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

MixMaSTR: a Software Package for Designing and Interpreting Forensic DNA Validation Studies

Gordon Research Conference (GRC)
Forensic Analysis of Human DNA
Newry, Maine (June 22 – 27, 2025)

For more information on the release date or to request a trial version, please reach out to us at: mixmastr@nist.gov.

Sarah Riman[1], Hari Iyer[1], Sarra Chouder [1], Edgar Robitaille[2], Benjamin J. Long [1], Peter M. Vallone[1]

[1] U.S. National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg, MD 20899-8314, USA
[2] Johns Hopkins University Whiting School of Engineering, Department of Computer Science, Baltimore, MD 21218 USA

Funding: This work was supported by the Accelerating Forensic Innovation for Impact through the *NIST Special Programs Office: Forensic Genetics.*
Disclaimer: Points of view in this poster are those of mine and do not necessarily represent the official position or policies of the U.S. Department of Commerce. Certain commercial software, instruments, and materials are identified in order to specify experimental procedures as completely as possible. In no case does such identification imply a recommendation or endorsement by NIST, nor does it imply that any of the materials, instruments, or equipment identified are necessarily the best available for the purpose. All work involving NIST samples has been reviewed and approved by the NIST Research Protections Office.

Background:

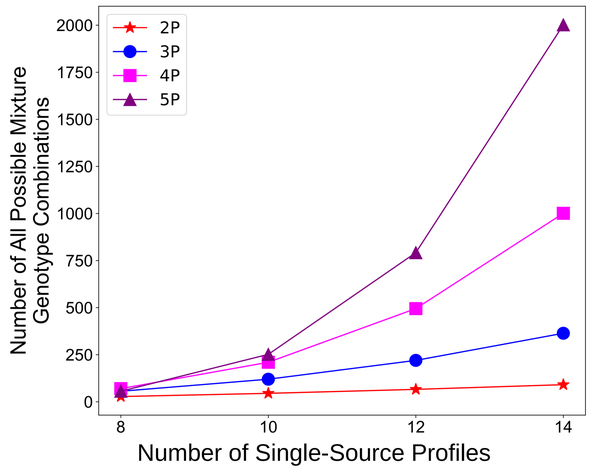
- Protocols implemented by forensic laboratories to analyze casework samples must be assessed for performance by validation studies.
- Validation is not a one-time process; laboratories revalidate their methods when changes or upgrades are made to their workflows.
- Forensic practitioners face the challenges of conducting validation studies, which require significant time and resources.
- The field lacks a much-needed open-source software tool to overcome these challenges.

Motivation:

- Develop a standalone, open-source, and easy-to-use application intended to facilitate the design of validation experiments and the interpretation and visualization of the resulting data.

MixMaSTR Key Features in Designing Validation Experiments

1 Construction of all possible mixture genotype combinations



Mixture Genotype Combinations Combinatorial Formula

Combinations, $ {}_nC_r = \frac{n!}{r!(n-r)!} $

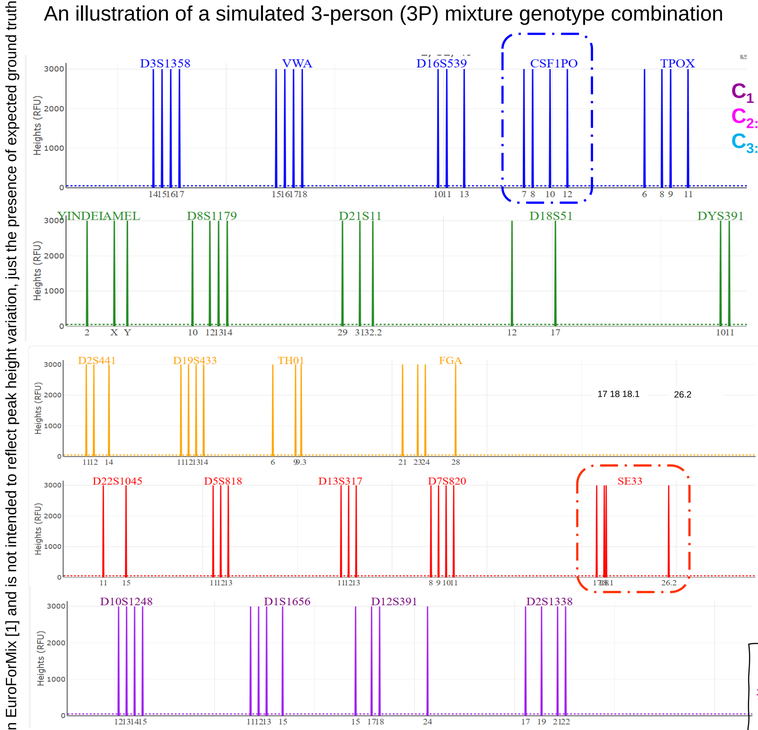
${}_nC_r$ = number of possible genotype combinations
n = number of single-source profiles
r = number of contributors

- Supports loading user-provided ***ground truth single-source profiles*** genotyped by any STR multiplex kit.
- Generates all possible mixture genotype combinations depending on the ***number of single-source profiles and NoCs*** chosen.

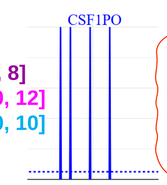
2 Computation of various metrics of interest for each constructed mixture combination

- Allele Sharing Ratio (ASR)
- Counts of homozygote genotypes
- Instances of rare alleles
- Instances of allele-allele 1-bp difference
- Maximum allele count (MAC) → MAC NoC

An illustration of a simulated 3-person (3P) mixture genotype combination



Locus metrics



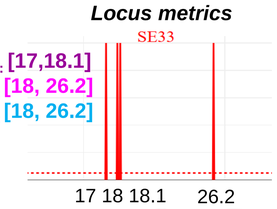
C_1: [7, 8]
C_2: [10, 12]
C_3: [10, 10]

- User selected or uploaded allele frequency/population database.
- User selected rare allele threshold.

Expected CSF1PO Alleles :	[7, 8, 10, 12]
# of alleles	4
Homozygosity	C3 [10,10] | 1
A-A 1 bp	[ ] | 0
Rare Alleles*	C1 [7] | 1

*Based on the NIST US Caucasian population data

Locus metrics



C_1: [17,18.1]
C_2: [18, 26.2]
C_3: [18, 26.2]

Expected SE33 Alleles :	[17, 18, 18.1, 26.2]
# of Alleles	4
Homozygosity	[ ] | 0
A-A 1 bp	C2 [18], C3 [18], C1 [18.1] | 1
Rare Alleles	[ ] | 0

Profile metrics: *Summary statistics across each simulated mixture combination (N=21 loci)*

Mixture Combinations	eNoC	MAC NoC	∑ Homozygote Genotypes	∑A-A 1 bp Difference	∑ Rare Alleles	ASR
C_1,C_2, C_3	3	2	13	1	1	0.56

This illustration was created in EuroForMix [1] and is not intended to reflect peak height variation, just the presence of expected ground truth genotypes.

3 Choosing a Validation Experimental Design

Using ***Statistical Theory of Experimental Design such as Space-Filling Design and Fractional Factorial Design*** and considering a ***laboratory available resources***, the software will output candidate experimental plans to ensure reasonable coverage of the factor space based on user specifications.

Illustration of 3P Mixture Experimental Design

The user will input the factor space desired to be covered as shown in this illustration

eNoC ☐1 ☐2 ☑3 ☐4 ☐5

# of PCR reactions 240

Serial dilution of DNA 10

DNA Quality ☑ ***0*** *(Pristine)* ☑ ***2*** *(2 Cs are degraded)* ☐ *Mixture prepared and then degraded* ☑ ***1*** *(1 C is degraded)* ☑ ***3*** *(All 3 Cs are degraded)*

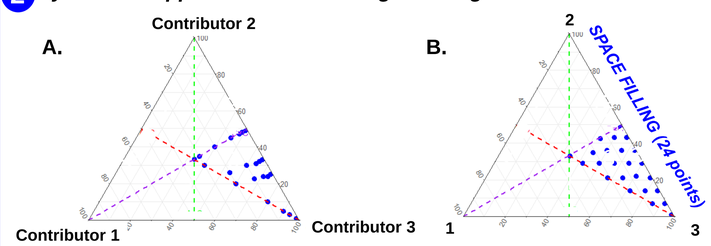
Mixture Ratios 24

MixMaSTR Output

1 ***Unique Mixture Genotype Combinations*** 24

Min	Max	
0.3	0.7	***ASR***
15	33	***Homozygosity***
2	5	***Rare Alleles***
0	1	***A-A 1bp***
1	3	***MAC NoC***

2 ***Systematic approach for examining coverage of mixture ratios***



(A) Ternary plot showing mixture ratios from a validation study (from a collaborator lab).
(B) Ternary plot showing mixture ratios provided by MixMaSTR.
Each point in blue represents a unique mixture ratio.

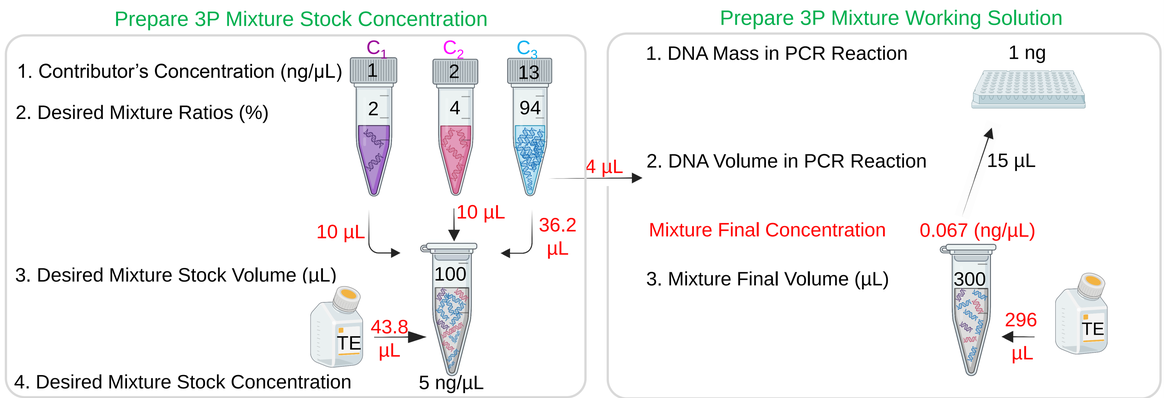
3 **A candidate experimental design**

MGC	C1	C2	C3	R1 (%)	R2 (%)	R3 (%)	Degradation
1	1	11	20d	7	7	86	1
2	2d	11d	20d	7	36	57	3
3	1	11	13	21	29	50	0
4	1	9d	11d	1	7	92	2
5	11d	12d	13d	1	1	98	3
6	1	11	15d	1	36	63	1
7	10	11	13d	21	35	44	1
8	4	12	20d	14	21	65	1
9	3d	5d	16d	1	14	85	3
10	5	10	12	14	43	44	0
11	5	9	19	7	29	64	0
12	1d	6d	7d	7	22	71	3
13	6d	13d	16d	1	49	50	3
14	4d	7d	14d	1	21	78	3
15	1	6	10	21	21	58	0
16	9	15d	18d	7	14	79	2
17	3	15d	17d	1	43	56	2
18	4	15	17	14	14	72	0
19	3	5d	15d	14	36	50	2
20	3	15	16d	14	29	57	1
21	3	7d	15d	29	29	42	2
22	4	14	19d	1	29	70	1
23	2	3	15	33	33	34	0
24	1	8d	12d	7	43	50	2

MGC = Mixture genotype combination
C = Contributor
R = Ratio
d = degradation

4 Mixture Calculations

The software will take user's requirements (e.g., C's concentration, desired mixture ratios) and constraints (minimum pipetting amounts, DNA mass in PCR reaction, minimum mixture stock solution) and provide an efficient strategy for making the desired mixtures.



[1] Ø. Bleka, G. Storvik, P. Gill, EuroForMix: An open source software based on a continuous model to evaluate STR DNA profiles from a mixture of contributors with artefacts, Forensic Science International: Genetics 21 (2016) 35-44.