jury than the LR.

A likelihood ratio is a ratio of two probabilities. In the numerator is the probability of a set of data conditional on one hypothesis; in the denominator is the probability of the same set of data conditional on a mutually exclusive hypothesis. For forensic DNA applications, the data are the alleles found in the evidence sample, the hypothesis in the numerator is that of the prosecutor (H_p), and the hypothesis in the denominator is that of the defense (H_d). The LR is a measure of the support for the prosecution hypothesis relative to that of the defense. If the LR is greater than one, H_p is better supported by the data than H_d; if the LR is less than one, H_d is better supported by the data than H_p. For single source evidence profiles, the H_p is typically that a particular suspect is the source of the crime scene DNA and H_d is that an unknown, unrelated person is the source of that DNA. For two-person evidence profiles, there are more options for H_p and H_d. First, H_p could be that the crime scene sample represents a mixture of DNA from the suspect and an unknown, unrelated person while H_d could be that the crime scene sample represents a mixture of DNA from two unknown, unrelated people. Alternatively, the prosecutor may assert that the sample represents a mixture of DNA from the suspect and a known person, for example a victim, and the defense may assert that it represents a mixture of DNA from the known person and an unknown, unrelated person or from two unknown, unrelated persons. For three-person evidence profiles, there are even more possibilities, as up to two known contributors may be included in either or both hypotheses. The number of contributors in the two hypotheses need not be the same and a known contributor that is included in either the numerator or the denominator does not need to be included in the other.

OCME has developed and validated the Forensic Statistical Tool (FST), a software tool for LR analysis of low and high template DNA profiles for single source samples and for mixtures of DNA from two or three contributors allowing for drop-out and drop-in. The first step of this development was to empirically determine drop-out rates for each locus, genotype and DNA template-quantity, and drop-in rates for HT-DNA and LT-DNA amplification conditions. These rate estimates were generated from duplicate or triplicate amplifications of over 700 samples, totaling more than 2000 amplifications. These samples include single source samples, as well as two- and three-person deliberate mixtures. Drop-out rates were estimated separately for each Identifiler® locus (Applied Biosystems, Carlsbad, CA), for single source template quantities ranging from 6.25 pg to 500 pg and for two- and three-person mixtures with template quantities ranging 25–500 pg. Separate estimates were obtained for the probability of partial heterozygous drop-out, complete heterozygous drop-out, and complete homozygous drop-out. Drop-out rates were estimated separately for deducible and non-deducible mixtures. Drop-in rates were estimated separately for 28 and 31 cycle amplification.

Empirically estimated drop-out and drop-in rates were incorporated into likelihood ratio frameworks including the appropriate number of contributors (one to three). For mixtures, LRs were formulated with and without assumed contributors ("knowns") in addition to the test profile, usually the suspect profile. For example, for three-person profiles, the LR can be set up as (1) test plus two unknowns versus three unknowns, (2) test plus known plus unknown versus known plus two unknowns, or (3) test plus two knowns versus two knowns plus one unknown.

To date, several other methods and software for LR analysis allowing for drop-out and drop-in have been published. These include True Allele, LoComatioN, the LRMix module within the Forensim package, and LikeLTD. True Allele [12–15] uses Markov chain Monte Carlo (MCMC) to deconvolute a DNA mixture, after which a LR for any set of one to three contributors specified by the user can be computed. In the calculation of the LR, a heuristic penalty for drop-out is applied; for every instance of drop-out that would be required to explain the mixture with a given set of hypothesized contributors, the resulting likelihood ratio is reduced by two orders of magnitude.

LoComatioN [16,17], LRMix [18], likeLTD [19], and FST do not deconvolute DNA mixtures, but simply compute a LR for scenarios specified by the user, allowing for mismatches between contributors' profiles and the DNA alleles labeled in the mixtures. The mismatches are accounted for by incorporating drop-out and drop-in probabilities in the LR calculation. While FST uses empirically determined drop-out and drop-in rates, LoComatioN and Forensim require the user to specify drop-out and drop-in probabilities. Forensim then calculates the LR for a range of drop-out rates and displays the results graphically. LikeLTD finds the drop-out probabilities and mixture ratios that maximize the likelihood under H_p and H_d. One difference between FST and the other three programs is that they adjust for intra- and inter-individual correlation in genotypes [16,20], while FST adjusts only for intra-individual correlation with a correction to the expected population frequencies of homozygous genotypes [11,20]. In addition, LoComatioN and LRMix model allelic drop-in using estimated allele frequencies, whereas FST does not consider the identity of drop-in alleles, simply that drop-in of one allele or of two or more alleles has occurred. In addition, other models [16–18] specifically exclude stutter from the definition of drop-in, whereas FST's drop-in definition includes stutter as well as extraneous peaks that are not in stutter position (see Section 2.2.3).

The FST software was validated by running the program 516 times using 454 mock evidence profiles, including single source samples, deliberate mixtures of DNA from two or three individuals, and samples developed from touched objects handled by one to four known individuals. These samples were composed of various combinations of 85 donors, representing a mix of Caucasian, Asian, African American, and Hispanic ethnicities. FST performance was first evaluated using each true DNA contributor's profile as the test profile and all results were compared to previously generated qualitative conclusions. Other validation steps explored the effect on the FST calculation if the source of the test sample did not contribute to the DNA mixture in question. This was achieved by running FST using each individual in a population database of 1246 non-contributors, collected by OCME and NIST [21] as test profiles. In total, more than 557,000 test runs of the program were performed as part of this validation. We demonstrate that FST is an effective tool for assigning weight of evidence to forensic DNA profiles using likelihood ratios.

2. Methods

2.1. Statistical methods

As a matter of background, in a conventional forensic DNA LR framework, the probability of a crime scene DNA profile (G) is

computed conditional on two competing hypotheses: that of the prosecution (H_p) and that of the defense (H_d). The LR is the ratio of these two probabilities: $LR = (Pr(G|H_p)/Pr(G|H_d))$. In a typical simple scenario, the prosecution asserts that the crime scene DNA belongs to the suspect (H_p:S) and the defense asserts that the crime scene DNA belongs to an unknown, unrelated individual (H_d:U). If the crime scene profile matches the suspect profile, then $Pr(G|H_p) = 1.0$ and $Pr(G|H_d) = P$, where P is the estimated population frequency of the crime scene profile (i.e., the random match probability, RMP). Thus, in this scenario, LR = (1/RMP) when the evidence profile matches the suspect profile. When a crime scene sample reflects a mixture, the conventional LR can also be used by specifying the appropriate H_p and H_d. For example, the prosecution may assert that the mixture includes the victim and suspect (H_p:$V + S$), while the defense may assert the mixture includes the victim and an unknown, unrelated individual (H_d:$V + U$).

Multiple replicate amplifications of an evidentiary sample can be considered within a LR framework. Let $\mathbf{R} = R_1, R_2, \ldots, R_n$ represent the alleles observed in amplification replicates 1 through n at a single locus in an evidentiary sample. At OCME, normally, $n = 3$ for LT-DNA samples and $n = 1$ or 2 for HT-DNA samples. The replicate data are used to compute $LR = (Pr(\mathbf{R}|H_p)/Pr(\mathbf{R}|H_d))$.

We have incorporated allelic drop-out and drop-in into the LR. A critical step in this process is to consider all possible genotypes for the unknown contributor(s) in the denominator (as well as in the numerator for more complex scenarios). If x distinct alleles are observed at a locus in the evidentiary profile, there are $m = (x(x+1)/2) + x + 1$ values comprising the set of possible genotypes of an unknown contributor. This calculation treats all unobserved alleles as a single 'other' allele. That is, an unknown contributor's genotype at the locus could include any pair wise combination of the observed alleles and the unobserved 'other' allele.

If the prosecution hypothesis is that the DNA in the evidence sample belongs to the suspect (H_p:S) and the defense hypothesis is that the DNA in the evidence sample belongs to an unknown, unrelated person (H_d:U), the likelihood ratio is formulated as:

$$LR = \frac{Pr(\boldsymbol{R}|S)}{\sum_{j=1}^{m} Pr(\mathbf{R}|U = G_j)Pr(U = G_j)},$$

where S represents the suspect's alleles, U represents the alleles of an unknown contributor, and G_j represents the jth possibility for the genotype of the unknown contributor. $Pr(U = G_j)$ is the expected population frequency of G_j, including a θ correction for intra-individual (but not inter-individual) population substructure, applied to homozygous genotypes, as described in Recommendation 4.1 of the second National Research Council Report [20] and the 2010 SWGDAM mixture interpretation guidelines [11].

If the prosecution hypothesis is that the DNA in the evidence sample belongs to the suspect and the victim (H_p:$S + V$) and the defense hypothesis is that the DNA in the evidence sample belongs to the victim and an unknown, unrelated person (H_d:$V + U$), the likelihood ratio is formulated as:

$$LR = \frac{Pr(\mathbf{R}|S, V)}{\sum_{j=1}^{m} Pr(\mathbf{R}|V; U = G_j)Pr(U = G_j)}.$$

If the prosecution hypothesis is that the DNA in the evidence sample belongs to the suspect and an unknown person (H_p:$S + U_1$) and the defense hypothesis is that the DNA in the evidence sample belongs to two unknown persons, (H_d:$U_2 + U_3$), the likelihood ratio is formulated as:

$$LR = \frac{\sum_{i=1}^{m} Pr(\mathbf{R}|S; U_1 = G_i)Pr(U_1 = G_i)}{\sum_{i=1}^{m}\sum_{j=1}^{m} Pr(\mathbf{R}|U_2 = G_i, U_3 = G_j)Pr(U_2 = G_i)Pr(U_3 = G_j)}.$$

Table 1
Numerator and denominator options supported by FST. "Known" refers to an elimination profile from an individual who is assumed to be a contributor to the evidence sample. "Comparison" refers to the comparison profile of interest (often the suspect). "Unknown" refers to a randomly selected individual from a population of individuals that are unrelated to the known, comparison, or one another.

Option	Numerator (prosecution hypothesis)	Denominator (defense hypothesis)
1	Comparison	Unknown
2	Comparison + known	Known + unknown
3	Comparison 1 + comparison 2	2 unknowns
4	Comparison + unknown	2 unknowns
5	Comparison + 2 unknowns	3 unknowns
6	Comparison + known + unknown	2 unknowns + known
7	Comparison + 2 known	Unknown + 2 knowns
8	Comparison + 3 known	Unknown + 3 knowns
9	Comparison + 2 knowns + unknown	2 unknowns + 2 knowns
10	Comparison + known + 2 unknowns	3 unknowns + known
11	Comparison + 3 unknowns	4 unknowns

All of the pairs of prosecution and defense hypotheses shown in Table 1 are formulated similarly.

Drop-out and drop-in rates are incorporated into $Pr(R_1, R_2, \ldots, R_n|S)$ and $Pr(R_1, R_2, \ldots, R_n|U = G_j)$. Separate parameters are used for the rate of partial drop-out of heterozygotes, complete drop-out of heterozygotes, and complete drop-out of homozygotes. Drop-out rates were estimated empirically as a function of the locus, the quantity of template DNA, the number of contributors to the sample, and the approximate mixture ratio, i.e., approximately one-to one (not deducible) or not one-to-one (deducible). For single source samples, drop-out rates were estimated for eight template DNA quantities, ranging from 6.25 to 500 pg. For mixtures, drop-out rates were estimated for six template DNA quantities, ranging from 25 to 500 pg. Drop-in rates were estimated separately for LT-DNA (less than or equal to 100 pg per reaction) and HT-DNA (greater than or equal to 100 pg per reaction) amplifications. Note that 100 pg samples were typed under LT-DNA and HT-DNA conditions and two sets of drop-out rates were established for this quantity of template DNA. In total, more than 2000 amplifications were performed for the estimation of drop-out and drop-in rates.

Listed below is a single-locus sample calculation of the LR for a single source evidence profile with H_p: suspect and H_d: unknown. The suspect's alleles at this locus are 11, 14 and the evidence alleles are replicate 1:11; replicate 2:11, 14; replicate 3:11, 14, 15. The drop-out and drop-in that would be required to explain the evidence if the prosecution hypothesis is correct (i.e., the factors in the numerator) are shown in Table 2. Conditional on the evidence sample DNA originating from the suspect, there would be one drop-out from a heterozygous locus and no drop-in in the first replicate, no drop-out from a heterozygous locus and no drop-in in the second replicate, no drop-out from a heterozygous locus and one drop-in in the third replicate. Thus, the numerator in this example is

$$[Pr(D_1)Pr(C_0)] \times [Pr(D_0)Pr(C_0)] \times [Pr(D_0)Pr(C_1)],$$

Table 2
Factors used in numerator (prosecution hypothesis) in example. D_0 and D_1 represent drop-out of zero and one alleles, respectively, from a heterozygous locus; C_0 and C_1 represent drop-in of zero and one alleles, respectively.

Replicate	Alleles labeled	Counts		Factor used	
		Drop-out	Drop-in	Drop-out	Drop-in
1	11	Yes (1)	No	D_1	C_0
2	11, 14	No	No	D_0	C_0
3	11, 14, 15	No	Yes (1)	D_0	C_1

Table 3
Factors used in denominator (defense hypothesis) in example. Variables are defined as in Table 2 plus D_2 which represents drop-out of both alleles from a heterozygous locus, D_{H0} and D_{H1}, which represent no drop-out and complete drop-out, respectively, from a homozygous locus, and C_{2+} which represents drop-in of two or more alleles.

Unknown Person's genotype	Genotype frequency estimate	Replicate 1		Replicate 2		Replicate 3	
		Drop-out	Drop-in	Drop-out	Drop-in	Drop-out	Drop-in
11, 11	P_{11}^2	D_{H0}	C_0	D_{H0}	C_1	D_{H0}	C_{2+}
11, 14	$2\ P_{11}\ P_{14}$	D_1	C_0	D_0	C_0	D_0	C_1
11, 15	$2\ P_{11}\ P_{15}$	D_1	C_0	D_1	C_1	D_0	C_1
11, *w*	$2\ P_{11}\ P_w$	D_1	C_0	D_1	C_1	D_1	C_{2+}
14, 14	P_{14}^2	D_{H1}	C_1	D_{H0}	C_1	D_{H0}	C_{2+}
14, 15	$2\ P_{14}\ P_{15}$	D_2	C_1	D_1	C_1	D_0	C_1
14, *w*	$2\ P_{14}\ P_w$	D_2	C_1	D_1	C_1	D_1	C_{2+}
15, 15	P_{15}^2	D_{H1}	C_1	D_{H1}	C_{2+}	D_{H0}	C_{2+}
15, *w*	$2\ P_{15}\ P_w$	D_2	C_1	D_2	C_{2+}	D_1	C_{2+}
w, *w*	P_w^2	D_{H1}	C_1	D_{H1}	C_{2+}	D_{H1}	C_{2+}

where D_0 and D_1 represent no drop-out and partial drop-out, respectively, from heterozygous loci and C_0 and C_1 represent drop-in of zero and one alleles, respectively.

For the denominator, FST allows the unknown contributor to have a genotype made up of any pair wise combination of the alleles observed in the replicates and any other allele, represented by the letter *w*. In this example, the unknown contributor could have any of the following genotypes: (11, 11), (11, 14), (11, 15), (11, *w*), (14, 14), (14, 15), (14, *w*), (15, 15), (15, *w*), (*w*, *w*), where *w* represents any allele other than 11, 14, or 15. For each possible unknown contributor genotype, FST determines what type of drop-out and drop-in, if any, would be required to explain the evidence profile if the true source of the DNA is an unknown person with that particular genotype. This is analogous to the analysis done using the suspect genotype in the numerator, shown above. The drop-out and drop-in factors for the denominator of this example are shown in Table 3. Using the notation from Table 3 and allowing the unknown contributor to have any of the genotypes listed above, the denominator in this example is shown below. The denominator includes a single term for each possible unknown contributor genotype. Each term includes a factor representing the population frequency estimate of the genotype and drop-out and drop-in probabilities for each replicate. Without incorporating the adjustment to the expected frequency of homozygote genotypes, the denominator is

$$\begin{aligned} &P_{11}^2 \times [Pr(D_{H0})Pr(C_0)] \times [Pr(D_{H0})Pr(C_1)] \times [Pr(D_{H0})Pr(C_{2+})] \\ &\quad + 2P_{11}P_{14} \times [Pr(D_1)Pr(C_0)] \times [Pr(D_0)Pr(C_0)] \times [Pr(D_0)Pr(C_1)] \\ &\quad + 2P_{11}P_{15} \times [Pr(D_1)Pr(C_0)] \times [Pr(D_1)Pr(C_1)] \\ &\quad \times [Pr(D_0)Pr(C_1)] \cdots P_w^2 \times [Pr(D_{H1})Pr(C_1)] \times [Pr(D_{H1})Pr(C_{2+})] \\ &\quad \times [Pr(D_{H1})Pr(C_{2+})] \end{aligned}$$

D_{H0} and D_{H1} represent no drop-out and total drop-out, respectively, from homozygous loci. Because these events are mutually exclusive and represent all possible outcomes, $Pr(D_{H0}) + Pr(D_{H1}) = 1.0$. D_0, D_1, and D_2 represent no drop-out, partial drop-out, and total drop-out, respectively, from heterozygous loci. Again, these events are mutually exclusive and represent all possible outcomes, so $Pr(D_0) + Pr(D_1) + Pr(D_2) = 1.0$. C_0, C_1, and C_{2+} represent drop-in of zero alleles, one allele, and two or more alleles, respectively, in a single amplification at a single locus. These events are also mutually exclusive and represent all possible outcomes, so $Pr(C_0) + Pr(C_1) + Pr(C_{2+}) = 1.0$.

The method described above has been implemented in a C# program with a web interface. The evidence profile, the comparison profile and profiles of known contributors (if applicable) can be entered manually or uploaded. The user selects prosecution and defense hypotheses, specifies whether or not a mixture is deducible, and enters the quantity of template DNA amplified in each reaction (up to three amplifications per evidence profile). For evidence samples with DNA template quantities that fall between those used for drop-out rate estimation, FST interpolates to determine the appropriate rate to use. In order to be conservative, FST uses the drop-out rate estimate minus one standard deviation for each locus, template DNA quantity, number of contributors, and ratio for mixed samples. The final program was tested by performing hand calculations to verify the expected result based on the algorithms explained above and the user set sample characteristics.

FST can perform a comparison between an evidence profile and a single test profile (with or without assumed contributors) or an evidence profile and a database of DNA profiles, such as a quality control database to check for possible sample contamination. We used this capability to determine the distribution of the LR when a non-contributor was included as the test sample. A database of 1246 population samples was used. This database included 546 profiles (115 Caucasian, 125 African American, 151 Hispanic, and 155 Asian contributors) collected locally and 700 profiles (302 Caucasian, 258 African American, and 140 Hispanic contributors) obtained through NIST's STRbase website [21].

2.2. Laboratory methods

2.2.1. DNA sample collection

DNA extracted from buccal swabs given by volunteer donors was used to generate defined DNA mixtures to be used for drop-out rate determination. Extraction, quantification, amplification, separation, and analysis protocols were identical to those used for the mock evidence samples, as described below. To ensure that mixture ratios were accurate, buccal swab extracts were quantified three times in triplicate and the average of the nine values was used for generating mixtures. Extracts were amplified within one week of quantitation or samples were re-quantified.

Mock evidence profiles were developed from mixtures of DNA extracted from post mortem blood or volunteer buccal specimens and from touched items. The test set of mock evidence samples included 98 two-person mixtures from buccal specimens, 102 three-person mixtures from buccal specimens, three two-person mixtures from post mortem blood, 97 two-person mixtures from buccal specimens with either one or both contributors degraded with UV C irradiation, 91 three-person mixtures from post mortem blood, 15 touched items handled by one person, 19 touched items handled by two persons, 39 touched items handled by three persons, and 31 touched items handled by four persons. In total, 350 mixtures and 104 touched items were included in the validation.

Touched items, either cleaned with 10% bleach followed by water and 70% ethanol or not cleaned, were handled by one or

consecutively by two, three, or four members of the laboratory as indicated. Touched samples were swabbed with the NYC OCME's swab (patent pending) pre-moistened with 0.01% SDS. Swabbing was performed with a light touch and, if applicable, along the grain of the item. If needed, more than one swab per item or section of an item was used. DNA was extracted from swabs within 1–2 days of collection.

2.2.2. DNA typing

All work was performed using contamination control measures as described by Caragine et al. [9]. Blood and buccal specimens and touched-item swabs were extracted with Chelex, Qiagen M48, or OCME's High Sensitivity Extraction protocol [22] using 0.05% SDS. The digest was purified twice and concentrated with a Microcon® 100 (Millipore, Billerica, MA, USA) pretreated with 1 μg of fish sperm DNA and eluted with 20 μL of irradiated water. Two μL of sample was measured on the Rotor-Gene Q 3000® (Qiagen, Valencia, CA, USA) using an Alu-based real time PCR assay based on the method described by Nicklas and Buel [23], with the exception of the addition of 0.3 μL of 100X SYBR green I (Molecular Probes) and 0.525 mg/mL BSA in a 25 μL reaction volume. With the exception of one set of the blood mixtures, prior to making mixtures, each contributor's DNA extract was measured in triplicate, three times. For one set of the blood mixtures, equivalent sized punches were taken from bloodstain cards and were combined in ratios such as 2:2:1, for example. The "punches" were extracted together and the resultant extract was measured.

Depending on the target quantity, samples were amplified for 28 cycles (≥100 pg, HT-DNA; ID28) or 31 cycles (≤100 pg, LT-DNA; ID31) using the AmpFlSTR® Identifiler® PCR Amplification Kit (Applied Biosystems, Carlsbad, CA). The manufacturer's recommendations were used with the exception of a two minute annealing time, a half-reaction volume, and 2.5 U of AmpliTaq Gold®. ID31 samples were amplified in triplicate and replicates were termed "a", "b", and "c". All other samples were amplified in duplicate.

Amplification products were separated by capillary electrophoresis on an Applied Biosystems (Carlsbad, CA) 3130xl Genetic Analyzer as described previously [9]. Injection parameters were modified based on DNA input and amplification parameters as follows: "ID31 low (L)": 100 pg and 75 pg samples – 1 kV for 22 s, "ID31 normal (N)": 50 pg and 25 pg samples – 3 kV for 20 s, or "ID31 high (H)": 12.5 pg and 6.25 pg samples – 6 kV for 30 s. For samples amplified using Identifiler® 28 cycles (ID28), 3 μL of PCR product were added to HIDI formamide and GeneScan® 500 LIZ® Size Standard (Applied Biosystems, Carlsbad, CA) for a total volume of 30 μL and unless otherwise indicated, injected according to the following parameters: "ID28 high (IR)": samples 200 pg and below – 1 kV for 22 s, or "ID28 normal (I)": samples >200 pg – 5 kV for 20 s.

Data were collected with non-variable binning that collects the highest signal for each color as opposed to variable binning that compensates for low red peak heights. Analysis was performed with Applied Biosystems GeneScan® and Genotyper® or GeneMapper® software with a 75 RFU detection threshold, a 251 baseline window size, a 10% general filter which removes peaks that are less than 10% of the highest peak at a locus, and the OCME's standard locus-specific stutter filters for Identifiler® 28 cycles and for Identifiler® 31 cycles. If multiple injections of a given PCR product were generated for a sample, for each locus the injection or amplification that showed the greatest number of labeled peaks that were not off scale or over saturated was used. For LT-DNA samples, a consensus profile was recorded which contained alleles that repeated in at least two of the three replicates. Additional details on OCME's LT-DNA protocols can be found in Caragine et al. [9].

2.2.3. Drop-out and drop-in rate determination

Drop-out rates were estimated for single source samples with 6.25 pg, 12.5 pg, 25 pg, 50 pg, 100 pg, 150 pg, 250 pg, and 500 pg of template DNA and for two- and three-person mixtures with 25 pg, 50 pg, 100 pg, 150 pg, 250 pg, and 500 pg of template DNA. Mixture ratios of 1:1 (non-deducible) and 4:1 (deducible) were used for two-person mixtures; ratios of 1:1:1 (non-deducible) and 5:1:1 (deducible) were used for three-person samples. Over 700 samples of various combinations of 85 contributors were amplified in duplicate or triplicate and analyzed for the purposes of drop-out and drop-in rate estimation. The 85 contributors represented the diverse population of New York City. For 72% of the samples, the ethnicity of the donor was known, as these donors were laboratory employees. The breakdown was as follows: 20% Asian, 16% black, 54% Caucasian, and 10% Hispanic [24]. The remaining samples were obtained at autopsy and represented a random draw from the population of New York City. According to the 2010 United States census, the population of New York City is 9.8% Asian, 26.6% black, 44.7% Caucasian, and 27% Hispanic [25].

Allele calls for the single source samples and the mixtures were made by the software and artifacts were edited without reference to the known contributors and were then compared to the profiles of the known contributors. Any of the known contributors' alleles that were not present or were present but did not meet the laboratory's limit of detection, 75 RFU, were considered drop-out. Any peaks above 75 RFU that could not be attributed to the known contributors were considered drop-in, whether or not they were in the stutter position of a true allele. Drop-in was modeled this way, as it is never possible to distinguish with certainty whether an extraneous allele in stutter position is actually stutter or not. In addition, for the purposes of the LR, the nature of an extraneous allele is not important.

Drop-out rates were calculated by quantity of template DNA from 25 pg to 500 pg for mixtures and 6.25–500 pg for single source samples. At each locus, the number of opportunities for homozygous and heterozygous drop-out (i.e., the number of homozygous and heterozygous contributors to the samples times the number of amplifications) were counted and used as denominators for the drop-out rates. The number of instances of partial heterozygous drop-out, total heterozygous drop-out, and homozygous drop-out were then tallied and used as the numerators for the drop-out rates. The standard deviation was also calculated and the rates minus one standard deviation were used for the FST program.

Drop-in rates were calculated for each locus and sample type and for 28 cycle and 31 cycle amplification protocols. Drop-in was calculated per locus per replicate. The number of single allele drop-in events and the number of events involving two or more drop-in alleles were counted separately.

2.2.4. Analysis of mock evidence samples

Using OCME's current casework guidelines, mixtures were categorized as deducible or non-deducible. Deducible mixtures can be deconvolved into the full or partial profiles of the major contributors, whereas non-deducible mixtures cannot. The mixture deconvolution was carried out by experienced analysts without any reference to known profiles. The profiles of known contributors were then compared to the mixtures and categorized as one of the following: (a) included/major: the deduced major contributor was consistent with the known contributor and/or all of the alleles of the known contributor were labeled in the mixture; (b) cannot be excluded: most of the known contributor's alleles were labeled in the mixture; (c) no conclusions: several of the known contributor's alleles were not labeled in the mixture but the individual could not be ruled out as a contributor; and (d) excluded: many of the known contributor's alleles were not

observed in the mixture. The purpose of the exercise was to perform a comparison to these types of qualitative assessments and demonstrate the effectiveness of probabilistic models.

FST was used to a generate likelihood ratio for each true contributor to each mock evidence profile, based on the apparent number of contributors and deducibility. For two-person profiles, the LR was formulated as suspect + unknown versus two unknowns. For three-person profiles, LR was formulated as suspect + two unknowns versus three unknowns. FST was also tested with a known profile, such as a victim profile, included in the model. For these scenarios, the LR for two-person profiles was formulated as suspect + victim versus victim + unknown and the LR for three-person profiles was formulated as suspect + victim + unknown versus victim + two unknowns.

Furthermore, each mock evidence sample was tested against the database of 1246 non-contributors to determine the distribution of LR expected when a suspect is actually not a contributor to the mixture obtained from the evidence. Therefore, the program was run an additional 1246 times for each evidence sample, using each individual in the population database in turn as the "suspect".

3. Results

For two-person and three-person deducible and non-deducible mixtures, the fraction of alleles called at each STR locus is shown in Fig. 1A–D. Rates are given for deducible two-person samples (A), non-deducible two-person samples (B), deducible three-person samples (C), and non-deducible three-person samples (D). Within each type of mixture, call rates are separated by quantity of template DNA, from 25 pg to 500 pg. Call rates were calculated as $(2A + B - C - 2D - E)/(2A + B)$, where A is the number of

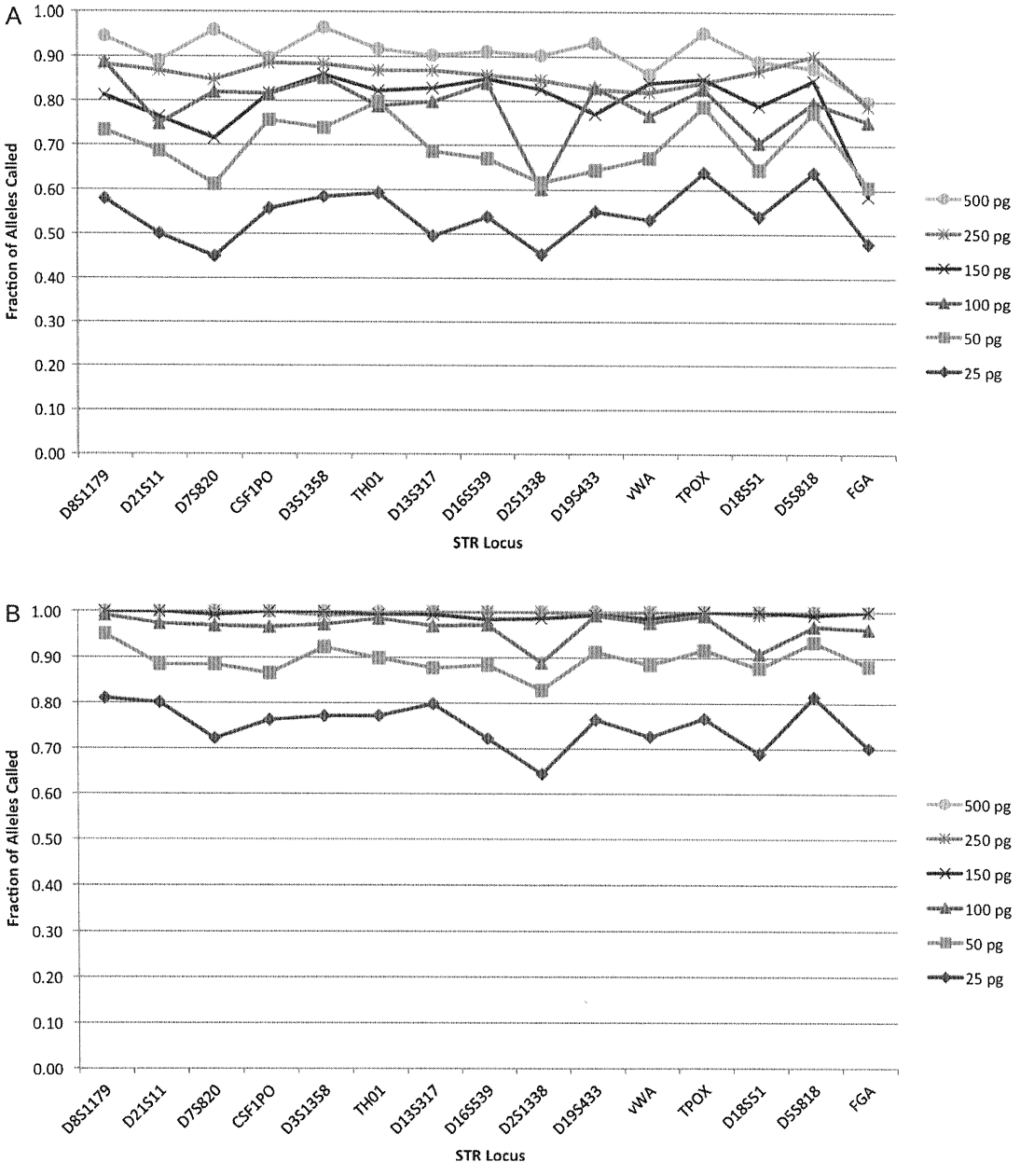


Fig. 1. Fraction of alleles called for minor contributors to two-person deducible mixtures (A), both contributors to two-person non-deducible mixtures (B), minor contributors to three-person deducible mixtures (C), and all contributors to three-person non-deducible mixtures (D).

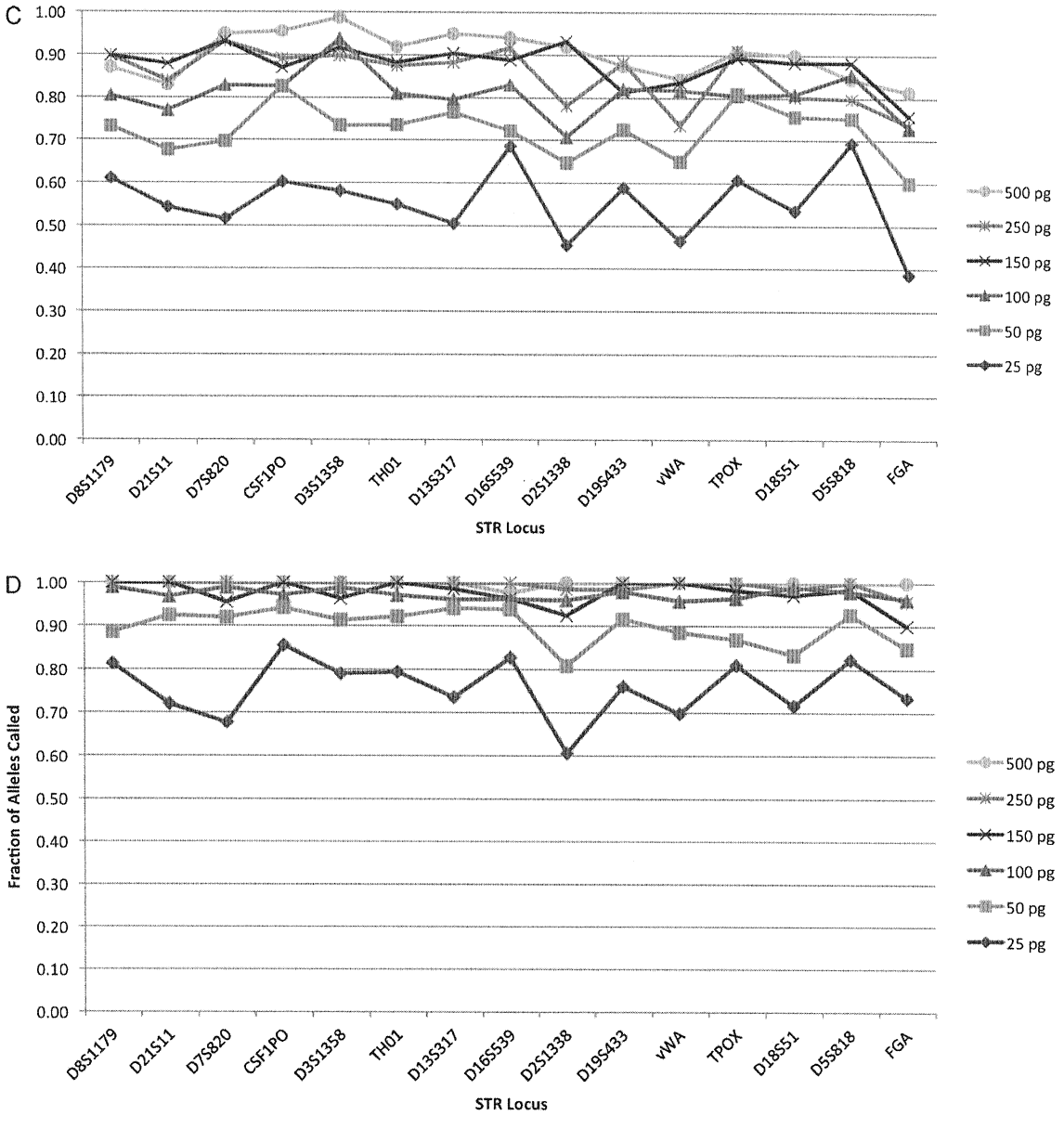


Fig. 1. (*Continued*).

heterozygous loci, B is the number of homozygous loci, C is the number of instances of partial heterozygous drop-out, D is the number of instances of complete heterozygous drop-out, and E is the number of instances of complete homozygous drop-out.

Drop-in rates did not vary significantly by locus, number of contributors, or DNA template quantity. For single source samples, no drop-in alleles were observed with 28 cycle amplification. The average drop-in rate for 31 cycle amplification of single source samples was 0.018 per locus per replicate. For two-person mixtures, the average drop-in rates were 0.009 and 0.041 per locus per replicate for 28 and 31 cycles, respectively. For two-person mixtures, these values were 0.027 and 0.050. Drop-in of two or more alleles at a single locus in a single replicate occurred infrequently, at a rate of approximately 0.005 per locus per replicate. Drop-in rates were not substantially and consistently different for different loci or starting template amounts, but were mainly a function of amplification cycle number. Therefore, only two sets of drop-in rates were incorporated into the FST software and the probability of drop-in of one allele per locus per replicate was set to 0.02 for 28 cycles and 0.035 for 31 cycles amplifications. For both 28 and 31 cycles, 0.005 is used for the probability of drop-in of two or more alleles per locus per replicate. Underestimation of the drop-in rates is likely to be conservative.

The majority of drop-in alleles were found in the minus four stutter position of a true allele. This was true for single source samples and for mixtures, although stutter was more prevalent in mixtures than in single source samples. For single source samples, 60% of drop-in alleles occurred in the minus four position of a true allele. For two-person and three-person mixtures, these values were 77% and 83%. This difference approaches statistical significance ($p = 0.059$). There is a statistically significant difference in minus four drop-in rates for single source samples versus two-person mixtures ($p = 0.009$) and for single source samples versus three-person mixtures ($p = 1.7 \times 10^{-4}$). The fraction of drop-in alleles found in the minus four stutter position was identical (79%) for low template and high template mixtures. Rates were similar across loci.

Fig. 2A and B shows the distribution of DNA template quantities amplified for two- (A) and three- (B) person samples in the validation of FST. The maximum length of the whiskers in the box plots is 1.5 times the interquartile range (the distance between the 25th and 75th percentiles). Observations that fall more than 1.5 times the interquartile range away from the 25th or 75th percentile are treated as outliers and indicated by dots beyond the ends of the whiskers. An effort was made to include samples with a wide range of template DNA amounts, with approximately equal numbers of samples in the low template (less than or equal to 100 pg per triplicate amplification) and high template (greater than or equal to 100 pg per duplicate amplification) ranges.

Fig. 3A–D shows the distribution of log likelihood ratios for true contributors with manual calls of "included/major", "cannot be excluded" (CBE), "no conclusions" (NoConc), and "excluded" for two-person deducible mixtures (A), two-person non-deducible mixtures (B), three-person deducible mixtures (C), and three-person non-deducible mixtures (D). Outliers are defined as in Fig. 2. Although there is overlap among the categories, the trend is for decreasing LRs with increasing drop-out.

Log LRs obtained for comparison of non-contributors to mixtures can be found in Fig. 4. Fig. 4A and B shows the distribution of log likelihood ratios for non-contributors to two-person deducible and non-deducible mixtures (A) and to three-person deducible and non-deducible mixtures (B). The number of comparisons to non-contributors was as follows: $N = 166{,}147$ for deducible two-person mixtures, $N = 176{,}158$ for non-deducible two-person mixtures, $N = 93{,}700$ for deducible three-person mixtures, and $N = 121{,}869$ for non-deducible three-person mixtures. Log LRs less than −44 are plotted at −44.

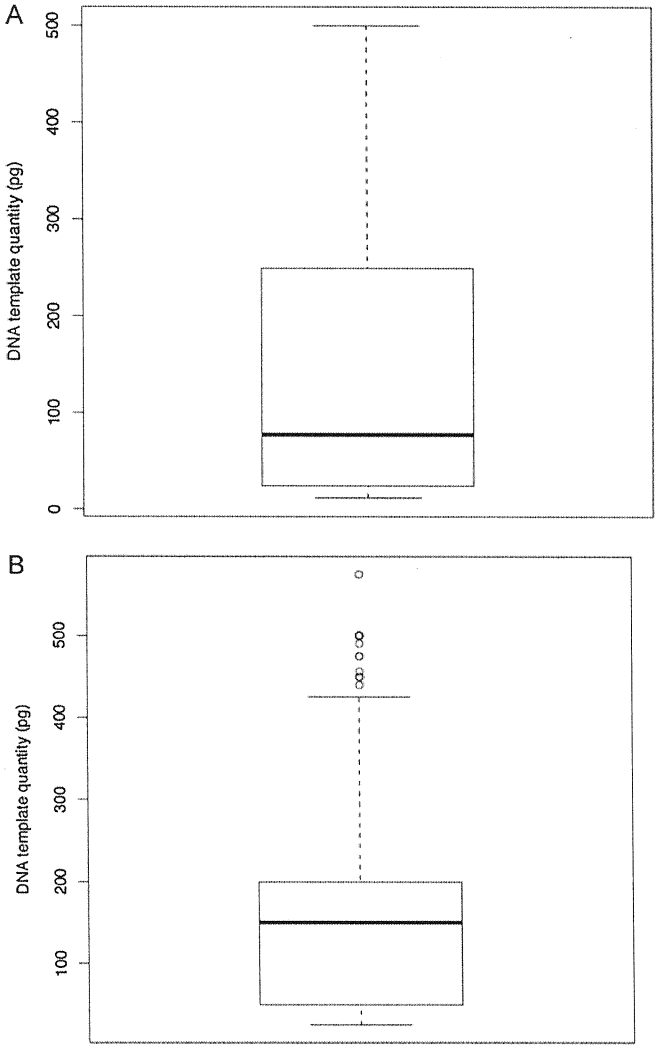


Fig. 2. Distribution of DNA template quantities amplified for two (A) and three (B) person samples. For two-person samples, $N = 272$; for three-person samples, $N = 208$.

Overall, a small number of non-contributors generated a LR greater than 1.0 due to allele sharing, for both the deduced and non-deduced parameters for both two- and three-person mixtures. Although this occurred more frequently with LT-DNA two-person mixtures than with HT-DNA two-person mixtures, there was not a significant difference between these sample types for three-person mixtures. Table 4A and B shows the frequency at which an LR greater than 1, 10, 100, 1000, and 10,000 was obtained for two person (A) and three-person (B) mock evidence samples when non-contributors were used as test samples. Values are shown for deducible and non-deducible mixtures. When non-contributors were tested, no LR's greater than 1000 were observed for two-person mixtures, and no LR's greater than 10,000 were observed for three-person deducible mixtures. Considering all 557,874 comparisons evaluated, the frequency for which a non-contributor generated an LR greater than 100 was 0.0020% and greater than 1000 was 0.00090%.

4. Discussion

4.1. Drop-out and drop-in rates

FST represents one of the first implementations of the likelihood ratio for analysis of forensic DNA mixtures incorporating allelic drop-out and drop-in. Our approach is unique in that drop-out and drop-in rates have been empirically estimated using our laboratory's casework protocols. The drop-out rate estimates employed by FST depend on DNA template quantity, the number of contributors to a sample, their approximate mixture ratio, STR locus, and the genotypes (heterozygote or homozygote) of the hypothesized contributors. Drop-in rates are a function only of the number of amplification cycles used, as drop-in was not substantially different for single source samples versus mixtures or for different STR loci. Our drop-in values (0.005–0.035) are consistent with Balding and Buckleton's [19] recommendation of less than 0.05 to reflect the relative rarity of this phenomenon.

In general, STR loci with longer amplicons demonstrated a trend toward higher drop-out rates than those with shorter amplicons. However, there are a few exceptions. For example, the vWA and TH01 loci have similar amplicon sizes, but the drop-out rates are higher for vWA than for TH01. This may be due to excessive allele sharing at TH01 relative to vWA. For heterozygous loci, a trend for the longer or shorter allele to preferentially drop out was not observed (data not shown). The samples used for the current analysis were pristine buccal swabs. Results may differ if samples displaying more of the complicated characteristics of many evidence samples. Further study of these phenomena is warranted. Consistent with our previously published results [9], drop-out rates increased with decreasing amounts of template DNA.

Although increased allele sharing in mixtures sometimes led to lower drop-out rates for three-person mixtures relative to two-person mixtures, drop-out rates were generally higher with increasing numbers of contributors for a fixed total amount of template DNA. One exception was for common alleles. For example, the 12 allele at the D5S818 locus has frequency 0.25–0.35 in the four populations included in the OCME database. For 25 pg three-person mixtures, the drop-out rate of this allele was 0.069, whereas for 25 pg two-person mixtures it was 0.139. This

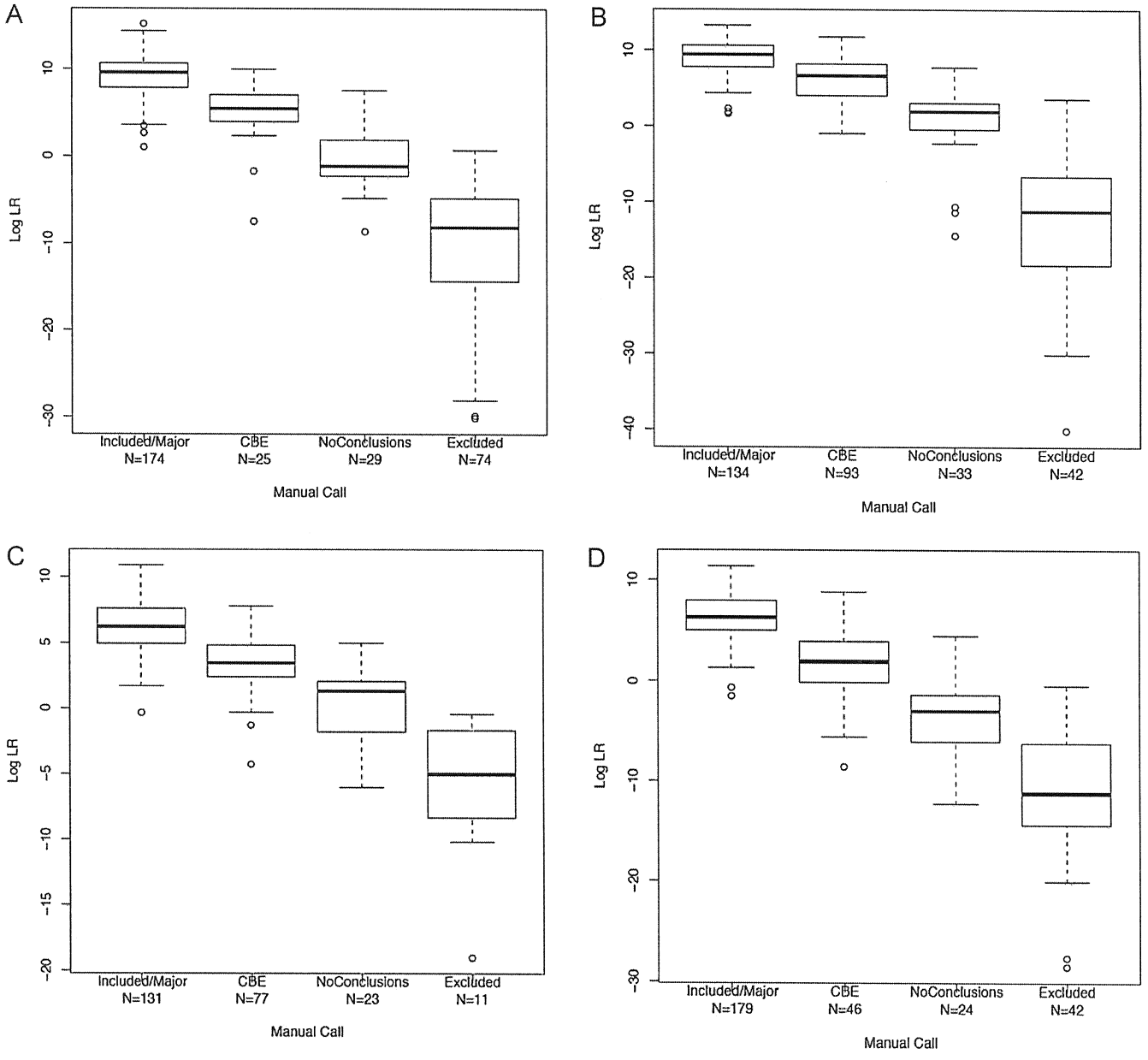


Fig. 3. Distribution of log likelihood ratios for true contributors with manual calls of "included/major", "cannot be excluded" (CBE), "no conclusions" (NoConc), and "excluded" for two-person deducible mixtures (A), two-person non-deducible mixtures (B), three-person deducible mixtures (C), and three-person non-deducible mixtures (D).

difference is likely due to more sharing of the 12 allele with three-person mixtures than with two-person mixtures. For less common alleles at this locus, drop-out rates were higher for three-person than for two-person mixtures.

Estimation of drop-out probabilities is a current topic of discussion in the forensic DNA literature [19,26–30]. Gill et al. [26] introduced a methodology for simulating stochastic variation introduced during each of the steps in STR amplification, including DNA extraction, aliquot sampling, and the PCR process itself. Such a model can be used to describe expected stutter peaks and heterozygote balance and to estimate the probability of drop-out as a function of experimental parameters such as the number of PCR cycles and quantity of template DNA. Gill and Buckleton [27] argue against a peak height threshold below which drop-out is considered, and above which it is not, as this creates a situation in which values just below and just above the threshold are treated very differently. Rather, they advocate for the development of continuous drop-out probability functions. Tvedebrink et al. [28] use such a continuous model with average peak height as a logistic regression predictor of drop-out probability at each locus. Tvedebrink et al. [29] extend the model to allow variable numbers of PCR cycles and to model the STR amplification process in greater detail. The LRMix program [30], on the other hand, does not attempt to model drop-out, but leaves estimation of drop-out rates to the user.

Rather than explicitly modeling the STR amplification process, estimating drop-out probability based on peak heights, and/or allowing users to estimate drop-out rates for a particular sample, we elected to empirically estimate drop-out rates as a function of the total amount of template DNA in a sample, the estimated number of contributors to the sample, their approximate ratio (equal or not equal) and STR locus. In developing and validating our

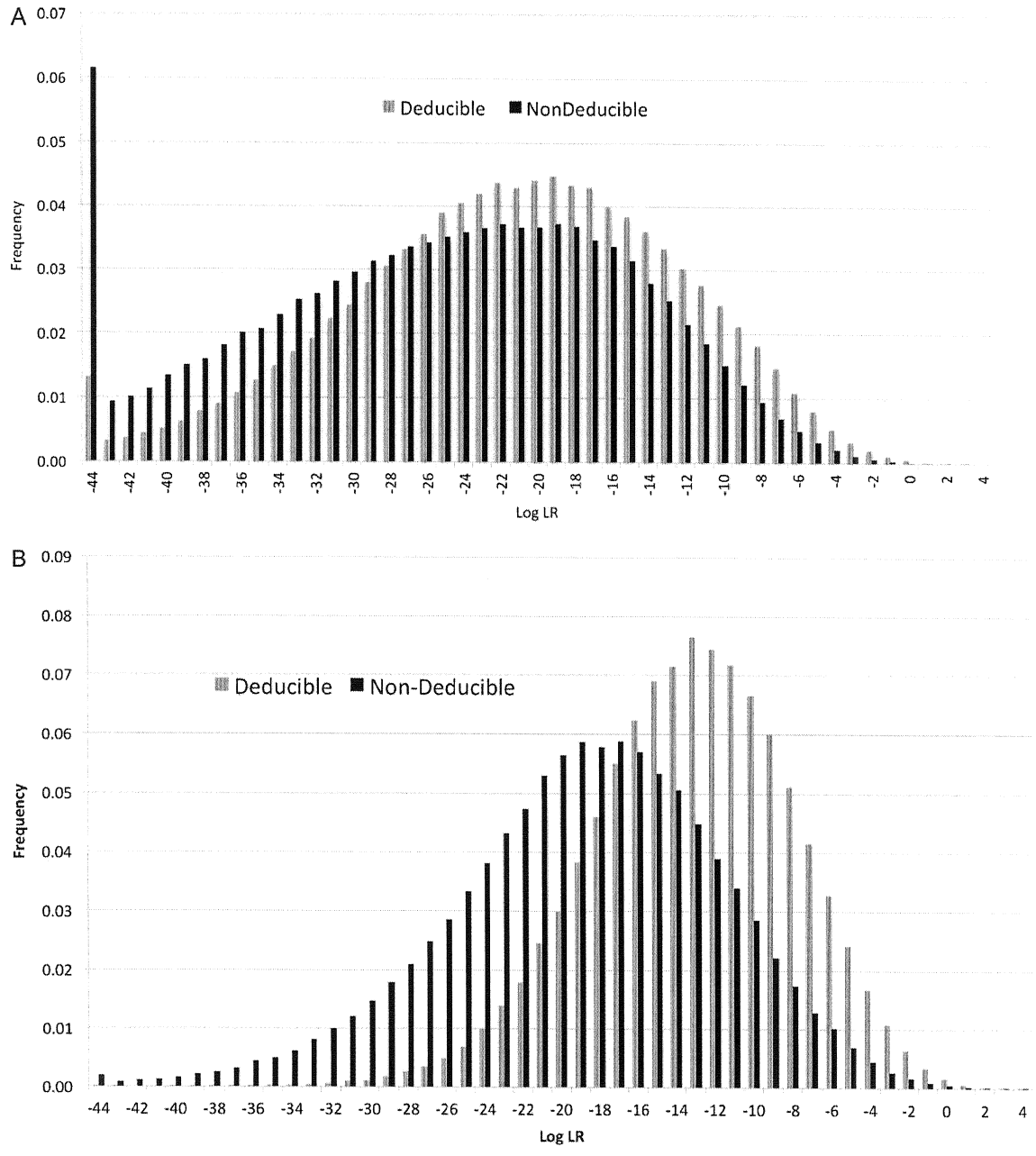


Fig. 4. Distribution of log likelihood ratios for non-contributors to two-person deducible and non-deducible mixtures (A) and to three person deducible and non-deducible mixtures (B). The number of comparisons to non-contributors was as follows: $N = 166{,}147$ for deducible two-person mixtures, $N = 176{,}158$ for non-deducible two-person mixtures, $N = 93{,}700$ for deducible three-person mixtures, and $N = 121{,}869$ for non-deducible three-person mixtures. Log LRs less than −44 are plotted at −44.

high sensitivity testing protocols, we found that peak heights can vary considerably among samples with similar quantities of input DNA and that tall peaks are not necessarily indicative of a low probability of drop-out [9].

OCME's use of a very sensitive and accurate template DNA quantitation system [23] allows reliable quantitation-based estimation of drop-out rates. While there is some variation from instrument to instrument, drop-out rate estimation was performed using data from our laboratory's more sensitive instruments and, thus, the resulting rates may be underestimates for the less sensitive instruments, which is a conservative approach.

Our model for drop-out is continuous, as recommended by Gill and Buckleton [27]. That is, there is no threshold beyond which drop-out is not considered. In order to compute a likelihood ratio for a mixture in which one or more of the hypothesized contributors' alleles are missing, a non-zero probability of drop-out must be used. We use minimum values of 0.05 for drop-out of one allele from heterozygotes and 0.02 for homozygous drop-out and for drop-out of both alleles from heterozygotes. These values were selected after exploration of a range of minimum values. If minimum values are extremely low, a single drop-out event can lead to a very low likelihood ratio, even for true contributors with many repeating alleles. If minimum values are very high, the model is too forgiving of drop-out, leading to low LRs for true contributors without missing alleles and high LRs for non-contributors.

4.2. Validation with mock evidence samples

When FST's performance was tested with mock casework samples that were treated according to OCME casework protocols,

Table 4
Frequency at which a likelihood ratio greater than 1, 10, 100, 1000, and 10,000 was obtained for two-person (A) and three-person (B) mock evidence samples when non-contributors were used as comparison samples. Values are shown separately for deducible and non-deducible mixtures.

LR greater than	Observed frequency (1 in x)	
	Deducible	Non-deducible
(A)		
1	5000	10,000
10	16,000	19,000
100	55,000	176,000
1000	>166,000	>176,000
10,000	>166,000	>176,000
(B)		
1	1200	3100
10	4400	7600
100	13,000	40,000
1000	31,000	60,000
10,000	>93,000	121,000

the LR calculations support the qualitative conclusions and help to refine the broad categories of "included/major", "cannot be excluded", "no conclusions", and "excluded". In fact, FST results were often more conservative than manual conclusions, particularly for samples at the top end of HT- and LT-DNA ranges where the FST drop-out rates are low. The LR values were also affected by the presence of common or rare alleles in the mixture.

To determine the distribution of LRs expected when a non-contributor is used as the test profile, the mock evidence profiles were tested against a database of 1,246 population samples. Each non-contributor was treated in turn as the "suspect" with a model that included no known contributors. This exercise is extremely valuable, in that it demonstrates the range of LRs that could be expected when the wrong person is included as a test profile in the model. That is, it gives an idea of what to expect if, for example, a suspect is actually not associated with a piece of tested evidence.

The highest LRs for non-contributors occurred with three-person mixtures, which often contain most of the common alleles at the STR loci. Thus, it would not be unusual to find alleles matching those of a non-contributor at many or all loci in a three-person sample. The highest LR obtained for any non-contributor (4.59×10^4) occurred with an item that had been handled by three-persons without first cleaning the item. This sample contained 575 pg of template DNA per amplification, so it was a high template sample that was amplified in duplicate. The non-contributor that produced the highest LR carried only common alleles that were shared with one or more of the true contributors to the mixture. None of the non-contributor's alleles were missing from the mixture and only two were non-repeating. Thus, this LR appropriately reflects the chance inclusion of this individual's alleles in the mixture.

4.3. Casework considerations

The DNA in evidence samples sometimes shows signs of degradation, which may include decreasing peak heights for loci with longer versus shorter amplicons (i.e., ski slope effect), fewer labeled alleles at longer loci than shorter loci, and STR results of lower quality than expected for the amount of DNA in the sample if the quantification system employed uses shorter amplicons than some of the loci in the STR kit [31,32]. Because degradation is a common phenomenon in evidence samples, a degradation module was incorporated into FST and tested. Using UV light, known samples were degraded and the drop-out rates and patterns of the samples were observed. It was determined that drop-out rates at some loci increased faster than others when samples were subjected to UV radiation, relative to samples that had not been treated. In the FST degradation module, locus-specific drop-out rates were adjusted according to the estimated degree of degradation of a sample, from moderate to severe. Degree of degradation was estimated based on the ratio of peak heights of the longest to the shortest locus in each Identifiler® color. Ultimately, it was determined that, in general, use of the degradation module as programmed resulted in LRs closer to 1.0 for both true contributors and non-contributors. That is, this approach did not increase the overall separation between true contributors and non-contributors (data not shown). This is an area in which further study is warranted, as improvement in quantification of degradation or identification of moderately to severely degraded samples coupled with changes to the degradation model might improve performance.

FST was approved for use with criminal casework samples in New York State by the New York State Forensic Science Commission in December 2010. It is currently in use for mixtures from which major and/or minor contributors cannot be deconvoluted and for comparisons to the minor contributor of mixtures from which a major, but not a minor, contributor can be deconvoluted. Random match probability is still used for single source samples and for deduced profiles.

As explained above all evidence profiles are finalized without consideration of any potential known contributor sample. Samples must be categorized as single source or mixtures and, for mixtures, the number of contributors must be estimated. Characteristics of a mixture can be used to determine whether to treat the mixture as two, three, or four people [24,33–36]. Yet, due to allele sharing amongst related and unrelated individuals, there will always be a level of uncertainty to this determination [33]. Using only the maximum number of alleles observed at any locus to estimate the number of contributors to a mixture will often lead to an underestimate [33,34]. Nevertheless, using the minimum number of contributors typically results in the lowest possible LR, the LR that most favors the defendant [34]. Tools to better estimate the number of contributors are being explored [24,35,36]. FST is currently online for analysis of two- or three-person mixtures. Validation of four-person models is currently in progress.

For mixtures, the analyst must also determine whether or not the profile(s) of major and/or minor contributor(s) can be deduced. If the test profile in a case matches a deduced profile, the random match probability is computed and FST is not used. If the test profile does not match a deduced profile, FST is used with drop-out rates for deducible mixtures. If individual contributors' profiles cannot be determined, FST is used with drop-out rates for non-deducible mixtures. During the FST LR calculation, drop-out is always considered as a possibility, regardless of whether or not any of the potential contributors' alleles are missing. Thus, the model is not selected based on the absence or presence of a suspect or another person's alleles.

Prior to the application of FST to casework samples appropriate prosecution and defense hypotheses must be selected. Within the context of a case, the analyst must determine whether or not to include known contributors in the formulation of the LR. For intimate samples or certain case scenarios, a victim or other individual may be included as a known contributor to a mixture, which will reduce the number of evidence alleles to be explained by the comparison sample. This will generally lead to larger LR values and must be used carefully. Therefore, for certain scenarios, OCME computes and reports the LR with and without including a known contributor, and reports both results.

OCME analysts receive extensive training on presentation of LR results in court. One must be extremely careful when stating any type of results in court and the LR is no different. Great care is taken to avoid transposing the conditional [37] when presenting results.

That is, analysts are trained to state that the DNA alleles labeled in the evidence are x times more probable if H_p is true than if H_d is true (or vice versa), rather than stating that H_p is x times more probable than H_d. To help jurors, judges and attorneys understand the magnitude of the results, support for H_p over H_d (or H_d over H_p) is described using the scale of limited support to very strong support suggested by Butler [38] and others.

As an additional interpretation measure, it could be considered to test an evidence mixture against the database of random non-contributors. If the database yields any results larger than those obtained using the test sample from the case, the rate at which this occurred can give an estimate of how often a randomly selected individual could be expected to generate an LR that is at least as large as that obtained using the test sample. For example, if three population samples out of the database of 1246 generated an LR larger than the test sample comparison, we could expect to obtain an LR at least this high for 1 in 415 randomly selected individuals. If the database did not yield any results larger than that of the test sample, the distance between the highest LR for a random population sample and the test sample could be presented. Since testing more contributors would yield a better estimate of the population frequency of the LR results, simulated databases for random non-contributors could also be used. Such simulated DNA profiles as well as Tippet plots were explored by Gill et al. [39] to formulate a different method to test robustness of a weight assessment. Although more work and validation is needed prior to implementation, these general approaches could help put the isolated LR result in perspective.

5. Conclusions

There are two main limitations to the current version of the FST application. First, correlation among genotypes of contributors to mixtures is not considered, which means the calculation is based on unrelated individuals. Each unknown person's genotype is treated as independent from the genotypes of all others in the model. Hypotheses that involve unknown individuals who are related cannot be explicitly modeled and for example, FST cannot accommodate a prosecution hypothesis that includes the victim and the suspect along with a defense hypothesis that includes the victim and the suspect's brother. In this type of situation, the analyst explains that we cannot perform that comparison with the current program, but that the question could probably be resolved with a sample from the brother.

The second limitation of the program is that the drop-out and drop-in rates employed by the program are specific to OCME's protocols, kits and equipment. Application to data generated in another laboratory would require assessment and perhaps adjustment of these rates, as would alteration of OCME's current protocols, kits or equipment. While the program is currently not available to other laboratories, sharing options are being explored and methodology to adapt the parameters for other laboratories or changes to protocols within our laboratory is in progress.

In summary, the FST program performed well over a wide range of DNA template amounts and mixture ratios and types. The results for the touched and purposefully degraded samples demonstrate that the program parameters used are suitable and conservative for degraded samples. The LRs generated when comparing a true contributor with a mixture were generally concordant with manual interpretations. The LR values for true contributors correlated well with the qualitative strength of the positive associations and the few outliers can be explained by the presence of frequent alleles and demonstrates the importance of adding a weight assessment. The addition of a large set of comparisons to random non-contributors identified several chance positive associations mostly to very complex mixtures, as can be expected based on allele sharing, but overall showed very good separation between LR values for contributors and non-contributors. This validation shows that the FST program is a useful tool for calculating a statistical weight for evidence and reference sample comparisons particularly for mixtures that cannot be deconvoluted, or have components that cannot be deconvoluted. The FST software provides a quantitative weight to interpretations that would otherwise be qualitative or where no conclusions could previously be drawn.

Conflict of interest statement

None of the authors has a financial or personal relationship with other people or organizations that could inappropriately influence or bias this paper.

References

[1] I. Findlay, A. Taylor, P. Quirke, R. Frazier, A. Urquhart, DNA fingerprinting from single cells, Nature 389 (1997) 555–556.
[2] P. Gill, J. Whitaker, C. Flaxman, N. Brown, J. Buckleton, An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA, Forensic Sci. Int. 112 (2000) 17–40.
[3] J.P. Whitaker, E.A. Cotton, P. Gill, A comparison of the characteristics of profiles produced with the AMPFLSTR SGM Plus multiplex system for both standard and low copy number (LCN) STR DNA analysis, Forensic Sci. Int. 123 (2001) 215–223.
[4] L. Forster, J. Thomson, S. Kutranov, Direct comparison of post-28-cycle PCR purification and modified capillary electrophoresis methods with the 34-cycle low copy number (LCN) method for analysis of trace forensic DNA samples, Forensic Sci. Int. Genet. 2 (2008) 318–328.
[5] G. Ruano, W. Fenton, K.K. Kidd, Biphasic amplification of very dilute DNA samples via 'booster' PCR, Nucleic Acids Res. 17 (1989) 5407.
[6] P. Taberlet, S. Griffin, B. Goossens, S. Questiau, V. Manceau, N. Escaravage, L.P. Waits, J. Bouvet, Reliable genotyping of samples with very low DNA quantities using PCR, Nucleic Acids Res. 24 (1996) 3189–3194.
[7] P.J. Smith, J. Ballantyne, Simplified low-copy-number DNA analysis by post-PCR purification, J. Forensic Sci. 52 (2007) 820–829.
[8] M. Prinz, L. Schiffner, J.A. Sebestyen, E. Bajda, J. Tamariz, R.C. Shaler, H. Baum, T. Caragine, Maximization of STR DNA typing success for touched objects, Int. Cong. Ser. 1288 (2006) 651–653.
[9] T. Caragine, R. Mikulasovich, J. Tamariz, E. Bajda, J. Sebestyen, H. Baum, M. Prinz, Validation of testing and interpretation protocols for low template DNA samples using AmpFlSTR Identifiler, Croat. Med. J. 50 (2009) 250–267.
[10] P. Gill, C.H. Brenner, J.S. Buckleton, A. Carracedo, M. Krawczak, W.R. Mayr, N. Morling, M. Prinz, P.M. Schneider, B.S. Weir, DNA commission of the International Society of Forensic Genetics: recommendations on the interpretation of mixtures, Forensic Sci. Int. 160 (2006) 90–101.
[11] Scientific Working Group on DNA Analysis Methods, SWGDAM Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories. Available from: http://www.fbi.gov/about-us/lab/codis/swgdam-interpretation-guidelines, 2010.
[12] M.W. Perlin, B. Szabady, Linear mixture analysis: a mathematical approach to resolving mixed DNA samples, J. Forensic Sci. 46 (2001) 1372–1378.
[13] M.W. Perlin, J.B. Kadane, R.W. Cotton, Match likelihood ratio for uncertain genotypes, Law Probab. Risk 8 (2009) 289–302.
[14] M.W. Perlin, A. Sinelnikov, An information gap in DNA evidence interpretation, PLoS One 4 (2009) e8327.
[15] M.W. Perlin, M.M. Legler, C.E. Spencer, J.L. Smith, W.P. Allan, J.L. Belrose, B.W. Duceman, Validating TrueAllele(R) DNA mixture interpretation, J. Forensic Sci. 56 (2011) 1430–1447.
[16] J.M. Curran, P. Gill, M.R. Bill, Interpretation of repeat measurement DNA evidence allowing for multiple contributors and population substructure, Forensic Sci. Int. 148 (2005) 47–53.
[17] P. Gill, A. Kirkham, J. Curran, LoComatioN: a software tool for the analysis of low copy number DNA profiles, Forensic Sci. Int. 166 (2007) 128–138.
[18] H. Haned, Forensim: an open-source initiative for the evaluation of statistical methods in forensic genetics, Forensic Sci. Int. Genet. 5 (2011) 265–268.
[19] D.J. Balding, J. Buckleton, Interpreting low template DNA profiles, Forensic Sci. Int. Genet. 4 (2009) 1–10.
[20] National Research Council, The Evaluation of Forensic DNA Evidence, National Academy Press, Washington, DC, 1996.
[21] J.M. Butler, R. Schoske, P.M. Vallone, J.W. Redman, M.C. Kline, Allele frequencies for 15 autosomal STR loci on US Caucasian, African American, and Hispanic populations, J. Forensic Sci. 48 (2003) 1–4.
[22] L.A. Schiffner, E.J. Bajda, M. Prinz, J. Sebestyen, R. Shaler, T.A. Caragine, Optimization of a simple, automatable extraction method to recover sufficient DNA from low copy number DNA samples for generation of short tandem repeat profiles, Croat. Med. J. 46 (2005) 578–586.
[23] J.A. Nicklas, E. Buel, Development of an Alu-based, real-time PCR method for quantitation of human DNA in forensic samples, J. Forensic Sci. 48 (2003) 1–9.

[24] J. Perez, A.A. Mitchell, N. Ducasse, J. Tamariz, T. Caragine, Estimating the number of contributors to two-, three-, and four-person mixtures containing DNA in high template and low template amounts, Croat. Med. J. 52 (2011) 314–326.
[25] U.S. Census Bureau, U.S. Census Bureau Delivers New York's 2010 Census Population Totals, Including First Look at Race and Hispanic Origin Data for 2010. Available from: http://2010.census.gov/news/releases/operations/cb11-cn122.html.
[26] P. Gill, J. Curran, K. Elliot, A graphical simulation model of the entire DNA process associated with the analysis of short tandem repeat loci, Nucleic Acids Res. 33 (2005) 632–643.
[27] P. Gill, J. Buckleton, A universal strategy to interpret DNA profiles that does not require a definition of low-copy-number, Forensic Sci. Int. Genet. 4 (2010) 221–227.
[28] T. Tvedebrink, P.S. Eriksen, H.S. Mogensen, N. Morling, Estimating the probability of allelic drop-out of STR alleles in forensic genetics, Forensic Sci. Int. Genet. 3 (2009) 222–226.
[29] T. Tvedebrink, P.S. Eriksen, M. Asplund, H.S. Mogensen, N. Morling, Allelic drop-out probabilities estimated by logistic regression-Further considerations and practical implementation, Forensic Sci. Int. Genet. 6 (2012) 263–267.
[30] H. Haned, T. Egeland, D. Pontier, L. Pene, P. Gill, Estimating drop-out probabilities in forensic DNA samples: a simulation approach to evaluate different models, Forensic Sci. Int. Genet. 5 (2011) 525–531.
[31] M. Graw, H.J. Weisser, S. Lutz, DNA typing of human remains found in damp environments, Forensic Sci. Int. 113 (2000) 91–95.
[32] P.M. Schneider, K. Bender, W.R. Mayr, W. Parson, B. Hoste, R. Decorte, J. Cordonnier, D. Vanek, N. Morling, M. Karjalainen, C. Marie-Paule Carlotti, M. Sabatier, C. Hohoff, H. Schmitter, W. Pflug, R. Wenzel, D. Patzelt, R. Lessig, P. Dobrowolski, G. O'Donnell, L. Garafano, M. Dobosz, P. De Knijff, B. Mevag, R. Pawlowski, L. Gusmao, M. Conceicao Vide, A. Alonso Alonso, O. Garcia Fernandez, P. Sanz Nicolas, A. Kihlgreen, W. Bar, V. Meier, A. Teyssier, R. Coquoz, C. Brandt, U. Germann, P. Gill, J. Hallett, M. Greenhalgh, STR analysis of artificially degraded DNA-results of a collaborative European exercise, Forensic Sci. Int. 139 (2004) 123–134.
[33] J.S. Buckleton, J.M. Curran, P. Gill, Towards understanding the effect of uncertainty in the number of contributors to DNA stains, Forensic Sci. Int. Genet. 1 (2007) 20–28.
[34] H. Haned, L. Pene, J.R. Lobry, A.B. Dufour, D. Pontier, Estimating the number of contributors to forensic DNA mixtures: does maximum likelihood perform better than maximum allele count? J. Forensic Sci. 56 (2011) 23–28.
[35] H. Haned, L. Pène, F. Sauvage, D. Pontier, The predictive value of the maximum likelihood estimator of the number of contributors to a DNA mixture, Forensic Sci. Int. Genet. 5 (2011) 281–284.
[36] 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.
[37] I.W. Evett, B.S. Weir, Interpreting DNA Evidence. Statistical Genetics for Forensic Scientists, Sinauer Associates, Sunderland, MA, 1998.
[38] J.M. Butler, Forensic DNA Typing, Second Edition: Biology, Technology, and Genetics of STR Markers, Elsevier Academic Press, Burlington, MA, 2005.
[39] P. Gill, J. Curran, C. Neumann, A. Kirkham, T. Clayton, J. Whitaker, J. Lambert, Interpretation of complex DNA profiles using empirical models and a method to measure their robustness, Forensic Sci. Int. Genet. 2 (2008) 91–103.

Austin Journal of Forensic Science and Criminology

Open Access

Austin Publishing Group

Review Article

Touch DNA in a Complicated Alleged Child Abuse Case

Coyle HM*
Department of Forensic Science, University of New Haven, USA

*Corresponding author: Miller Coyle H, Department of Forensic Science, Henry C. Lee College of Criminal Justice & Forensic Sciences, University of New Haven, 300 Boston Post Road, West Haven, CT 06516, USA

Received: June 10, 2015; **Accepted:** October 20, 2015; **Published:** October 22, 2015

Abstract

Touch DNA can be of use in establishing what may have occurred through reconstruction of events based on biological evidence transfer. However, interpretation of results and patterns must be approached with some caution as in the alleged child abuse case detailed here. This case was brought forward as a touch DNA and body fluid case where the male in question was a father reported to have forced a young child to perform oral sex on him. Her pajamas were collected and evaluated for presence of body fluids and associated DNA. The sleeves of the pajamas tested positive for amylase, a potential indicator of saliva and DNA of both victim and father combined. Initially, it was thought that this would be clear evidence to bring forth to trial; however, as the remaining stains were tested on the pajamas, reconstruction of events became substantially altered. Up to six family member DNA profiles were recovered off the child's garments and also a semen stain from a half-brother.

Keywords: DNA; Familial DNA; Touch DNA; Sexual assault; Body fluids; Pattern interpretation

Abbreviations

DNA: Deoxyribonucleic Acid; FST: Forensic Statistical Tool; LR: Likelihood Ratio

Introduction

The enzyme alpha amylase is a non-specific indicator of saliva and was used to screen sixteen stains identified on a young child's pajama shirt. Of the sixteen stains, six were positive for amylase, two stains were inconclusive, and eight stains were negative. Subsequent DNA testing of the stains revealed various DNA mixtures with at least six family member detected by standard forensic DNA methods. On the pajama pants, fifteen stains were tested for the presence of amylase and seven were positive; one stain was also positive for semen. Alpha amylase is an enzyme produced by salivary glands and using forensic Phadebas tests, false positives have been detected from urine, sweat and fecal matter [1].

Since the amylase diffusion test indicates but is not a conclusive identification for saliva, some interpretation of the DNA profiles associated with the stains was important for forming conclusions about the case. Vaginal secretions and bacteria, both commonly found on worn clothing, also will yield a positive result for amylase [1]. When DNA is recovered from a stained area, it may be from the same source as the body fluid and deposited at the same time as the fluid. Alternatively, deposit may occur as an independent event through touching whereby shed epithelial cells are being placed in the same area either before or after the saliva and appear by DNA test methods as an inadvertent mixture.

Given the accusation of forced oral sex with the juvenile victim, the case went forward for prosecution with the focus being on the two amylase positive stains on the wrist area of the right and left pajama sleeves (Table 1, shirt stains 1 and 2) that contained the DNA mixture of father and victim (shirt stain 1) and victim only (shirt stain 2). However, when the case was evaluated holistically and with a broad overview of all the stains combined with the DNA results, (sixteen on the shirt, and fifteen on the pants) and a surprise semen contributor on the pajama pants(who was also the witness against the father), the interpretation of the DNA case became significantly more complex.

The six family members included: (a) the victim, (b) an uncle, (c) an aunt, (d) the biological mother, (e) the half-brother and (f) the biological father. The majority of the DNA mixtures consisted of two individuals but due to large percentages of shared DNA, the results were often not sufficiently distinctive enough to do more than include a relative as a potential contributor in the non-deductible mixture (Table 1). Although probabilistic genotyping software is one method for teasing apart DNA mixture interpretation, when high percentages of shared genetic information exist as with this case, the software often cannot conclusively establish with DNA profiles which relative is the definite contributor [2]. DNA mixtures and stain interpretation for reconstruction of events remains a challenge in forensic science in these circumstances. If one relative possesses a rare allele, then that individual may be distinguishable from the others, however, there is some associated error rate when genetic relatives are involved [3-5]. The Forensic Statistical Tool (FST) software recommends not using their mathematical algorithms for calculation of Likelihood Ratios (LR) when genetic relatives are involved as with shared genetics the error rates would be substantially greater due to coincidental matching; this is evident by accidental matches to non-contributors in validation studies [6]. True Allele software has one published case where genetic relatives could be distinguished and the difference is likely due to use of no analytical threshold and the presence of a detectable allele difference between the related individuals [7]. In this case scenario, you can see without the father's DNA information, it would be impossible to distinguish maternal parentage between the mother and aunt for the victim which does lend some concern for immigration casework since the shared alleles are 50% identical by descent and either individual could have contributed the necessary allele per locus. If these two individuals' DNA were present in a DNA mixture, it would be difficult to discern individuality.

Citation: Coyle HM. Touch DNA in a Complicated Alleged Child Abuse Case. Austin J Forensic Sci Criminol. 2015; 2(5): 1042.

Table 1: Allele Values for Touch DNA samples.

Locus	Victim	Uncle	Aunt	Mother	Father	Half-Brother	Shirt Stain 1	Shirt Stain 2	Pant Semen Stain 15
D3S1358	16	16	16	16	16	15,16	16	16	15,16
D16S539	11,13	12,13	11,12	12,13	11,12	11,13	11,12,13	11,13	11,12,13
Sex	X,X	X,Y	X,X	X,X	X,X	X,Y	X,Y	X,X	X,Y
TH01	6,9	9.3	6,9	6,9.3	7,9	6,8	6,7,9	6,9	6,8,9,9.3
TPOX	9,11	11	8,9	8,11	9,10	8	9,10,11	9,11	8,9,11
CSF1PO	9,12	12	9,10	10,12	12,Z*	12	9,12	9,12	9,12
D7S820	8,9	8,11	8,12	8,11	12,Z*	8,10	8,9,12	8,9	8,9,10,11
VWA	14,17	17	14,18	14,17	17,18	11,17	14,17,18	14,17	11,14,17
FGA	22,24	24	22,24	24	22	21,24	22,24	22,24	21,22,24
D8S1179	12,15	10,12	12,15	12	14,15	12,14	12,14,15	12,15	12,14,15
D21S11	28,29	28,30	29,30	28,30	29	28	28,29	28,29	28,29
D18S51	17,18	17,18	14,17	14,18	17,20	12,18	17,18,20	17,18	12,17,18,23
D5S818	10	12,13	10,13	10,13	10,13	10,12	10,13	10	10,12,13
D13S317	11,12	11,13	9,11	11,13	9,12	11,13	9,11,12	11,12	9,11,12,13
D2S1338	17,20	20,21	17,20	20,23	17,20	20,23	17,20	17,20	17,20,21,23
D19S433	11,16.2	13,16.2	11,16.2	13,16.2	11,14	13,15.2	1,14,16.2	11,16.2	11,13,15.2,16.2

Z*, refers to a missing allele or allele drop-out during processing of the DNA sample

The outcome for this case was an acquittal based on the inability to firmly associate the amylase enzyme activity from the shirt stains to the small child wiping her mouth on the sleeve after eating or to bacteria or to forced oral sex. Other contributing factors to the decision included the highly confounding touch DNA results from thirty-one stains indicating that multiple family members, including the father, had detectable DNA on her pajamas which could have been deposited under any number of circumstances. The most compelling biological evidence was a semen stain from the young half-brother on the hip of the victim's pajama bottoms which strongly suggested a different scenario than the oral sex allegation. This was initially a challenging case to resolve due to a family custody battle with additional immigration issues but the DNA and body fluid analysis was helpful in refuting the allegations of oral child sex abuse brought against the father given the semen evidence.

Acknowledgment

Thank you to attorney Alfred Dorfman, 26 Court Street, Brooklyn, NY 11242 for sending this case for DNA review.

References

1. Casey D, Price J. The sensitivity and specificity of the RSID-saliva kit for the detection of human salivary amylase in the Forensic Science Laboratory, Dublin, Ireland. Forensic Science International. 2010; 194: 67-71.
2. Tvedebrink T, Bright JA, Buckleton JS, Curran JM, Morling N. The effect of wild card designations and rare alleles in forensic DNA database searches. Forensic Science International Genetics. 2015; 16: 98-104.
3. Perlin MW, Dormer K, Hornyak J, Schiermeier-Wood L, Greenspoon S. TrueAllele casework on Virginia DNA mixture evidence: computer and manual interpretation in 72 reported criminal cases. PLoS ONE. 2014; 9: 1-16.
4. Mitchell AA, Tamariz J, O'Connell K, Ducasse N, Budimlija Z, Prinz M, et al. Validation of a DNA mixture statistics tool incorporating allelic drop-out and drop-in. Forensic Science International Genetics. 2012; 6: 749-761.
5. Miller Coyle H. Sources of computational error in probabilistic genotyping software used for DNA mixture interpretation. International Research Journal of Computer Science. 2015; 5: 12-16.
6. Mitchell AA, Tamariz J, O'Connell K, Ducasse N, Prinz M, Caragine T. Likelihood ratio statistics for DNA mixtures allowing for drop-out and drop-in. Forensic Science International Genetics Supplement Series. 2011; 3: e240-e241.
7. Cybergenetics Newsletter. DNA intelligence and forensic failure: what you don't know can kill you. 2011.

Citation: Coyle HM. Touch DNA in a Complicated Alleged Child Abuse Case. Austin J Forensic Sci Criminol. 2015; 2(5): 1042.

Development of a Simulation Model to Assess the Impact of Contamination in Casework Using STRs

ARTICLE *in* JOURNAL OF FORENSIC SCIENCES · JUNE 2004

Impact Factor: 1.16 · DOI: 10.1520/JFS2003366 · Source: PubMed

2 AUTHORS:

Peter Gill
Norwegian Institute of Public Health
206 PUBLICATIONS **10,823** CITATIONS
SEE PROFILE

Amanda J Kirkham
University of Birmingham
14 PUBLICATIONS **217** CITATIONS
SEE PROFILE

J Forensic Sci, May 2004, Vol. 49, No. 3
Paper ID JFS2003366
Available online at: www.astm.org

Peter Gill,[1] *Ph.D. and Amanda Kirkham,*[1] *B.Sc.*

Development of a Simulation Model to Assess the Impact of Contamination in Casework Using STRs

ABSTRACT: Because contamination is usually tube-specific, negative controls cannot give assurance that an associated batch of extracted casework material is contaminant-free. However, it is possible to use them to predict the level of overall (undetected) contamination that is processed by an operational DNA unit. A MATLAB®-based program was used to combine results of negative controls with actual casework DNA profiles to assess the probability that laboratory contaminants will give rise to reportable profiles (along with their likelihood ratios). Using data from an operational DNA unit as an example, it was demonstrated that the risk is inextricably linked to guidelines used to interpret DNA profiles. We have demonstrated how computer-based models can predict the levels of contamination expected in the process and, in addition, how the process can be made more robust by changing reporting guidelines. There is a need to compare DNA profiles against staff and plasticware elimination databases in order to determine sources of contamination. The likeliest outcome of a contamination event is false exclusion.

KEYWORDS: forensic sciences, contamination, short tandem repeats, negative control, national DNA database, simulation

In this paper we use a computer simulation model to estimate potential levels of contamination that may be encountered in routine casework; we also demonstrate how the effect of contamination can be mitigated by reference to the simulation model.

First of all, it is important to make the distinction clear between the meaning of the terms "contamination" and "adventitious transfer" of DNA. Gill et al. (1) originally outlined a model where the definition of contamination was specifically restricted to describe the introduction of extraneous DNA into the process directly as a result of the intervention of a police investigator or scientist. The alternative kind of contamination, namely adventitious transfer, refers to the transfer of DNA from sources that may be unconnected to the case and the investigator, for example by secondary transfer from the perpetrator (2). The crucial point is that adventitious transfer occurs only before the crime scene is established, whereas contamination can occur only afterwards.

At the biochemical level, contamination may manifest itself in two nonexclusive ways. Either as allele "drop-in"—described by Gill et al. (3) where spurious alleles from multiple independent sources may find their way into a polymerase chain reaction (PCR)—or as a partial or complete DNA profile from a single (rather than from multiple) source. There are certain other generalizations to be made. First, no process is entirely free of contamination. Secondly, contaminants tend to be quantitatively at low level. This means that casework profiles that are also low level, partial, and consequently difficult to interpret, tend to be affected more than full profiles that are associated with substantial amounts of DNA. Thirdly, contamination can be characterized by reference to negative controls.

There are various sources of contamination that are known to occur. Each can be monitored by reference to staff elimination databases and negative controls in order to estimate or capture the effect:

1. Transfer from scientists and police investigators, e.g., skin flakes or saliva spray—prevented by good laboratory practice (gowns and face masks) and monitored by reference to staff databases that consist of relevant personnel.
2. Reagent contamination. Plasticware may be contaminated during the manufacturing process—sterile does not equate to DNA-free. Schmidt et al. (4) showed that reaction tubes used to analyze mtDNA were often contaminated. We have subsequently demonstrated several casework-related STR contamination events were derived from staff of a reaction tube manufacturer (after voluntary screening was carried out). Monitoring of consumables prior to introduction into casework (on a batch basis) is important to establish their quality and to ensure that manufacturers are not inadvertently contaminating their products. Ideally, a comparison of casework samples and negative controls against staff elimination databases from reagent manufacturers is needed.
3. The contaminant may arise from another sample that has been processed concurrently and cross contamination has resulted (e.g., lane-to-lane leakage on a flatbed gel or between adjacent wells of microtitre plates on automated systems). This can be monitored by software that compares samples within a batch for potential duplicate profiles—interpretation of mixtures may be necessary.

Multiplexes such as Applied Biosystems AmpF*l*STR SGM plus (5) are very sensitive using manufacturer recommended methods, down to 250 pg, or less than 100 pg if increased PCR cycles (3) are utilized—this means that just several cells may be required to give a signal. Low level contamination tends to be sporadic (i.e., tube specific). The logical consequence of this is that a negative control does not give confidence that the associated batch of extractions is contaminant free. The only exception to this is when the contaminant is observed both in the negative control and in an associated extraction tube(s), in which case the contamination is considered gross, i.e., multiple events that are possibly reagent based.

[1] Forensic Science Service, Trident Court, Birmingham Business Park, 2960 Solihull Parkway, Solihull, B37 7YN, UK.

Received 1 Nov. 2003; and in revised form 19 Jan. 2004; accepted 19 Jan. 2004; published 7 April 2004.

The purpose of this paper is to illustrate how negative controls can be used to predict the performance of a DNA unit in terms of an estimate of the expected number of false positive results along with their associated likelihood ratios. In addition we show how the use of reporting guidelines affects the results. Because contamination tends to be low level, this means that the lower the reporting threshold, the greater the chance that a contaminant will be reported.

For example, by analyzing negative control data, and casework data from a Forensic Science Service (FSS) operational laboratory, we demonstrated that under a reporting guideline utilizing a 50 rfu peak height lower threshold there was a chance of approximately 1 in 1000 that a sample would be a false positive with a likelihood ratio greater than 10^7. However, simply by raising the reporting lower threshold from 50 rfu to 80 rfu was sufficient to negate the effect.

This does not mean that low-level profiles cannot or should not be reported, rather the relevance of low-level DNA evidence is less certain. In particular, interpretation at the source level (i.e., association of the DNA profile with a particular body fluid) is uncertain. The most probable outcome of a contamination event is false exclusion. A framework to report low level DNA samples under hierarchy of propositions principles using Bayesian networks is given by Evett et al. (6)

Method

Negative Controls

To estimate the level of false positive reporting, a suite of MATLAB® (The Mathworks Inc., Natick, MA) programs (SIMLAB) was written to simulate the casework environment. The specific purpose was to predict the overall level of false positive results that may occur as a result of sporadic and undetected laboratory contamination. The term sporadic and undetected contamination specifically relates to:

1. Extraneous DNA in plasticware, solutions or other reagents that have been introduced during the manufacturing process (i.e., external to the laboratory).
2. Contamination from personnel who are not part of a staff elimination database and therefore undetected.

A contaminant may be detected if:

3. The contaminant has arisen from an operator who is on a staff elimination database—either the forensic lab or plasticware manufacturer.
4. The contaminant has come from another sample that has been processed concurrently and cross contamination has resulted (e.g., lane-to-lane leakage on flatbed gel or a microtitre plate used in automated systems).

In this paper only inexplicable contamination events (in 1,2 above) are simulated; i.e., profiles explained by staff contamination and cross contamination are not included in the analysis.

To determine overall levels of contamination and to estimate the associated effects, we examined negative extraction controls. These controls are a microcosm of casework samples that are assumed devoid of DNA. Nevertheless, reportable DNA profiles are sometimes observed. Because negative controls are processed in the same way as casework samples, they can be used to estimate the level of contamination in casework samples over the same period of time.

Most, if not all, contamination events seen in negative controls are sporadic single-tube events—i.e., the contaminant is specific to one tube only. Generally, this means that the contaminant has no relevance to the associated batch of extracted samples.

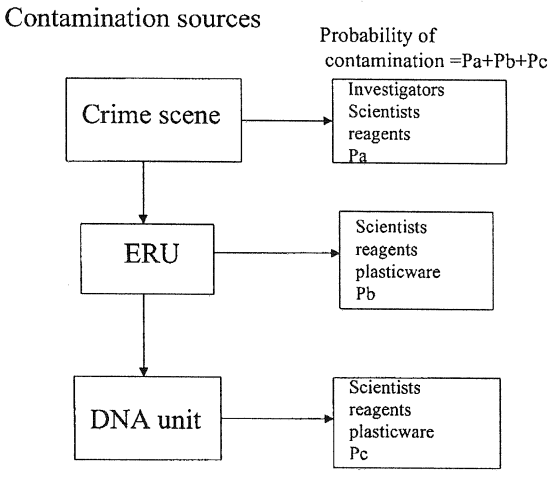


FIG. 1—*Flow diagram to outline potential sources of contamination.*

This does not mean to say that negative controls are useless, far from it; *however, we have to interpret them in a way that relates to the entire DNA process, not just in relation to a specific batch of samples with which the negative controls have been co-processed.* It may be believed that quality is assured simply by eliminating a batch of samples where the negative control gives a result—but this ultimately leads to a false sense of security. To date, the effect of sporadic contamination on the DNA process has not been considered and is therefore poorly understood.

How Contamination Arises and How the Process Can be Monitored to Produce Predictive Models of the Consequences

To be completely effective, any model must take account of the entire process. In the Forensic Science Service (FSS) the case is submitted to an evidence recovery unit (ERU) where it is evaluated and stains are then submitted to the DNA unit for processing. We can break down the origin of contaminants into three discrete consecutive categories (Fig. 1):

1. At the crime scene—contamination by investigating officers or reagents used to collect evidence (P_a).
2. Similarly, when the case is transferred to the ERU, where evidence is evaluated, contamination may result from scientists, reagents, or plasticware (P_b).
3. In the DNA unit the same as category 2 applies (P_c).

Each category comprises several subcategories; for example, P_b could comprise separate probabilities for swabs, plastic tubes, scientists, other reagents, etc. Over time, contaminants will pass from the crime scene to the ERU and finally to the DNA unit for analysis; at each stage there is additional opportunity for contamination to occur so that the process is additive.

To summarize, the chance of contamination (C) is subdivided into several non-exclusive categories such that $C = P_a + P_b + P_c$ (Fig. 1). If a control is to reflect the entire process it purports to control, ideally it should be prepared at the crime scene in order to capture P_a. Consideration also needs to be made about the kind of negative control that is employed.

For example, if a moistened swab is used to collect evidence, then the ideal negative control would be an additional blank swab also

moistened with water and concurrently prepared. Similarly, within the ERU, negative controls to measure P_b could also be prepared and cascaded to the DNA unit. If the only part of the process that produces negative controls is the DNA unit itself, then only P_c is estimated and this is therefore an underestimate of the total risk of contamination.

Method to Estimate P_c

The method illustrates how to estimate P_c measured from negative controls that are prepared and generated by the DNA analysis unit itself. In principle, P_a and P_b can also be estimated provided the appropriate negative controls are collected and processed. The standard AmpF*l*STR SGM plus® (Applied Biosystems) system was used as described by Cotton et al. (5) using the standard 28 PCR cycles.

Collection of Data

Two subsets of data were collected from a DNA unit:

1. All 295 negative controls over the period of time for which an assessment was made (a period of five months—10 May 02 to 14 Dec. 02), including samples for which no signal was obtained.
2. A random collection of 50 casework DNA profiles over the same period of time, including samples that failed to give any signal.

Details of SIMLAB Computer Program

Each profile is represented in a single row of a spreadsheet format where each locus is defined by three parameters—allele designation, peak height, and peak area. Only true allelic peaks that are within their expected size ranges are used in the analysis.

Profiles from negative controls are similarly prepared.

All blank negatives and casework samples are included in the simulation regardless of whether they generate profiles. Thus in a sample of 295 negative controls 26 showed evidence of contamination. Conversely, in the sample of casework profiles, five failed to show a result.

Pairwise comparison was used to combine casework with contaminants observed in negative controls. For 50 casework samples and 295 negative controls there are 50 × 295 simulated casework/contaminant profiles. There are four possible outcomes: casework sample only, contaminant only, casework/contaminant mixture, or a blank result with no profile apparent.

The simulated profiles are analyzed further. For each sample, the mixture proportion (M_x) is approximately estimated by summing peak areas of contaminants and casework samples and calculating:

$$M_x = \text{peak heights negative controls/peak heights casework samples}$$

(Note: this could be improved by calculating M_x per locus and treating each locus separately.)

Once M_x has been calculated, then reporting guidelines are applied for each locus in turn as follows:

(*a*) Is the allele peak height greater than the threshold reporting level (e.g., 50 rfu)? Below this level the allele is deemed inconclusive and not reportable (falls into the low copy number (LCN) category not considered further by the program).

(*b*) Is the locus a mixture?

Further reporting guidelines are applied as follows: If the locus is a mixture, then is the major component a contaminant and is it reportable? To fulfill this requirement the peak height must be greater than the threshold reporting level and must also be distinct from the minor component of the mixture. To measure whether the contaminant profile is distinct from the casework sample a rule is applied to the effect that M_x must be >0.5; i.e., on average, contaminant peaks must be 50% higher than casework peak heights in order to be accepted. Peaks that fail the rule are deemed inconclusive.

(*c*) The surviving peaks are deemed reportable and then the major profile is converted into a likelihood ratio. Note that an important feature of the guideline is the ability to run "what-if" scenarios simply by changing the reporting parameters to any desired level. Changing reporting parameters alters the number of reportable peaks and the outputs are given in terms of reportable profiles that can be converted into likelihood ratios (or match probabilities).

Results and Discussion

Pre-assessment of the Data

(*a*) *Casework:* Out of the 50 samples analyzed, five failed to give a result; i.e., the probability of a sample failing to give any profile (P_F) = 5/50 = 0.1.

(*b*) *Negatives:* Out of 295 samples analyzed, a total of 26 samples gave a signal. This means that the probability of negative control giving a profile (P_N) of one or more alleles is 26/295 = 0.088.

A contaminant is only detected or known to have occurred if it is found in a negative control tube purported to be free of DNA. The difficulty is that it is not possible to assess directly whether a casework sample is affected by sporadic contamination, as there is no supporting information. However, even though we cannot know which particular casework tube is contaminated, unless it matches a profile on a staff elimination database, we can assess the probability (P_c) of any given tube being affected because negative samples are simply a subset of casework samples. It is the same as the probability that a negative control is contaminated (where the contamination event may be one or more alleles)

$$P_N = P_c = 0.088$$

If a casework sample is contaminated then this will result in one of two different outcomes:

(*a*) If the casework sample is devoid of DNA, then only the contaminant will be visible and the profile appears unmixed. The chance of this occurrence (P_S) is the probability of contamination multiplied by the probability of a casework sample failing to give a profile:

$$P_S = P_N \times P_F$$

Specifically, in the DNA unit assessed $P_S = 0.009$ (or approximately 0.9% of samples will be contaminated and profile does not appear admixed).

(*b*) In addition, P_N–P_S, i.e., 0.088 − 0.0088 = 0.079 (or 7.9%) of casework samples will contain sporadic contamination in admixture with a casework profile.

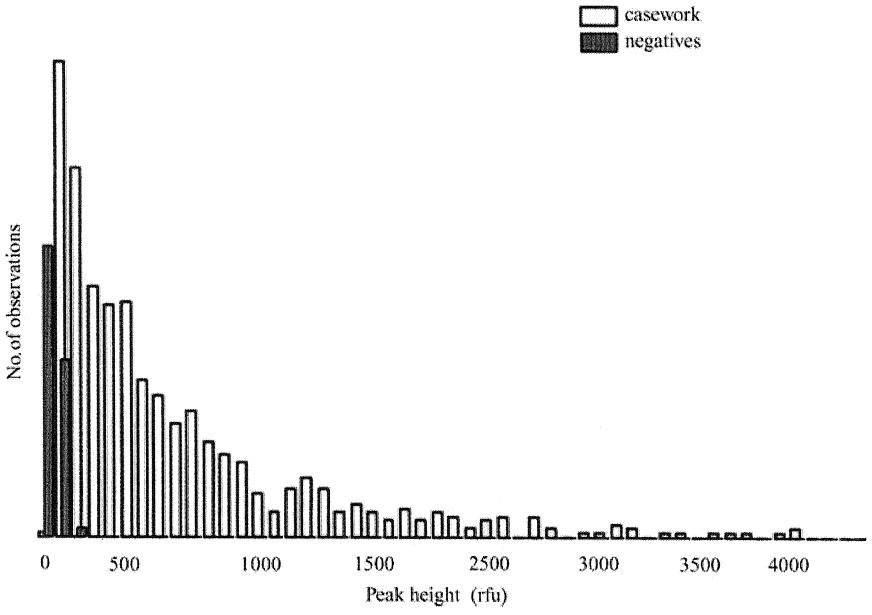


FIG. 2—*Comparison of peak heights >0 across all loci, plotted for negative controls (a population of 26 samples) and for casework samples (a population of 50 samples).*

Negative Controls Define When the Low Copy Number Framework Should Be Used for Reporting Purposes

The low copy number (LCN) probabilistic framework of reporting was originally developed by Gill et al. (3,7) in order to enable reporting of sub c. 100 pg quantities of DNA . The method was specifically designed to take account of sporadic contamination ("drop-in") and stochastic effects that led to allelic "dropout"; probabilities of both were factored in into likelihood ratio calculations. Duplicate analysis and derivation of a consensus profile are used. Furthermore, we recognized that the relevance of the evidence (origin of the sample tested and circumstances of deposition) was less certain compared with substantial profiles from large quantities of DNA derived from an obvious body fluid stain. Consequently, statements are specially tailored to reflect this uncertainty. In the United Kingdom, LCN is routinely reported to the courts. Although it is widely perceived that LCN refers to the use of DNA profiling under conditions that increase the sensitivity of detection by increasing the number of PCR cycles or increasing the injection time with CE instrumentation, in fact this is not the case. LCN has nothing to do with the ability to detect minute quantities of DNA; rather the decision to use the low copy number analysis and reporting framework is dominated by "drop-in" or contamination and this also occurs with conventional DNA profiling (note all data analyzed in this paper were derived by conventional methods). This is why it is useful to collate information from negative controls, recognizing that this is a direct mimic of the casework position. For example, if peaks c. 80 rfu are observed in 1 in 1000 negatives, then the same frequency of contamination is expected in casework samples. The 80 rfu observation in this example defines the "drop-in" threshold below which any alleles would be reported using the LCN framework, to take account of the fact that some may be present due to contamination. This example is somewhat oversimplified but serves to make the point. Whereas sporadic contamination will affect samples on a regular basis, it does not give an indication of the actual impact on casework reporting. To do this assessment properly, it is necessary to evaluate negative control and casework data in much greater detail with special emphasis on their relative peak areas or heights. Peak heights of negative control and casework data were combined and plotted across loci (Fig. 2). Note that the current data set used in the current analysis was rather limited. More data are preferable because a more comprehensive analysis could ensue; e.g., separate characterization of individual loci would be advantageous. The purpose of this paper is to provide a demonstration, in principle, of the proposed methodology, rather than a definitive analysis.

The majority (58%) of contaminant peak heights were <50 rfu. However, 42% were greater and consequently overlapped with case stain peak heights up to c. 150 rfu. We can summarize that in casework c. 17% of alleles were >50 <150 and c. 70% of data were <250 peak height (Fig. 2).

Can a Mixture of a Contaminant with an Evidential Sample Result in Mistyping?

The question arises, as a consequent of the foregoing, whether a mixture of a contaminant with a case sample can give a misleading result.

A MATLAB® program was used to rank the sum of peak heights of strongest weakest negative controls (Table 1) and weakest–strongest casework samples (Table 2), respectively. The worst scenario occurs when a strong contaminant combines with a weak or absent casework sample.

Characterization of Mixtures

Mixtures were simulated by using pairwise comparisons of casework versus negative control data (including all examples where profiles were absent). This means that from 50 casework samples and 295 negative controls, pairwise comparisons generated $50 \times 295 = 14750$ mixtures.

The simulated results comprised:

- Unmixed case samples only (82%)
- Unmixed contamination (0.9%)
- A mixture of case sample and contaminant (7.9%)
- No DNA profile detected (9.1%)

TABLE 1—*Case samples ranked in order of increasing summed peak height with numbers of alleles scored above a given peak height.*

No. allele>0 rfu	No. alleles>25 rfu	No. alleles>50 rfu	No. alleles>100 rfu	No. alleles>150 rfu	Sample No.
0	0	0	0	0	5
0	0	0	0	0	10
0	0	0	0	0	38
0	0	0	0	0	44
0	0	0	0	0	49
1	1	1	0	0	23
2	2	2	0	0	30
4	4	4	3	1	29
5	5	5	3	3	47
7	7	7	3	3	32
13	13	13	7	2	31
13	13	13	8	4	14
16	16	15	6	3	9
13	13	13	10	5	42
13	13	13	10	5	50
14	14	14	9	7	15
10	10	10	9	9	39
17	17	17	15	14	4
19	19	19	16	14	8
19	19	19	15	12	35
22	22	22	21	19	12
19	19	19	19	19	22
12	12	12	9	8	17
14	14	14	14	14	34
16	16	16	16	16	13
21	21	21	21	21	45
18	18	18	18	18	41
18	18	18	16	16	43
22	22	22	20	19	7
15	15	15	15	15	26
20	20	20	20	20	2
20	20	20	20	20	40
21	21	21	21	21	46
19	19	19	19	19	27
20	20	20	20	20	16
20	20	20	20	20	11
18	18	18	18	18	24
20	20	20	20	20	37
20	20	20	20	19	18
22	22	22	22	22	33
22	22	22	22	22	48
20	20	20	20	20	3
19	19	19	19	19	25
21	21	21	21	21	6
18	18	18	18	18	19
19	19	19	19	19	20
20	20	20	20	20	28
21	21	21	21	21	1
19	19	19	19	19	21
21	21	21	21	21	36

The mixture proportion (M_x) was calculated as M_x = sum peak heights contaminant/sum peak heights casework samples (8)—the distribution of M_x is given in Table 3. Most mixtures gave $M_x < 1$, which means that the casework sample was usually the major component. In approximately 1 in 500 cases the major component was the laboratory contaminant; in the most extreme example $M_x = 25$.

Likelihood Ratios of Reportable Contaminant Profiles

Finally, we ask the question, What does this mean in practice in terms of casework reporting and in terms of the national DNA database (NDNADB)? Cases are reported in terms of likelihood ratios. We calculated the likelihood ratio of DNA profiles that originated from contamination. Following laboratory guidelines, alleles were not incorporated into calculations unless above the LCN threshold (currently 50 rfu). Only unmixed contaminants were collated in Table 4; inclusion of major contaminant profiles in mixture data made little difference because of their relative rarity. From Table 4 and Fig. 3, the chance of a laboratory contaminant resulting in a reportable profile LR$>10^7$ was approximately 1 in 1000. We then assessed the effect of changing the current 50 rfu reporting guideline. We demonstrated that the risk reduced if the guideline was increased—for example, if the LCN guideline was rfu = 80, then the maximum LR observed was 10^2.

Data analysis showed that 82% samples were case samples only, and 0.9% of samples will give an unmixed profile from a laboratory contaminant. The overwhelming majority of mixtures will be $M_x < 1$, which means that the casework sample was the major component in most cases. In approximately 1 in 500 cases the major component was the laboratory contaminant. Consequently, this means that by far the greatest problem is when the laboratory

TABLE 2—*Negative controls ranked in descending order of intensity, taken from a population of 295 negative controls—only 26 controls that gave a signal are listed (i.e., 275 controls were blank).*

Sum Peak Height	No. alleles > 0	>25	>50	>100	>150	>200	>250	Sample No.
1217	16	16	16	5	3	0	0	26
1109	17	17	9	4	2	2	1	15
695	11	10	6	1	1	1	1	3
481	11	11	3	0	0	0	0	16
413	8	6	2	2	1	0	0	1
334	10	9	0	0	0	0	0	8
290	7	6	2	0	0	0	0	14
242	7	5	1	0	0	0	0	4
234	5	4	3	0	0	0	0	2
226	4	4	2	0	0	0	0	9
140	3	3	1	0	0	0	0	20
104	4	2	0	0	0	0	0	17
67	2	2	0	0	0	0	0	6
64	1	1	1	0	0	0	0	18
60	1	1	1	0	0	0	0	5
51	2	1	0	0	0	0	0	24
50	1	1	0	0	0	0	0	7
50	2	1	0	0	0	0	0	10
50	1	1	0	0	0	0	0	19
46	2	1	0	0	0	0	0	13
36	1	1	0	0	0	0	0	12
36	1	1	0	0	0	0	0	25
33	1	1	0	0	0	0	0	21
33	1	1	0	0	0	0	0	23
31	1	1	0	0	0	0	0	11
16	1	0	0	0	0	0	0	22

TABLE 3—*Analysis of observations relative to M_x = sum peak heights negatives/sum peak heights casework samples. When in admixture, laboratory contaminants give profiles that are greater in size than casework samples when $M_x > 1$ (marked in bold type).*

Mixture (M_x)	No. Observations	Probability
Case sample only	12105	0.8207
Negative sample only	130	0.0088
No sample	1345	0.0912
<=0.1 >0	973	0.0660
<=0.2 >0.1	73	0.0049
<=0.3 >0.2	30	0.0020
<=0.4 >0.3	19	0.0013
<=0.5 >0.4	9	0.0006
<=0.6 >0.5	9	0.0006
<=0.7 >0.6	9	0.0006
<=0.8 >0.7	4	0.0003
<=0.9 >0.8	6	0.0004
<=1 >0.9	5	0.0003
<=2 >1	**11**	**0.0007**
<=10 >2	**19**	**0.0013**
<=25 >10	**3**	**0.0002**
Total	14750	

TABLE 4—*Probability estimates for achieving a given likelihood ratio where a laboratory contaminant is responsible for the major (unmixed) profile.*

Log 10 LR	Guideline (rfu)					
	rfu = 50	rfu = 60	rfu = 70	rfu = 80	rfu = 90	rfu = 100
1	0.00746	0.00502	0.00319	0.00251	0.00339	0.00339
2	0.00217	0.00095	0.00014	0.00088	0	0
3	0.00027	0.00041	0.00210	0	0	0
4	0.00095	0.00183	0	0	0	0
5	0.00014	0	0	0	0	0
6	0.00007	0	0	0	0	0
7	0.00020	0	0	0	0	0
8	0.00088	0	0	0	0	0
9	0	0	0	0	0	0

contaminant appears as a nonmixed sample. Simply reducing the failure rate will reduce the risks of reporting a contaminant.

Thompson et al. (9) suggest that any level of contamination reduces likelihood ratios. Whereas we would not disagree with the principle, each case needs separate consideration. In particular, we have shown that there is no one simple error rate that can be universally applied regardless of the circumstances—it is dependent upon several factors, in particular the quantity of DNA analyzed and the associated reporting lower threshold limit. In addition, replicate analyses of different stains or areas of the same stain will also significantly reduce the impact of potential error. Good quality management system (QMS) feedback results in continually updated procedures and in turn reduces potential error rates. The key to understanding errors, their impact, and enabling continuous improvement of existing analytical processes is dependent upon instigation of monitoring systems such as those outlined here. Once the data are collected they can be used to inform Bayesian networks (6) that incorporate the contamination rates into probabilistic formulae.

The primary risk of contamination is wrongful exclusion, particularly if the contaminant masks the perpetrator's profile. However, it is important to consider that most contamination incidents will result in partial DNA profiles for which random searches of DNA database will produce adventitious matches with low match probabilities. (Given that there are c. 2 m samples on the UK national DNA database a sample with match probability c. 10^{-6} would often result in one or more adventitious matches.) In itself this should not be problematic provided the non-DNA evidence is always carefully considered within the context of an "intelligence database." The primary purpose is to supply a list of potential suspects for further investigation. This follows principles

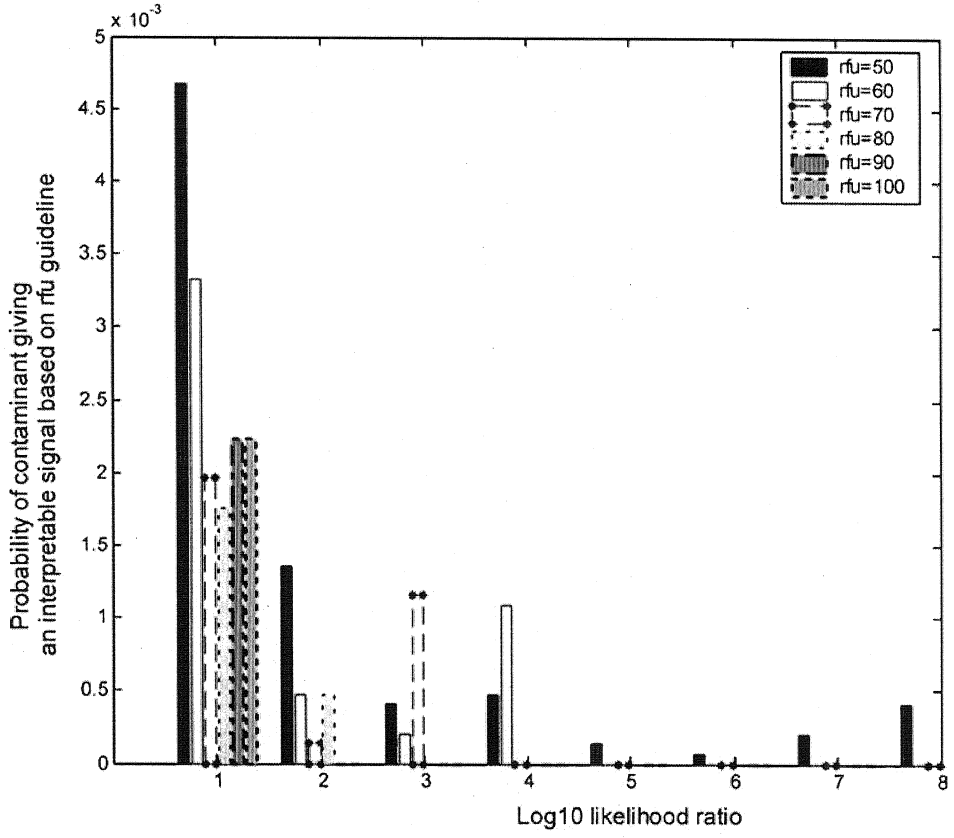


FIG. 3—*Histogram showing probability of a contaminant giving a reportable result (measured as log_{10} LR) relative to the reporting guideline (currently 50 rfu).*

of "hierarchy of propositions" developed by Cooke et al. (10,11) and Evett et al. (12) which has led to a much deeper understanding of the interpretation process. The potential of a result arising from a contamination incident is considered on a case-by-case basis. The approach of combining different kinds of evidence and incorporating contamination as part of the framework can be formalized by utilizing a probabilistic Bayesian network approach (6), and we consider this to be the best way forward.

Finally, it should be recognized that laboratory contamination is impossible to avoid completely but its extent is generally unknown unless proactively assessed—the probability of contamination must always be greater than zero. The effect can be mitigated by comparing all casework profiles against staff databases (including manufacturers of plasticware) but this can capture only a proportion of the events. It is possible to assess the prevalence of contamination and its effect using expert systems that analyze negative controls.

References

1. Gill P. Role of short tandem repeat DNA in forensic casework in the UK—past, present and future perspectives. Biotechniques 2002;32:366–85. [PubMed]
2. Lowe A, Murray C, Whitaker J, Tully G, Gill P. The propensity of individuals to deposit DNA and secondary transfer of low level DNA from individuals to inert surfaces. Forensic Sci Int 2000;129:25–34.
3. Gill P, Whitaker J, Flaxman C, Brown N, Buckleton J. An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA. Forensic Sci Int 2000;112:17–40. [PubMed]
4. Schmidt T, Hummel S, Herrmann B. Evidence of contamination in PCR laboratory disposables. Naturwissenschaften 1995;82:423–31. [PubMed]
5. Cotton EA, Allsop JL, Guest JL, Frazier RRE, Koumi P, Callow IP, et al. Validation of the Ampf/STR SGM plus system for use in forensic casework. Forensic Sci Int 2000;112:151–61. [PubMed]
6. Evett IW, Gill PD, Jackson G, Whitaker J, Champod C. Interpreting small quantities of DNA: the hierarchy of propositions and the use of Bayesian networks. J Forensic Sci 2002;47:520–30. [PubMed]
7. Gill P. Application of low copy number DNA profiling. Croatian Med J 2001;42:229–32.
8. Gill P, Sparkes R, Pinchin R, Clayton T, Whitaker J, Buckleton J. Interpreting simple STR mixtures using allele peak areas. Forensic Sci Int 1998;91:41–53. [PubMed]
9. Thompson WC, Taroni F, Aitken CG. How the probability of a false positive affects the value of DNA evidence. J Forensic Sci 2003;48:47–54. [PubMed]
10. Cooke R, Evett IW, Jackson G, Jones PJ, Lambert JA. A hierarchy of propositions: Deciding which level to address in casework. Sci Just 1998; 38:231–9.
11. Cooke R, Evett IW, Jackson G, Jones PJ, Lambert JA. Case preassessment and review in a two-way transfer case. Sci Just 1999;39:103–11.
12. Evett IW, Jackson G, Lambert JA. More on the hierarchy of propositions: Exploring the distinction between explanations and propositions. Sci Just 2000;40:3–10.

Additional information and reprint requests:
Peter Gill, Ph.D.
Forensic Science Service
Trident Court
Birmingham Business Park
2960 Solihull Parkway
Solihull, B37 7YN
United Kingdom