Forensic Science International: Genetics 37 (2018) 81–94

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsigen

# NIST interlaboratory studies involving DNA mixtures (MIX05 and MIX13): Variation observed and lessons learned

John M. Butler[a,*], Margaret C. Kline[b], Michael D. Coble[b,1]

[a] National Institute of Standards and Technology, Special Programs Office, Gaithersburg, MD 20899, United States
[b] National Institute of Standards and Technology, Applied Genetics Group, Gaithersburg, MD 20899, United States

ARTICLE INFO

*Keywords:*
Forensic DNA
DNA mixture
Mixture interpretation
Interlaboratory study
Collaborative exercise
MIX05
MIX13
Forensic science

ABSTRACT

Interlaboratory studies are a type of collaborative exercise in which many laboratories are presented with the same set of data to interpret, and the results they produce are examined to get a "big picture" view of the effectiveness and accuracy of analytical protocols used across participating laboratories. In 2005 and again in 2013, the Applied Genetics Group of the National Institute of Standards and Technology (NIST) conducted interlaboratory studies involving DNA mixture interpretation. In the 2005 NIST MIX05 study, 69 laboratories interpreted data in the form of electropherograms of two-person DNA mixtures representing four different mock sexual assault cases with different contributor ratios. In the 2013 NIST MIX13 study,108 laboratories interpreted electropherogram data for five different case scenarios involving two, three, or four contributors, with some of the contributors potentially related. This paper describes the design of these studies, the variations observed among laboratory results, and lessons learned.

## 1. Introduction

Interlaboratory comparisonstudies, which are sometimes referred to as collaborative exercises or round-robin studies, provide a useful way to demonstrate that multiple laboratories can generate comparable results with the same provided samples, and are cited as valuable methods for assessing measurement reproducibility in accredited laboratories [1]. Interlaboratory studies are regularly used in clinical DNA diagnostics (e.g., [2]) and other scientific communities. Given that DNA databases used in criminal investigations compile data from many jurisdictions, it is valuable to assess the degree to which laboratories across jurisdictions produce comparable analytical results.

Interlaboratory studies, which are typically voluntary, assess progress on the standardization of methods across laboratories and enable technical and statistical issues to be ascertained and discussed. Most analysts focus on their own laboratory protocols and rarely get an opportunity to determine how their laboratory performs relative to others. Intralaboratory evaluations examine performance across analysts within the same laboratory and can be useful in assessing whether further training on following protocols is needed to improve consistency. Both intra- and inter-laboratory studies can help better understand causes of variability among laboratories and analysts – and hopefully lead to improvement of the entire community.

It is important to recognize that interlaboratory studies tend to be research-focused and are not meant to evaluate the performance of individual analysts. Although errors made by laboratories are noted in interlaboratory study publications, finding these errors is not typically the primary objective of a study. Any errors detected reveal opportunities for improvement (see Ref. [3]) based on the research question being explored and cannot normally be used to formally assess operational error rates for a general activity as has been advocated for proficiency test data that is produced under standard conditions (see Ref. [4]).

DNA mixtures arise when biological material from two or more individuals contributes to the sample being tested, and different types or categories of mixtures have been proposed [5]. Methods for deconvoluting mixtures were first described about two decades ago [6]. In 2006, the DNA Commission of the International Society of Forensic Genetics (ISFG) stated in their "Recommendations on the interpretation of mixtures" article that "our discussions have highlighted a significant need for continuing education and research into this area" [7]. Interlaboratory studies can enable monitoring of variability in practice and overall laboratory performance with different types of DNA mixtures.

About a dozen interlaboratory studies exploring DNA mixture interpretation with short tandem repeat (STR) markers have been performed over the past two decades [8–18] to examine various aspects of

* Corresponding author at: NIST Special Programs Office, 100 Bureau Drive, Mail Stop 4701, Gaithersburg, MD 20899-4701, United States.
*E-mail address:* john.butler@nist.gov (J.M. Butler).
[1] Current address: University of North Texas Health Science Center, Fort Worth, TX 76107, United States.

https://doi.org/10.1016/j.fsigen.2018.07.024
Received 26 April 2018; Received in revised form 24 July 2018; Accepted 31 July 2018
Available online 01 August 2018

**Table 1**
Interlaboratory studies involving STR multiplexes and DNA mixtures.

| Study (when conducted) | Publication | # Labs (# data sets) | # Mixtures | Samples Provided and Study Purpose |
|---|---|---|---|---|
| STR triplex CTT (Dec 1995–May 1996) | Kline et al. [8] | 34 (46) | 0 | 4 single-source DNA extracts, 4 single-source stains to explore factors affecting sizing variability |
| NIST Mixed Stain Study (MSS) 1 (April–Nov 1997) | Duewer et al. [9] | 22 (37) | 5 | Buffy coat cells on S&S 903 paper; 6 single-source, 4 two-source mixtures, and 1 three-source mixture to explore donor types obtained given a complete set of reference sources |
| NIST MSS2 (Jan–May 1999) | Duewer et al. [9] | 45 (70) | 2 | Part A: 4 single-source stains, 1 two-source stain, 1 three-source stain; Part B: 5 vials of a four-level DNA concentration series to explore donor types obtained given an incomplete set of reference sources (Part A) and to examine performance of DNA quantitation assays (Part B) |
| NIST MSS3 (Dec 2000–Oct 2001) | Kline et al. [10] and Duewer et al. [11] | 74 (117) | 6 | DNA extracts; 1 single-source, 5 two-source mixtures (3:1 to 10:1 component ratios), and 1 three-source mixture (4:2:1) to explore the effect of quantitation on STR typing performance |
| GHEP-MIX01 (2010) | Crespillo et al. [12] | 32 (32) | 4 | Questionnaire and data (.fsa files) provided with 2 STR kits (Identifiler and PP16) for 4 two-source mixtures (1F:5 M, 1F:10 F, 1F:1 M, 5F:1 M) to explore errors (discrepancies) obtained during mixture interpretation |
| GHEP-MIX02 (2011) | Crespillo et al. [12] | 24 (24) | 2 | Questionnaire and data (.fsa files) provided with 1 STR kit (Identifiler) for 1 two-person mixture (1M:5 F) and 1 three-person mixture (2F:1M:1 M) to explore statistical treatment of results under a common set of hypotheses |
| GHEP-MIX03 (2012) | Crespillo et al. [12] | 17 (17) | 3 | Questionnaire and data (.fsa files) provided with 2 STR kits (Identifiler Plus and NGM) for 2 two-person mixtures (1F:5 M, 1F:10 F) and 1 three-person mixture (1F:3M:7 M) to explore statistical treatment of results under an open set of hypotheses for the likelihood ratios used |
| EuroForGen-NoE (2013) | Prieto et al. [13] | 18 (20); 18 (22) | 2 | Data (csv format) provided for 16 STR loci with case scenarios; two exercises each involving a two-person mixture were supplied along with victim and suspect profiles, population allele frequencies, and LRmix software to explore impact of training and whether standardization of an approach could be demonstrated |
| UK Forensic Science Regulator (2014) | Unpublished report | 8 (18) | 5 | DNA extracts provided with case scenarios for 2 two-person mixtures (4:1, 2:1) and 3 three-person mixtures (6:4:1, 6:3:1, 7:1.5:1) to explore variability across UK forensic science providers |
| DFSC Mixture Study (2014–2015) | Aranda [14] *(presentation only)* | 55 (185) | 6 | Data provided for 15 STR loci (Identifiler Plus) involving 4 two-person mixtures and 2 three-person mixtures to explore intra- and inter-laboratory variation in genotype determinations to better understand the current state and potential limitations of mixture interpretation |
| STRmix study (2014) | Cooper et al. [15] | 12 (20) | 3 | Data provided for 15 STR loci (Identifiler) involving three casework samples (ground truth not known) to explore the improved level of agreement that was possible within and between laboratories using a common probabilistic genotyping software program |
| 22nd GHEP-ISFG IE Basic (2014) | Toscanini et al. [16] | 72 | 1 | Two-source stain: 2:1 mixture (v/v) saliva/blood; results generated with autosomal STRs, Y-STRs, X-STRs, and mtDNA; to explore various approaches being used for mixture interpretation and technical difficulties observed |
| 22nd GHEP-ISFG IE Advanced (2014) | Toscanini et al. [16] | 52 | 1 | Two-source stain: 4:1 mixture (v/v) saliva/semen; results generated with autosomal STRs, Y-STRs, X-STRs, and mtDNA; to explore various approaches being used for mixture interpretation and technical difficulties observed |
| GHEP-MIX06 (2015) | Barrio et al. [18] | 25 | 2 | Data (pdf files) provided for 15 STR loci (NGM) involving a three-person (7M:3F:1 M) mixture and 17-YSTRs (Yfiler) involving a two-male (3:1) mixture; participants were provided with mock case information and analytical, stochastic, and stutter thresholds used; to explore how results would be reported if this exercise were a real case |
| NFI-led study (2016) | Benschop et al. [17] | 3 (26) | 10 | Data (pdf files) provided for 15 STR loci (NGM) with replicates involving 2 two-person (1:1, 5:1), 4 three-person (1:1:1, 5:1:0.2, 10:1:1, 10:1:1), 2 four-person (5:1:1:1, 5:1:1:1), and 2 five-person (2:2:1:1:1, 2:2:1:1:1) mixtures and some person of interest reference profiles to explore intra- and inter-laboratory variability |
| NIST MIX05 (Feb–Sept 2005) | *This article (and several presentations)* | 69 (75) | 4 | Data (.fsa files) provided from 6 STR kits; 4 two-person mixture "evidence" profiles (3F:1 M, 1F:3 M, 1F:1 M, 7F:1 M) with female "victim" reference profiles; no "suspect" male reference profiles supplied for comparison purposes to explore mixture deconvolution approaches |
| NIST MIX13 (Aug–Dec 2013) | *This article (and several presentations)* | 108 (163) | 5 | Data (.fsa files) provided from 2 STR kits with case scenarios; 5 "cases" involving two- (1:1, 3.5:1), three- (6:1.5:1, 7:2:1), or four- contributors (1:1:1:1) with "person of interest" reference profiles, some of which were not in the mixtures to explore variability in overall mixture interpretation |