Upgrading to STRmix™ Version 2.11: A Chance to Do It All Again

Ut N. Dinh, MS and Alicia M. Cadenas, MS • Signature Science Forensic DNA Laboratory, Austin, TX

ABSTRACT

In life second chances may not always come along but in the field of forensics, where new or improved technology continues to expand the limits of DNA testing, that is routinely not the case. One such technology is probabilistic genotyping, specifically STRmix, which has undergone various version upgrades since its original introduction to forensic DNA testing laboratories. With each new version release, laboratories must consider which additional validation studies should be undertaken prior to implementation in casework. Standards and guidelines are in place for the validation of new methodology in the laboratory to assist in forming a basis for establishing reliable methods, but it is ultimately up to the laboratory to determine which studies are relevant for a given method and to identify additional studies that can further assist in defining method limitations/parameters.

Signature Science (SigSci), a private Forensic DNA Laboratory based in Austin, Texas, initially validated STRmix v2.5 in 2018, however, validation efforts have been ongoing. Several validation addendums have been completed as well as an upgrade to STRmix™ v2.11. A thorough validation was initially performed prior to casework implementation to include the examination of known and non-probative evidence samples and investigations into reproducibility and precision, sensitivity and stochastic effects, and mixture data. As is often the case with newly implemented methodology, SigSci identified a need to explore and characterize additional parameters of the STRmix software functionality prior to any version upgrades. Interpretation guidelines were updated to evaluate uninformative comparisons as a result of potential familial relationships in casework mixture data. A validation addendum was also performed to evaluate drop-in and stutter modeling.

In addition to the original and addendum studies that were performed when validating v2.5, the v2.11 validation studies included the characterization of additional parameters. Familial relationships, specifically relationship likelihood ratios generated, were explored to establish a familial diagnostic range for use in casework. Additional challenging samples affected by degradation and/or inhibition were included to further assess features within STRmix modeling. Finally, a wider range of mixture samples were prepared and analyzed for the integration of interpretation thresholds in the laboratory. These validation efforts ultimately establish more robust and definitive guidelines for DNA profile interpretation, specifically regarding samples and/or contributors not suitable for comparison.

INTRODUCTION

In addition to our original validation studies when validating STRmix 2.5 we explored:

- Degraded and/or inhibited samples
- Wider set of mixture samples
- Familial diagnostics
- New interpretation guidelines
- Expanded drop-in study

This poster will focus on familial diagnostics due to space constraints.

Forensic DNA analysis has evolved significantly with the advent of probabilistic genotypic software such as STRmix. Traditional methods of DNA interpretation struggled with complex mixtures, particularly when contributors were related. STRmix provides a powerful approach by generating likelihood ratios (LRs) to quantify the strength of evidence for or against a given hypothesis. However, familial relationships introduce additional complexity, as related individuals may exhibit similar genetic profiles, potentially leading to misinterpretation.

This study focuses on evaluating familial likelihood ratios in STRmix and establishing diagnostic thresholds to flag potential misassignments. A range of mixtures from the sensitivity study were supplemented with 41 mixtures across two amplification kits. The 41 mixtures included familial and non-familial donors to determine empirical LR ratios for true contributors and to identify cases where comparisons to relatives may inflate LRs.

Reported LR	The DNA profile obtained from the sample is approximately 130 million times more likely if the sample originated from the POI and an unknown person than if it originated from two unknown persons.
Sibling LR	The DNA profile obtained from the sample is approximately 50 times more likely if the sample originated from the **POI and an unknown**, unrelated person than if it originated from a **sibling of the POI and an unknown**, unrelated person.

MATERIALS AND METHODS

402 mixtures from the sensitivity study (2p, 3p and 4p) were evaluated for TRUE NOC=APPARENT NOC and were included along with 41 mixtures comprising 29 familial and 12 non-familial mixtures of 2, 3, 4, and 5 persons. All mixtures were amplified using GlobalFiler and Investigator 24plex QS systems and deconvoluted through STRmix v 2.11. LRs were generated for all known contributors and relatives of known contributors (when applicable).

Of the familial mixtures, the TRUE NOC was apparent for a majority of the mixtures; as a result of allele sharing a subset of the mixtures appeared to be TRUE NOC-1. These mixtures were evaluated at both TRUE NOC and APPARENT NOC.

Family A		
Mixture	Proportions	Known Contributor
1	1 : 1	D3 : S3
2	1 : 3	D3 : S3
3	1 : 10	D3 : S3
4	1 : 1	M2 : M3
5	1 : 3	M2 : M3
6	1 : 10	M2 : M3
7	1 : 1	M3 : UK1
8	1 : 10	M3 : UK1
9	1 : 1 : 1	M1 : M2 : M3
10	1 : 2 : 5	M1 : M2 : M3
11	1 : 1 : 3	M2 : M3 : UK1
12	1 : 1 : 1	M3 : UK1 : UK2
13	1 : 1 : 3	UK1 : M3 : UK2
14	1 : 2 : 3 : 5	M3 : UK1 : UK2 : UK3
15	1 : 2 : 3 : 5	UK1 : UK2 : M3 : UK3
16	1 : 1 : 1 : 1	M1 : M2 : M3 : UK1
17	1 : 2 : 3 : 5	M1 : M2 : M3 : UK1
18	1 : 1 : 1 : 1	D1 : D2 : D3 : UK1
19	1 : 2 : 3 : 5	D1 : D2 : D3 : UK1
20	1 : 1 : 1 : 1	D2 : S2 : D3 : S3
21	1 : 1 : 1 : 1 : 1	D1 : D2 : S2 : D3 : S3

Family B		
Mixture	Proportions	Known Contributor
22	1 : 1	D4 : UK1
23	1 : 3	D4 : UK1
24	1 : 10	D4 : UK1
25	1 : 1	M4 : D4
26	1 : 3	M4 : D4
27	1 : 10	M4 : D4
28	1 : 1 : 1	M4 : D4 : GD4
29	1 : 2 : 5	M4 : D4 : GD4
30	1 : 1 : 1	GD4 : GS4 : D4
31	1 : 2 : 5	GD4 : GS4 : D4
32	1 : 1 : 3	D4 : UK1 : UK2
33	1 : 1 : 1	M4 : D4 : UK1
34	1 : 2 : 5	UK1 : D4 : UK2
35	1 : 1 : 1 : 1	M4 : D4 : GD4 : UK1
36	1 : 2 : 3 : 5	M4 : D4 : GD4 : UK1
37	1 : 1 : 1 : 1	D4 : UK1 : UK2 : UK3
38	1 : 2 : 3 : 5	D4 : UK1 : UK2 : UK3
39	1 : 2 : 3 : 5	UK1 : UK2 : D4 : UK3
40	1 : 1 : 1 : 1	UK1 : M4 : D4 : UK2
41	1 : 2 : 3 : 5	UK1 : M4 : D4 : UK2

M=Mother D=Daughter S=Son GD=Granddaughter
GS=Grandson UK=Unrelated Individual

Establishing Variance Ranges

Evaluate all mixtures where true contributors are being accurately assigned. → **Calculate the Log-Stratified LR / Log-Stratified Relative LRs.** → **Sibling and Parent/Child LRs were most affected. Therefore, diagnostic ranges were developed for these relationships**

When POI LR ≥100, and Parent/Child (P/C) and Sibling LRs are >1, Standards from first degree relatives are requested when:

$$\frac{\text{Average Log } (\textit{Stratified Total LR})}{\text{Average Log } (\textit{Stratified Sibling LR})} = 3.11 \qquad \frac{\text{Average Log } (\textit{Stratified Unified LR})}{\text{Average Log } (\textit{Stratified Sibling LR})} = 2.71$$

$$\frac{\text{Average Log } (\textit{Stratified Total LR})}{\text{Average Log } (\textit{Stratified P/C LR})} = 2.13 \qquad \frac{\text{Average Log } (\textit{Stratified Unified LR})}{\text{Average Log } (\textit{Stratified P/C LR})} = 1.89$$

When POI LR≥100, and Parent/Child and Sibling LRs are <1, Standards from first degree relatives are requested.

FALSE EXCLUSIONS

Of 125 comparisons of true known contributors:

- 5 false exclusions were observed at the TRUE NOC
- 22 false exclusions were observed at the APPARENT NOC

Case Study – Mix9 – A Tale of Three Sisters

1:1:1 M1:M2:M3 – Apparent 3:1 2p Mixture

If red flags are seen in casework, reference samples from first degree relatives would be requested

NOC	Proportions			LR			Parent/Child LR			Sibling LR		
	C1	C2	C3	M1	M2	M3	M1	M2	M3	M1	M2	M3
3	61.7	37.7	0.52	0.99	0.62	0.60	1.90	**1.47E-12**	**8.76E-12**	**8.64E-16**	**7.48E-16**	**7.65E-16**
2	61.7	38.2	0	0	0	0	0	0	0	0	0	0

Diagnostic Red Flags:

- = Parent/Child
- = Sibling

Other Red Flags:

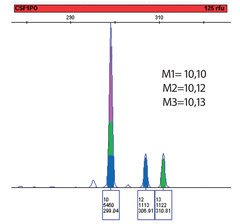
- For deconvolutions at apparent NOC of 2, all three sisters were assigned as Contributor 1
- Full representation of each donor with LR=0 at one or two loci due to flipping between C1 and C2

D2S1338 M1 = 16, 20 M2 = 17, 23 M3 = 17, 23

Contributor	Genotype	Weight	Component ≥ 99%
Contributor 1 (61.75%) (Q)	17, 23	100.00%	17, 23
Contributor 2 (38.25%) (Q)	16, 20	100.00%	16, 20

CSF1PO M1 = 10, 10 M2 = 10, 12 M3 = 10, 13

Contributor	Genotype	Weight	Component ≥ 99%
Contributor 1 (61.75%) (Q)	10, 10	100.00%	10, 10
Contributor 2 (38.25%) (Q)	12, 13	100.00%	12, 13



Observed Mx Priors Calculations

Total RFU @ CSF1PO = (5450 + 1113 + 1122) = 7685

M1 = 10,10 = (5450-1113-1122)/7685= 0.42

M2 = 10,12 = (1113 * 2) / 7685 = 0.29

M3 = 10,13 = (1122 * 2) / 7685 = 0.29

Assumption Runs

NOC 3 Assumed Donor	Proportions			LR		
	C1	C2	C3	M1	M2	M3
M1	51.3	48.1	0.52	--	0	0
M2	57.1	42.4	0.44	0	--	0
M3	64.2	35.3	0.41	0	0	--
M1,M2	33.7	38.3	27.9	--	--	5.01E+24
M2,M3	32.1	34.0	33.7	2.82E+27	--	--
M1,M3	31.5	36.4	31.9	--	4.98E+25	--

I. Assuming two sisters in the mixture resolved the proportions and false exclusions.

Informed Mx Priors Runs

NOC	Input Mx Priors			LR		
	C1	C2	C3	M1	M2	M3
3	0.34	0.33	0.33	7.92E+18	2.55E+17	1.76E+17
3	0.42	0.29	0.29	6.19E+18	1.08E+17	1.18E+17

II. Using Mx Priors resolved the false exclusions.

FALSE INCLUSIONS

Of 230 comparisons of true known non-contributors:

- 13 false inclusions were observed at the TRUE NOC
- 1 false inclusion was observed at the APPARENT NOC

All false inclusions of TRUE and APPARENT NOC were positively flagged by familial diagnostic ratio or by relative LRs <1.

Case Study – Mix31 – My Favorite Uncle

1:2:5 GD4:GS4:D4 – Brother/Uncle (S4) Assigned as Contributor 2

If red flags are seen in casework, reference samples from first degree relatives would be requested

LR Proportion Stratified

Relation of Unknown in HD to POI	Lower HPD
Unrelated	**9.4412E11**
Sibling	**6.6845E3**
Parent/Child	**3.2033E5**
Half Sibling	6.7637E7
Grandparent/Grandchild	6.7637E7
Uncle or Aunt/Nephew or Niece	6.7637E7
Cousin	3.4761E9
Unified	**5.2901E11**

Log (9.44E11) / Log (6.68E3) = 3.13

Log (9.44E11) / Log(3.20E5) = 2.17

Diagnostic Red Flags:

$$\frac{\text{Log } (\textit{Stratified LR})}{\text{Log } (\textit{Stratified Sib LR})} = 3.13 \qquad \frac{\text{Log } (\textit{Stratified LR})}{\text{Log } (\textit{Stratified P/C LR})} = 2.17$$

$$\frac{\text{Log } (\textit{Stratified Unified LR})}{\text{Log } (\textit{Stratified Sibling LR})} = 3.06 \qquad \frac{\text{Log } (\textit{Stratified Unified LR})}{\text{Log } (\textit{Stratified P/C LR})} = 2.12$$

Other Red Flags:

Unintuitive Genotype weights, visual comparison of S4 (i.e. D8S1179, S4 = 13, 16)

D8S1179

Contributor	Genotype	Weight	Component ≥ 99%
Contributor 1 (60.56%) (Q)	15, 16	95.35%	15, F
	15, 15	4.65%	
Contributor 2 (25.05%) (Q)	14, 15	47.46%	INC
	13, 14	18.19%	
	15, 15	17.62%	
	13, 15	10.42%	
	16, 16	3.11%	
	14, 16	1.68%	
	15, 16	0.56%	
	13, 16	**0.50%**	
	14, 14	0.45%	
	13, 13	0.00%	
Contributor 3 (25.05%) (Q)	13, 15	44.29%	INC
	13, 14	22.81%	
	15, 15	11.82%	
	14, 15	9.19%	
	13, 16	4.51%	
	14, 14	2.44%	
	14, 16	1.02%	
	15, 16	0.98%	
	13, 13	0.80%	
	16, 16	0.48%	
	13, Q	0.34%	
	15, Q	0.34%	
	Q, Q	0.34%	
	14, Q	0.33%	
	16, Q	0.32%	

CONCLUSION

This study highlights the importance of understanding familial likelihood ratios in forensic DNA analysis using STRmix. Our findings demonstrate that the presence of related individuals in a DNA mixture can significantly influence likelihood ratio calculations, sometimes leading to inflated LRs. By defining diagnostic thresholds and evaluating empirical LR ranges, we have established a framework to flag inflated LRs for true non-contributors and false exclusions of true contributors (LR=0). The value of using the stratified total LR versus the stratified unified LR is being further investigated to determine which will be used in casework. Initial evaluations showed benefits for both as a diagnostic. Implementing familial diagnostic flags provides forensic analysts with additional tools to scrutinize evidence more effectively. These findings support the continued refinement of STRmix interpretation guidelines and emphasize the need for additional testing when familial relationships may be a factor in casework. Future studies should focus on expanding validation datasets and refining threshold values to improve forensic DNA interpretation further. By refining these interpretations, we aim to improve forensic DNA analysis and provide guidance for casework involving complex relationships.