RESULTS FROM THE 1999 NIST MIXED-STAIN STUDY #2: DNA QUANTITATION, DIFFERENTIAL EXTRACTION, AND IDENTIFICATION OF THE UNKNOWN CONTRIBUTORS

Margaret C. Kline, Janette W. Redman, David L. Duewer, and Dennis J. Reeder

Chemical Science and Technology Laboratory
National Institute of Standards and Technology
Gaithersburg, MD 20899

Abstract:

The National Institute of Standards and Technology (NIST) has provided a number of educational interlaboratory exercises to the forensic community. The recent NIST Mixed Stain Study #2 consisted of three separate sets of materials. Set 1) three stains on washed cotton sheeting, representing a sexual assault by an unknown assailant; a blood stain from the known female "victim", a neat semen stain from a known suspect, and a mixed blood/semen stain of the above knowns plus an additional unknown semen specimen. Set 2) three stains on washed cotton sheeting, representing a sexual assault; a bloodstain from the known female "victim", a neat semen stain from a known suspect, and a mixed blood/semen stain where the semen was different from the male reference. Set 3) five 20 µL DNA TE buffer extracts in screw-capped vials with DNA concentrations stated to range from ≈0.2 to ≈20 ng/µL. Two of these Set 3 samples were, in fact, replicates.

Introduction:

Forty-five local, state, federal, and commercial forensic laboratories participated in this study (of fifty-two laboratories receiving samples) from January to May, 1999. Table 1 profiles the participating laboratories. Participants were requested to: 1) specify all possible types for all donors represented in each sample of Sets 1 and 2 for all loci routinely assayed, 2) provide CODIS profile(s) to search for the suspect(s) in the mixed stain samples, 3) estimate the amount of recoverable DNA per sample (ng/stain), and 4) estimate the concentration (ng/µL) of DNA in all Set 3 samples.

As anticipated, no participant mis-typed any single-donor sample. However, many participants did not attempt to fully type the "unknown male" in the mixed stain of Set 1 – and several participants who did attempt to type this "unknown male" made one or more incorrect assignments.

There were large variations in the quantities of DNA recovered and completeness of the differential extractions in the Set 1 and 2 stains. The median reported DNA concentrations for the TE buffer extracts of Set 3 agree well with the known gravimetric values but the range in reported values was disturbingly large. Many participants did not clearly identify the replicate samples.

Table 1
Participant Profile

Participants	N	%
State/Provincial	26	58
City/County	10	22
Private	5	11
Federal	4	9
Total	45	100

Multiplex(es)	N	%
AmpfℓSTR Profiler™ Plus &COfiler™	23	52
Promega PowerPlex™	11	25
AmpfℓSTR Profiler™ Plus	6	14
AmpfℓSTR Blue™ & Green™ I	2	5
AmpfℓSTR Profiler™ Plus & CTT	1	2
(AmpliType® DQA1, PM; AmpliFLP™ D1S80)	1	2
Total	44	100

Instrument	N	%
ABI 310	21	47
ABI 377	11	24
Hitachi FMBIO®	11	24
ABI 373	1	2
MD FluorImager	1	2
Total	45	100

Quantitation Detection – Method	N	%
ECL – QuantiBlot®	20	43
TMB – QuantiBlot®	13	28
ECL – ACES™	8	17
ECL – slot blot	2	4
F – yield gel	2	4
F – microtiter	1	2
Radio – slot blot	1	2
Total	47	100

*Certain commercial equipment, instruments, or materials are identified in this report to specify adequately the results. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

Samples Distributed in NIST Mixed Stain Study #2 (MSS#2)

Set 1 – Stains on Cotton Cloth
Sample F Reference Female: blood (donor A)
Sample G Reference Male: semen (donor 1087) and blue dextran dye
Sample H Blood (donor A) and semen (donors 1087 and 1039)

Set 2 – Stains on Cotton Cloth
Sample J Reference Female: blood (donor B)
Sample K Reference Male: semen (donor 1131) and blue dextran dye

Sample L Blood (donor B) and semen (donor 1140)

Set 3 – Buffered DNA Solutions in Sealed Vials
Sample M 2.5 ng/μL
Sample N 1.0 ng/μL
Sample O 5.0 ng/μL
Sample P 1.0 ng/μL
Sample Q 0.5 ng/μL

Preparation of Stains
All work was carried out in a laminar flow hood. Seventy-five samples of each stain were prepared on 4 cm × 4 cm squares of white cotton cloth that had been bleached, twice washed, and UV sterilized. After preparation, stains were dried for 2 hours at ambient temperature, placed in brown paper coin envelopes, and stored at –20 °C in barrier bags sealed under vacuum. Table 2 presents quantitative details.

Sample F: Twenty μL aliquots of fresh blood from female donor A. The blood was stirred continuously with a stir-bar to maintain a uniform cell suspension.

Sample G: Twenty μL aliquots of a solution prepared by mixing 200 μL of semen from donor 1087 with 1.3 mL of Phosphate Buffer Saline (PBS) and blue dextran dye (to make the stain visible). This mixture was stirred continuously with a stir-bar to maintain a uniform cell suspension.

Sample H: Twenty-two μL aliquots of a solution prepared by mixing 1.5 mL of fresh female blood from donor A with 270 μL of semen from donor 1087 and 225 μL of semen from donor 1039. This mixture was stirred continuously with a stir-bar to maintain a uniform cell suspension.

Sample J was prepared as Sample F using blood from female donor B.

Sample K was prepared as Sample G using semen from donor 1131.

Sample L was prepared as Sample H using blood from female donor B blood and 142 μL of semen from donor 1140.

Table 2
Quantitative Preparation of Stains

Stain	Component	Aliquot ?L	White cells or s M / mL	Volume mL	# genome[a] M / mL	DNA per stain[b] ng
F	A, blood	20	6.6	1.5	6.6	910
G	1087, semen	20	128	0.2	8.5	1180
	PBS/dye			1.3		
			Total	1.5		
H	A, blood	22	6.6	1.5	5.0	750
	1087, semen		128	0.27	8.7	1310
	1039, semen		145	0.225	8.2	1240
			Total	2.0	21.8	3300
J	B, blood	20	12.2	1.5	12.2	1680
K	1131, semen	20	82.5	0.2	5.5	760
	PBS/dye			1.3		

			Total	1.5		
L	B, blood	22	12.2	1.5	11.1	1690
	1140, semen		80	0.142	3.5	520
			Total	1.64	14.6	2210

[a] One sperm contains one-half of a diploid genome

[b] $$\frac{\text{ng DNA}}{\text{human diploid genome}} = \left(6.8\times10^{9}\,\frac{\text{basepair}}{\text{human diploid genome}}\right)\left(1.013\times10^{-12}\,\frac{\text{ng}}{\text{basepair}}\right)$$

See: http://www.cbs.dtu.dk/nikob/science/nuc_table.html

Preparation of Buffered DNA Solutions

The transmittance and absorbance scales of a UV/vis spectrophotometer were confirmed using NIST SRM® 2031a, Metal-on-Fused-Silica Filters for Spectrophotometry, and the wavelength scale was verified using NIST SRM® 2034, Holmium Oxide Solution Wavelength Standard from 240 nm to 650 nm.

Two DNAs were mixed together in TE buffer. The total DNA concentration and purity of the mixture were verified by spectrophotometry and yield gels. Aliquots of this DNA mixture were diluted with TE buffer to produce the desired concentrations. The replicate Samples N and P were produced as a single lot that was distributed into randomly labeled (N or P) tubes.

Typing Results

Single-donor samples: As anticipated, typing of the single-donor samples was a non-issue for the participating laboratories. Tables 3a and 3b present the types for all donors for the 13 core CODIS STR loci and amelogenin. However, there are procedural problems that need to be addressed by some of the participating laboratories.

- **Transcription errors** in the allele assignments for one or more samples were made by four participants. All of these apparent mis-typings were traced to transcription and/or table labeling errors made while documenting results.
- **Stutter** was handled in different ways by different laboratories. Some laboratories "called" the stutter peaks and noted that they were "probably stutter."
- **Three-banded patterns**: Sample J has a {14, 15, 18} three-banded pattern at locus vWA, with allele 15 of lower intensity than alleles 14 and 18. All laboratories were warned of the presence of a third allele in this sample; some laboratories reported the allele 15 as a "weak allele", some did not report it consistently. Figure 1 shows vWA electrophergrams for 1) a participant who reported the 15 allele in the female fraction of sample L but not in the reference Sample J and 2) a participant who explicitly "struckout" the 15 allele in sample J.

  In the ~ 2000 samples we have analyzed in the past two years, we have encountered only one three-banded pattern of equal intensities and at least five that have one band less intense than the other two.

Multiple-Donor Samples:

- **Peak heights** in the mixtures were not completely reliable indicators of allele association. The "dominant donor" alleles at one locus were not necessarily dominant at all loci (Figure 2) nor were the same patterns always observed in different multiplexes in the same laboratory (Figure 3).
- **Preferential amplification**: One participant amplified only three of the four male donor D21S11 alleles in the male fraction of Sample H. Figure 4 compares electropherograms provided by this laboratory and a reference laboratory at amelogenin, D8S1179, D21S11, and D18S51. The peak height patterns are similar for both laboratories except at locus D21S11. At D21S11, the reference electropherogram has four alleles {27, 28, 29, 32.2} with similar peak heights of 700 RFU while the other shows three alleles {27, 28, 32.2} with different peak heights from 200 to 400 RFU.

We have seen entire locus drop out as the amount of DNA amplified was increased in one single donor sample at the D8S1179 locus.

? **Allele Assignments** Thirty-one of the 44 participants attempted to assign types to the "unknown male" in Samples H and/or L. All of these participants included the correct assignment for "unknown male" of sample L; most participants provided a unique assignment. Fewer participants provided a unique assignment for Sample H. Four participants mis-specified the type at one locus (vWA-1, D5S818-2, D8S1179-1). All of these mis-assignments can be attributed to multiple shared alleles.

Table 3a
Types of Set 1 Sample Components

Locus	F #A	G #1087	H? #A	H? #1087, #1039	Unknown #1039
AMEL	X,X	X,Y	X,X	X,X,Y,Y	X,Y
CSF1PO	12,14	10,13	12,14	10,12,12,13	12,12
D3S1358	14,17	16,18	14,17	16,16,18,19	16,19
D5S818	11,11	11,11	11,11	11,11,11,12	11,12
D7S820	8,10	8,11	8,10	8,10,11,12	10,12
D8S1179	11,14	13,13	11,14	11,13,13,14	11,14
D13S317	10,12	8,10	10,12	8,10,11,11	11,11
D16S539	12,12	12,13	12,12	11,11,12,13	11,11
D18S51	15,22	15,15	15,22	13,15,15,18	13,18
D21S11	29,30	28,32.2	29,30	27,28,29,32.2	27,29
FGA	20,22	21,25	20,22	21,21,23,25	21,23
TH01	6,9.3	8,9.3	6,9.3	6,6,8,9.3	6,6
TPOX	8,11	8,11	8,11	8,8,11,12	8,12
vWA	16,17	16,18	16,17	16,17,18,18	17,18

Table 3b
Types of Set 2 Sample Donors

Locus	J #B	K #1131	L? #B	L? #1140	Unknown #1140
AMEL	X,X	X,Y	X,X	X,Y	X,Y
CSF1PO	13,13	10,11	13,13	10,12	10,12
D3S1358	17,18	16,17	17,18	16,16	16,16
D5S818	11,12	10,11	11,12	11,13	11,13
D7S820	8,9	9,12	8,9	10,11	10,11
D8S1179	12,13	14,14	12,13	13,15	13,15
D13S317	11,11	8,12	11,11	12,12	12,12
D16S539	11,13	12,13	11,13	11,12	11,12
D18S51	12,16	12,19	12,16	16,16	16,16
D21S11	28,32.2	29,29	28,32.2	30,30	30,30
FGA	23,23	23,24	23,23	21,25	21,25
THO1	7,9.3	9.3,9.3	7,9.3	6,6	6,6
TPOX	8,9	8,9	8,9	8,8	8,8
vWA	14,15,18	17,18	14,15,18	16,18	16,18

Quantification Results

Set 3, Individual Samples: Table 4 summarizes the reported DNA concentrations of the five Set 3 samples. Agreement between the nominal (what we believe went into the tubes) and the median of the measured values (what you think came out of the tubes) is very good for all but the lowest concentration. The discrepancy at the lowest level may represent DNA binding to the sample tube.

Table 4
DNA Concentration

Sample	Nominal ng/?L	N	Reported ng/?L Median	Min	Max
O	5.00	44	5.65	0.312	27.50
M	2.50	44	3.00	0.200	10.50
N	1.00	44	1.00	0.063	9.50
P	1.00	44	0.95	0.015	5.00
Q	0.50	43	0.25	0.020	2.90

Set 3, Repeat Samples: Samples N and P have the same DNA concentration. Figure 5 is a "Youden Plot" of each participant's result for Sample P graphed against the result for Sample N. Only one participant reported greatly divergent results for these two repeat samples; however, six participants were consistently either much higher or much lower than the consensus result. The circle is an approximate 95% repeatability bound on the majority distribution, based upon a repeatability standard deviation of approximately *a factor* of 1.9. (Note the logarithmic distribution of the data: the one standard deviation repeatability range for a given value is $\text{Value}/1.9 \le \text{Value} \le 1.9 \times \text{Value}$.)

Sets 1 and 2, Total recoverable DNA: Table 5 summarizes the reported DNA concentrations of the Set 1 and 2 samples. On average, the total estimated recovery accounted for about 15% of the nominal amount of DNA in the stain.
Based upon limited results, a large fraction of the total DNA remains in the cloth after differential extraction.

A number of participants reported the DNA concentration of their extracts rather than total ng/stain; where possible, we have done the appropriate conversion

$$\text{Total ng DNA/stain} = \left(\frac{\text{Area}_{\text{Stain}}}{\text{Area}_{\text{Extracted}}}\right)\left(\frac{\text{ng DNA}}{\mu\text{L}}\right)\mu\text{L} .$$

For the majority of participants, there was agreement between the results for the Set 1 and 2 recoveries and the Set 3 "calibration curve" when both are expressed relative to the consensus (median) result.

Table 5
DNA Quantity

Sample	Design ng/Stain	N	Reported ng/Stain Median	Min	Max
F	909	40	368	2.3	3000
G	1176	40	146	20.0	8800
H	3304	39	387	18.8	2600
$H_?$	752	27	160	3.3	1625
$H_?$	2552	29	150	9.4	2000
H_{cloth}		3	375	125	411
J	1681	40	367	9.1	4000
K	758	40	193	12.5	8800
L	2213	39	320	6.5	1700
$L_?$	1689	27	181	3.3	1417
$L_?$	524	28	52	3.3	283
L_{cloth}		3	200	200	480

Sets 1 and 2, Differential Extraction: Table 5 also summarizes the reported DNA concentrations of the differential extractions of the mixed stain Samples H and L. All participants achieved a "clean" male fraction. By the nature of the semen (over age material from a commercial sperm bank), all "female/epithelial" fractions contained a larger than typical amount of DNA from the male donors.

Fig. 1 Calling the three banded pattern

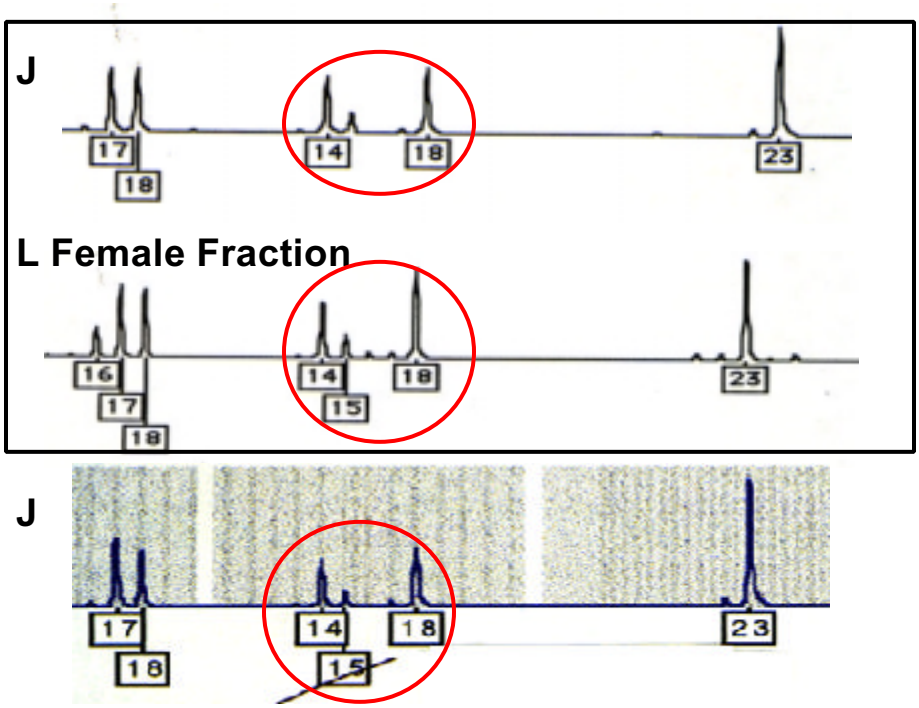


Fig 2 Peak Heights match Alleles to Donors ?

H Male Fraction

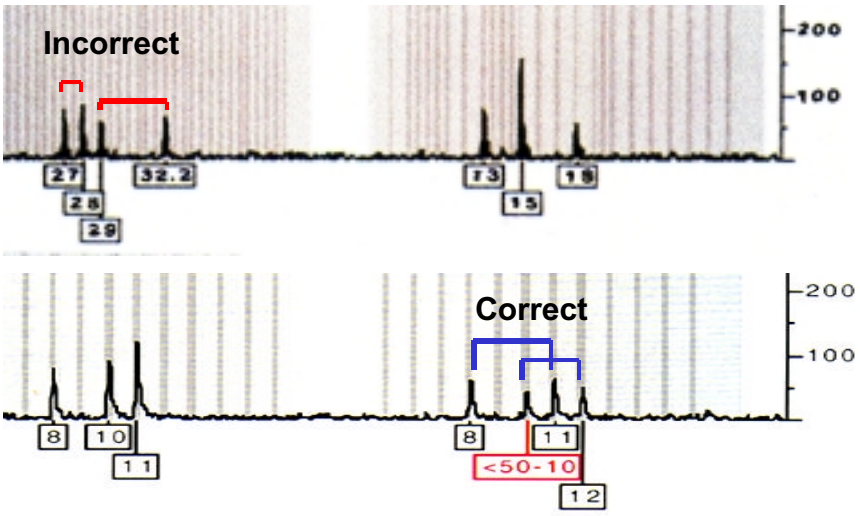


Fig. 3 Peak Height Variability with “Kit” Used

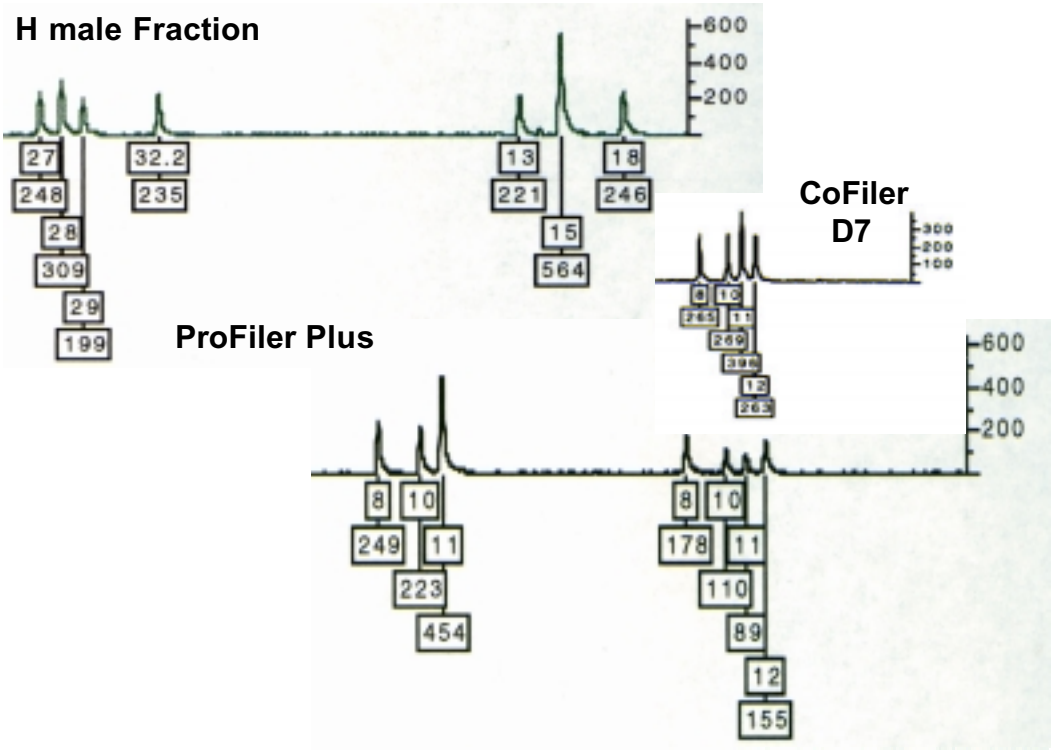
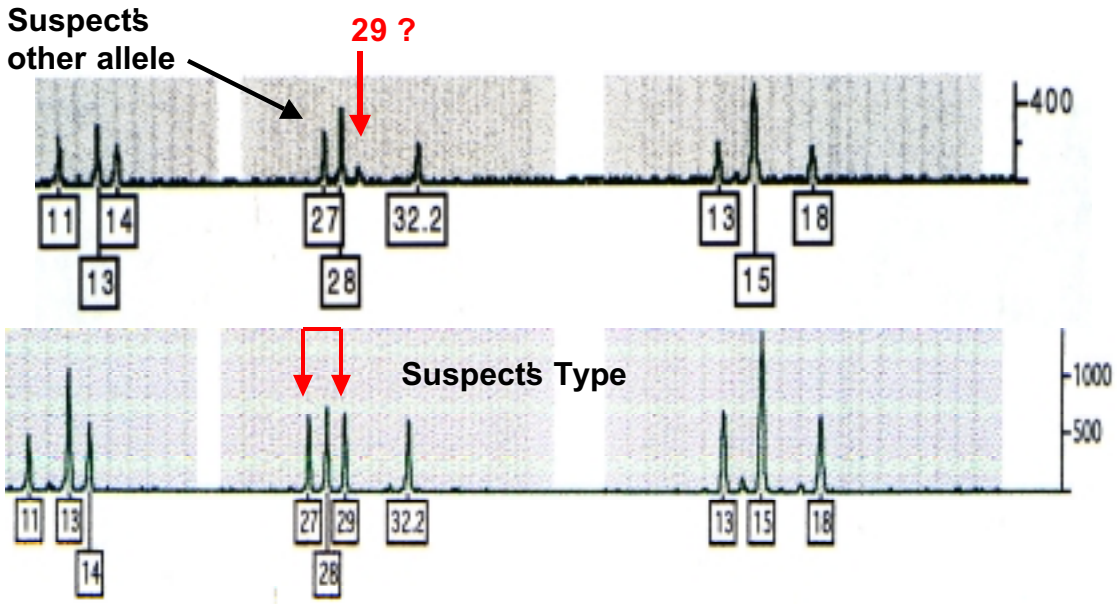


Fig. 4 Non-Amplifying Allele D21S11 locus



H Male Fraction from two different laboratories

Fig. 5 “Youden Plot” Sample P plotted against Sample N