

REPORT DOCUMENTATION PAGE			Form Approved OMB NO. 0704-0188		
<p>The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA, 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</p>					
1. REPORT DATE (DD-MM-YYYY) 28-09-2017		2. REPORT TYPE Final Report		3. DATES COVERED (From - To) 8-Sep-2014 - 7-Feb-2017	
4. TITLE AND SUBTITLE Final Report: A Tool for Determining the Number of Contributors: Interpreting Complex, Compromised Low-Template DNA Samples			5a. CONTRACT NUMBER		
			5b. GRANT NUMBER W911NF-14-C-0096		
			5c. PROGRAM ELEMENT NUMBER 111111		
6. AUTHORS Catherine Grgicak and Desmond Lun			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAMES AND ADDRESSES Trustees of Boston University, BUMC Office of Sponsored Programs 85 East Newton Street, M-921 Boston, MA 02118 -2340			8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES) U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211			10. SPONSOR/MONITOR'S ACRONYM(S) ARO		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S) 65847-LS-RIF.12		
12. DISTRIBUTION AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.					
14. ABSTRACT In forensic DNA analysis, the interpretation of a sample acquired from the environment may be dependent upon the assumption on the number of individuals from which the evidence arose. Degraded and damaged samples often exhibit signal that is lower than expected at the high molecular weights, resulting in a 'ski slope' effect. In this work we developed a method, and corresponding software tool, that provides a mechanism to determine the number of contributors (NOC) to an unknown stain, regardless of condition. This was accomplished by assessing all known confounding factors related to complex, low template DNA mixture interpretation. The software, named NOCI4					
15. SUBJECT TERMS forensic science, DNA mixtures, number of contributors					
16. SECURITY CLASSIFICATION OF:		17. LIMITATION OF ABSTRACT		15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT UU	b. ABSTRACT UU	c. THIS PAGE UU	UU		Catherine Grgicak
					19b. TELEPHONE NUMBER 617-638-1968

**RPPR Final Report**  
as of 17-Oct-2017

Agency Code:

Proposal Number: 65847LSRIF

**Agreement Number: W911NF-14-C-0096**

**INVESTIGATOR(S):**

**Name:** Ph.D. Catherine M. Grgicak

**Email:** cgrgicak@bu.edu

**Phone Number:** 6176381968

**Principal:** Y

Organization: **Trustees of Boston University, BUMC**

Address: Office of Sponsored Programs, Boston, MA 021182340

Country: USA

DUNS Number: 604483045

EIN: 1042103547

**Report Date:** 07-Mar-2017

Date Received: 28-Sep-2017

**Final Report** for Period Beginning 08-Sep-2014 and Ending 07-Feb-2017

**Title:** A Tool for Determining the Number of Contributors: Interpreting Complex, Compromised Low-Template DNA Samples

**Begin Performance Period:** 08-Sep-2014

**End Performance Period:** 07-Feb-2017

**Report Term:** 0-Other

Submitted By: Ph.D. Catherine Grgicak

Email: cgrgicak@bu.edu

Phone: (617) 638-1968

**Distribution Statement:** 1-Approved for public release; distribution is unlimited.

**STEM Degrees:**

**STEM Participants:**

- Major Goals:**
- a) Perfect the models which describe stutter within forensically relevant STR loci, and test the incorporation of a mixed model for stutter into NOCI. Modify the NOCI code/algorithm to incorporate forward stutter into the calculation,
  - b) Provide technical documentation for the software code,
  - c) Provide fully functional NOCI software, a NOCI technical manual and NOCI training material,
  - d) Prepare and generate 3,499 amplified work products of compromised, STR samples using the Identifiler Plus Amplification Kit. Generate 3,490 amplified work products of compromised, STR samples using the GlobalFiler (or other mega-plex) Amplification kit.
  - e) Incorporate a quality parameter such that the distribution on the NOCs (number of contributors) for all samples, including degraded and inhibited samples, can accurately be computed,
  - f) Decrease the run-time of NOCI,
  - g) Development of the NOCI interface. Key features of the completed software GUI include,
    - i) Graphical interface for calibration.
    - ii) Batch input.
    - iii) Graphical interface for output.
    - iv) Ability of user to use the option of an analytical threshold.

**Accomplishments:** See attachment

**Training Opportunities:** Nothing to Report

## RPPR Final Report as of 17-Oct-2017

**Results Dissemination:** October 2-5th, 2016 27th International Symposium on Human Identification. The Development and Release of a Collection of Computational Tools and A Large-Scale Empirical Database for Validation: The PROVEDIt Initiative. Madison, WI.

June 20th, 2016 1st Gordon Research Conference: Forensic Analysis of Human DNA – Catherine M. Grgicak, Ken R. Duffy, Ullrich Moenich, Muriel Medard, Desmond S. Lun, Neil Gurrum, Jennifer Sheehan, Kelsey C. Peters. “Forensic DNA Research, Validation and Pedagogy Using Empirical and Simulated Data: Understanding the Behavior of Mixtures” Waterville Valley, NH

June 22nd, 2016 1st Gordon Research Conference: Forensic Analysis of Human DNA – Slim I. Karkar, Harish Swaminathan, Lauren Alfonse, Catherine M. Grgicak, Desmond S. Lun. “NOCl: A Computational Tool to Infer the Number of Contributors to a Forensic Sample” Waterville Valley, NH

July 12th, 2016 28th International Biometric Conference WNAR Session: A continuous approach to interpreting forensic DNA profiles. Catherine M. Grgicak, Ken R. Duffy, Ullrich Mönich, Muriel Médard, Desmond S. Lun, Neil Gurrum, Kelsey C. Peters “Effects of laboratory and analysis decisions on the LR and its distribution” Victoria, British Columbia, Canada

**Honors and Awards:** Nothing to Report

**Protocol Activity Status:**

**Technology Transfer:** Nothing to Report

### **PARTICIPANTS:**

**Participant Type:** Faculty

**Participant:** Catherine M. Grgicak

**Person Months Worked:**

Project Contribution:

International Collaboration:

International Travel:

National Academy Member:

Other Collaborators:

**Funding Support:**

**Participant Type:** Faculty

**Participant:** Desmond Lun

**Person Months Worked:**

Project Contribution:

International Collaboration:

International Travel:

National Academy Member:

Other Collaborators:

**Funding Support:**

**Participant Type:** Graduate Student (research assistant)

**Participant:** Kelsey Peters

**Person Months Worked:**

Project Contribution:

International Collaboration:

International Travel:

National Academy Member:

Other Collaborators:

**Funding Support:**

**Participant Type:** Graduate Student (research assistant)

**Participant:** Xia Yearwood-Garcia

**RPPR Final Report**  
as of 17-Oct-2017

**Person Months Worked:**

Project Contribution:  
International Collaboration:  
International Travel:  
National Academy Member:  
Other Collaborators:

**Funding Support:**

**Participant Type:** Graduate Student (research assistant)

**Participant:** James Kelley

**Person Months Worked:**

Project Contribution:  
International Collaboration:  
International Travel:  
National Academy Member:  
Other Collaborators:

**Funding Support:**

**Participant Type:** Postdoctoral (scholar, fellow or other postdoctoral position)

**Participant:** Slim Karkar

**Person Months Worked:**

Project Contribution:  
International Collaboration:  
International Travel:  
National Academy Member:  
Other Collaborators:

**Funding Support:**

**CONFERENCE PAPERS:**

**Publication Type:** Conference Paper or Presentation

**Publication Status:** 0-Other

**Conference Name:** 1st Gordon Research Conference: Forensic Analysis of Human DNA

Date Received: 25-Aug-2016      Conference Date: 21-Jun-2016      Date Published: 21-Jun-2016

Conference Location: Waterville Valley, NH

**Paper Title:** Forensic DNA Research, Validation and Pedagogy Using Empirical and Simulated Data

**Authors:** Catherine M. Grgicak, Ken R. Duffy, Ullrich Moenich, Muriel Medard, Desmond S. Lun, Neil Gurrarn, Jer

Acknowledged Federal Support: **Y**

**Publication Type:** Conference Paper or Presentation

**Publication Status:** 0-Other

**Conference Name:** 1st Gordon Research Conference: Forensic Analysis of Human DNA

Date Received: 25-Aug-2016      Conference Date: 25-Aug-2016      Date Published: 25-Aug-2016

Conference Location: Waterville Valley, NH

**Paper Title:** NOCI: A Computational Tool to Infer the Number of Contributors to a Forensic Sample

**Authors:** Slim I. Karkar, Harish Swaminathan, Lauren Alfonse, Catherine M. Grgicak, Desmond S. Lun.

Acknowledged Federal Support: **Y**

## RPPR Final Report as of 17-Oct-2017

**Publication Type:** Conference Paper or Presentation **Publication Status:** 0-Other  
**Conference Name:** 28th International Biometric Conference WNAR Session: A continuous approach to interpreting forensic DNA profiles.  
Date Received: 25-Aug-2016 Conference Date: 12-Jul-2016 Date Published: 12-Jul-2016  
Conference Location: Victoria, British Columbia, Canada  
**Paper Title:** Effects of laboratory and analysis decisions on the LR and its distribution  
**Authors:** Catherine M. Grgicak, Ken R. Duffy, Ullrich Mönich, Muriel Médard, Desmond S. Lun, Neil Gurrām, Kels  
Acknowledged Federal Support: **Y**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published  
**Conference Name:** International Symposium on Human Identification  
Date Received: 17-Mar-2017 Conference Date: 27-Sep-2016 Date Published: 01-Oct-2016  
Conference Location: Minneapolis, MN  
**Paper Title:** The Development and Release of a Collection of Computational Tools and a Large-Scale Empirical Data Set for Validation: The PROVEDIt Initiative  
**Authors:** Lauren E. Alfonse, Amanda D. Garrett, Harish Swaminathan, Kelsey C. Peters, Genevieve Wellner, Xia  
Acknowledged Federal Support: **Y**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published  
**Conference Name:** American Academy of Forensic Sciences  
Date Received: 17-Mar-2017 Conference Date: 21-Feb-2017 Date Published: 25-Feb-2017  
Conference Location: New Orleans, LA  
**Paper Title:** The PROVEDIt Initiative: Development and Release of a Collection of Computational Tools and a Large-Scale Empirical Data Set for Forensic Research and Validation  
**Authors:** Lauren E. Alfonse, Amanda D. Garrett, Harish Swaminathan, Kelsey C. Peters, Genevieve Wellner, Lau  
Acknowledged Federal Support: **Y**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published  
**Conference Name:** American Academy of Forensic Sciences  
Date Received: 17-Mar-2017 Conference Date: 21-Feb-2017 Date Published: 21-Feb-2017  
Conference Location: New Orleans, LA  
**Paper Title:** Exploring the Relationship between qPCR and STR data for Compromised, Low-Template DNA Samples  
**Authors:** Lauren E. Alfonse, Amanda D. Garrett, and Catherine M. Grgicak  
Acknowledged Federal Support: **Y**

### DISSERTATIONS:

**Publication Type:** Thesis or Dissertation  
**Institution:** Massachusetts Institute of Technology  
Date Received: 04-Aug-2016 Completion Date: 6/2/16 3:59AM  
**Title:** A Mathematical Model of Polymerase Chain Reaction Induced Stutter  
**Authors:** Neil, Gurrām  
Acknowledged Federal Support: **N**

**Publication Type:** Thesis or Dissertation  
**Institution:** Boston University School of Medicine  
Date Received: 04-Aug-2016 Completion Date: 8/4/16 3:59AM  
**Title:** Exploring the Sources of Peak Height Reduction During Low-template, Compromised DNA Data Analysis  
**Authors:** Lauren M. Taranow  
Acknowledged Federal Support: **N**

**RPPR Final Report**  
as of 17-Oct-2017

**Publication Type:** Thesis or Dissertation

**Institution:** Boston University School of Medicine

Date Received: 20-Mar-2017

Completion Date: 3/17/17 4:00AM

**Title:** Characterizing Low Copy DNA Signal Using Simulated and Experimental Data

**Authors:** Kelsey C. Peters

Acknowledged Federal Support: **N**



**Boston University School of Medicine**  
Program in Biomedical Forensic Sciences  
Laboratory for Forensic Research and Technology  
Development

**Final Technical Report**

**NOCIt. A Tool for Determining the  
Number of Contributors: Interpreting  
Complex, Compromised, Low-Template  
Samples**

(Contract: W911NF-14-C-0096)

**Prepared for: Army Research Office**

**For the Period: Sept 8, 2014 to February 7, 2017**  
**Submitted by: Catherine M. Grgicak and Desmond S. Lun**

Approved for public release; distribution is unlimited.

## Contents

Abstract.....	2
List of Figures.....	3
List of Tables.....	6
Preface.....	6
Acknowledgements.....	7
Body Matter.....	7
5.2.1 Summary.....	7
5.2.2 Introduction.....	10
5.2.2 Methods, Assumptions and Procedures.....	16
5.2.3 Results and Discussion.....	27
5.2.3.3 Discussion.....	30
References.....	37

## Abstract

In forensic DNA analysis, the interpretation of a sample acquired from the environment may be dependent upon the assumption on the number of individuals from which the evidence arose. Degraded and damaged samples often exhibit signal that is lower than expected at the high molecular weights, resulting in a ‘ski slope’ effect. In this work we developed a method, and corresponding software tool, that provides a mechanism to determine the number of contributors (NOC) to an unknown stain, regardless of condition. This was accomplished by assessing all known confounding factors related to complex, low-template DNA mixture interpretation. The software, named NOCI, considers signal interferences from reverse stutter products, baseline noise and allele drop-out. It also models the expected allele peak heights or areas at all molecular weights and utilizes this information to provide an APP (*a posteriori* probability) that a certain NOC gave rise to the evidence. We extended the capabilities of NOCI to provide laboratory and algorithmic solutions to the issues associated the interpretation of compromised, complex and low-template DNA samples. We test the new implementation of NOCI on a set of 221 samples containing 15 forensically relevant STR (short tandem repeat) regions of interest and 248 samples containing 21 forensically relevant STR loci. We validate NOCI by examining: 1) accuracy; 2) repeatability; and 3) effects of low-level DNA quantities on the ability to infer the number of contributors to the evidence. We show that the accuracy of NOCI is ~90% across all samples tested. These samples were specifically chosen as they were known to contain minor components > 0.03 ng of DNA. The laboratory conditions used to generate the data were 28 PCR cycles coupled with a 20s/3kV 3130 Genetic Analyzer injection or 29 PCR cycles coupled with a 25s/1.2kV 3500 Genetic Analyzer injection. The sample sets contained non-damaged, damaged and differentially damaged samples. In all cases, as long as there was sufficient signal from the minor contributor, NOCI was able to detect the true NOC 90% of the time. Next, repeatability of these samples was tested by running the two sample sets in triplicate and plotting the differences in the posterior probabilities. We show that NOCI produced repeatable results for

most samples. In 9 of 468 cases, we observed a substantive change in the maximum a posteriori probability (MAP). In these 9 instances the samples contained 4- or 5- contributors.

## List of Figures

Figure 1. An exemplar interpretation pipeline that includes NOCI. All data are analyzed using a peak detection software of choice. The calibration data is garnered from single-source profiles of known genotype analyzed using the lowest possible signal threshold setting. Well-characterized artifacts, such as pull-up and minus A, are filtered manually or with CleanIt and user-defined criteria. These calibration data then parameterize the models utilized by NOCI and CEESIt. NOCI determines the APP (a posteriori probability) distribution on the NOC from data acquired from an unknown sample, which may contain any number of contributors in any proportion. As with the calibration data, the STR data acquired from the environmental sample will undergo pre-processing steps in which peak detection and artefact filtering are completed. Unlike the calibration data, however, a validated high pass signal threshold, also known as an analytical threshold (AT), may be applied. The NOCI output informs the analyst as to the NOC assumption(s). CEESIt, or alternative probabilistic software can be used to compute the LR (likelihood ratio). For this pipeline, data from calibration and unknown samples are expected to be acquired using the the same DNA laboratory processing protocols (i.e., same STR assays, cycle numbers, and electrophoresis settings).

Figure 2. The probability determining the NOC is (■)3; (■)4; and (■)5 for 150,000 simulated true NOC=5 samples with increasing levels of drop-out using MAC. The level of drop-out is indicated by a, b and c. These are the probabilities of drop-out for different contributors within the mixtures. As a, b or c increase, so does the level and rate of NOC misclassification.

Figure 3. Allele stacking and allele dropout. In plots (A), (C), and (D) the data are color-coded by the known number of contributors and separated according to the PCR amplification kit utilized. (A) The expected number of alleles per profile, given the genotypes of the contributors in the mixture (left axis), connected to the number of times signal was detected in allele positions (right axis). (B) An EPG of three representative loci obtained for a five-person mixture containing equal parts from each person: (■) Contributor 1; (■) Contributor 2; (■) Contributor 3; (■) Contributor 4; and (■) Contributor 5. (C) The maximum number of detected alleles at a locus for each sample against the rate of non-detection. The magnitude of the y-axis is not significant as the points were jittered for visualization purposes. (D) The proportion of samples originating from the known number of contributors versus the maximum number of alleles detected at a locus for all samples, which represents the upper bound on the minimum number of contributors that would explain the profile. In no instance are greater than eight alleles at a locus observed, despite the presence of five-person genotype combinations in the database.

Figure 3. The probability of allele drop-out with respect to dilution factor,  $r$ , for a sample containing  $[5.94r]$  copies of DNA at the start of sampling. A dilution factor of 1 denotes no dilution takes place between quantification and amplification.

Figure 4. Schematic depicting the general workflow and a summary of the sample types. (A) Profiles were created using two methods. First, 4,236 DEM (DNA Extract Mixtures) STR profiles were generated by extracting DNA from 69 one-person samples using standard organic

extraction procedures. Absolute DNA quantification was performed using the 7500 Real-Time PCR System and the Quantifiler<sup>®</sup> Duo DNA Quantification Kit according to manufacturer's recommendations and an external calibrator. Second, 20,928 STR profiles garnered from the extraction of whole blood mixtures (WBM) were generated. Developing a process that included the dilution and mixing of whole bloods, rather than DNA extracts, allowed for the acquisition of DNA profiles unaffected by dilutions effects, more closely resembling low-copy environmental samples. To generate these profiles, the quantity of DNA per volume ( $\mu\text{L}$ ) of whole blood was determined for 50 single-source donors via qPCR. These concentrations were utilized to determine the quantity of whole bloods to be diluted, mixed (if necessary), and then extracted. To generate compromised profiles, a subset of samples were enzymatically degraded, sonicated, UV irradiated, or treated with humic acid, a known PCR inhibitor. Samples were extracted using the standard silica-based methodologies. Absolute DNA quantification was performed using the 7500 Real-Time PCR System and the Quantifiler<sup>®</sup> Trio DNA Quantification Kit according to the recommendations of the manufacturer and an external calibrator. For both the DEM and WBM processing schemes, the specified target copy numbers of DNA were amplified using the desired STR multiplex, and samples were then electrophoresed on the 3130 or 3500 Genetic Analyzers. (B) A density plot of the proportion of the minor contributor versus the total copy number, as per qPCR, for two- to five-person mixtures amplified with common human identity STR multiplexes, separated by the total number of PCR cycles utilized. The number of samples and the Genetic Analyzer type are also presented. (C) The percentage of samples processed on the 3130 Genetic Analyzer. (D) The percentage of samples processed on the 3500 Genetic Analyzer.

Figure 5. The probability of allele drop-out with respect to dilution factor,  $r$ , for a sample containing  $\lfloor 5.94r \rfloor$  copies of DNA at the start of sampling. A dilution factor of 1 denotes no dilution takes place between quantification and amplification.

Figure 6. NOCIt "Filter" and "NOCIt" tab settings utilized to test accuracy and repeatability.

Figure 7. Example of a sample progressing through the interpretation pipeline. (A) The GeneMapper<sup>®</sup> ID-X v1.4 electropherogram of nine representative STR loci of a two-person mixture containing equal parts from each contributor wherein no more than four peaks were in obligate allele positions. The total target mass was 0.047 ng. Each obligate allele peak is labelled with the allele designation and peak height above the peak, while the known genotypes for each contributor,  $s_1$  and  $s_2$ , are listed below. The data were exported from the peak detection software and filtered with CleanIt. (B) In the forensic setting, neither the number of contributors nor the contributor genotypes are known a priori. The filtered data are imported into NOCIt, and the a posteriori distribution on the NOC output is shown. NOCIt results suggest that this sample could have arisen from two or three contributors. (C) Both assumptions are used to compute the evidentiary summary statistics in CEESIt for contributors,  $s_1$  and  $s_2$ , as potential donors to the mixture. When the smaller, and in this case correct, NOC is assumed (i.e., two, top two panels), the LR distribution shifts left and the  $\text{LR}_{\text{POI}}$  increases. When a larger number is assumed (i.e., three, bottom two panels), the  $\text{LR}_{\text{POI}}$  decreases for both  $s_1$  and  $s_2$  while the probability that a random person would result in  $\text{LR} > 1$  increases.

Figure 8. NOCIt PDF report, showing file tag and hash values, the log likelihood and the a posteriori probability on the NOC. Also shown is the AT threshold applied to the data.

Figure 9. The CSV batch output file exhibiting most of the information outlined in the pdf output file for all samples run within a batch.

Figure 10. (left panel) The accuracy for 1- to 5-person samples amplified using the Identifiler Plus kit and injected for 20 s at 3kV on a 3130 Genetic Analyzer. (right panel) The accuracy for the same samples as a function of sample type. The term pristine signifies the samples were not subjected to any environmental insults and the signal was of high quality. The compromised category represents all samples that were inhibited, degraded, or UV-Vis damaged. The last category represents differentially degraded samples.

Figure 11. Heat map and accuracies for samples containing various masses of minor components against the true NOC.

Figure 12. (left panel) The accuracy for 1- to 5- person samples amplified using the GlobalFiler kit and injected for 25 s at 1.2kV on a 3500 Genetic Analyzer. (right panel) The accuracy for the same samples as a function of sample type. The term pristine signifies that the samples were not subjected to any environmental insults and the signal was of high quality. The compromised category represents all samples that were inhibited, degraded, or UV-Vis damaged. The last category represents differentially degraded samples.

Figure 13. Heat map and accuracies for samples containing various masses of minor components against the true NOC.

Figure 14. The largest MAP minus the smallest MAP for a set of triplicate runs on NOCIt for the Identifiler Plus test set.

Figure 15. The largest MAP minus the smallest MAP for a set of triplicate runs on NOCIt for the Globalfiler test set.

Figure 16. Accuracy on ca. 1600 samples, wherein the results are classified as accurate if  $APP(NOC_{True}) \geq 1\%$  for NOCIt and MLE. The accuracy results for the MAC method are classified as 'accurate' if the minimum NOC equals the actual NOC. Numbers of samples for each NOC are on top of the bars and the x-axis depicts the true NOC that form the mixture.

Figure 17. Accuracy of NOCIt on the entire dataset as  $APP(True\ NOC) \geq 1\%$ , where samples are gathered by minor contributor DNA mass (in ng) and separated by  $NOC_{True}$  from 1 to 5 (left to right). IP= Samples analyzed with Identifiler Plus kit, GF = samples analyzed with Global Filer kit. The values on the bottom of the bars represent the number of samples in that category.

Figure 18. The degree of sloping observed in the STR profile. The exponent in the decay in fluorescence as a function of molecular weight,  $B$ , ranged from 0.02 to -0.04.  $B$  values significantly below zero correspond to compromised samples and indicate reduction in RFU signal as the length of the amplicons increase. (A) An electropherogram obtained from an untreated sample amplified at 0.25 ng (~40 copies). There is good intra-locus peak height balance across all heterozygous loci, and the total RFU signal is approximately equivalent across

all loci labeled with the same dye, which is represented in the  $B$  value (-0.0005). (B) An electropherogram obtained from a sample treated with 30 sonication cycles amplified at 0.25 ng. The decrease in peak height as the fragment length increases is apparent and characteristic of the “sloping effect” observed in degraded profiles; this is represented by the highly negative  $B$  term obtained (-0.02). (C) The correlation between QI (Quality Index factor, per qPCR) and  $B$ . The treatment protocol, PCR amplification kit, and number of profiles are noted for each plot. The plots are color-coded by density, where yellow and purple represent areas of highest and lowest sample density, respectively.

Figure 19. The accuracy of NOCIt as a function of degradation index,  $B$ , and the actual NOC ( $x$ -axis). The values on top of the figure signify the sample count that fell into that category. Top panel: the samples amplified with Identifiler Plus and run using a 3130 Genetic Analyzer (20s). Bottom panel: all samples amplified with GlobalFiler and run using a 3500 Genetic Analyzer (25s).

Figure 20. Top panel: the number of samples categorized as originating from 1, 2, 3, 4 or 5 contributors according to  $\text{NOCIt}_{\text{MAP}}$ . The colors indicate the percentage of samples - out of 5 NOCIt replicate runs - within that category that resulted in the same NOC results. White coloration demonstrates the results were repeatable for all samples. The  $x$ -axis represents the true NOC. Top panel: Identifiler Plus. Bottom panel: GlobalFiler.

## List of Tables

Table 1. Possible genotype combinations that explain how two random individuals could give rise to the allele signal 13, 16 at the D8S1179 locus.

## Preface

The purpose of this work was to develop a mechanism by which to compute the a posteriori probability (APP) distribution on the number of contributors (NOC) to DNA signal. We accomplished this by, first, generating a large-scale dataset of forensically relevant electropherograms (EPGs). These EPGs generated from well-characterized single-source samples were used to develop the models utilized to compute the posterior probability. These samples were also utilized to parameterize NOCIt models. The mixture sets were utilized to test the accuracy and interpretation capability of the software tool. Through this large empirical dataset, named PROVEDIt, we are able to measure signal detection rates for many laboratory conditions.

NOCIt software development included the expansion and implementation of models that rely on the signal rather than on approximations of the DNA target quantity acquired by qPCR. We implemented an efficient MC (Monte Carlo) based sampling methodology to improve run times and we improved the GUI (Graphical User Interface) by introduction of batched-processing. We ensured the software has been validated for operations by performing a variety of software tests. Last we perform a series of accuracy, repeatability and comparison tests demonstrating the utility of the NOCIt for forensic purposes.

NOCIt v3.1 was developed with contributions by: Slim Karkar, James Kelley, Lauren E. Alfonse, Xia Yearwood-Garcia and Amanda D. Garrett.

## Acknowledgements

We thank Harish Swaminathan whose foundational work formed the basis of this project. We also acknowledge the contributions of Ken Duffy, Muriel Medard, Kelsey Peters, Neil Gurram, Lauren Taranow and Engine Room Technology for their assistance and helpful discussions.

## Body Matter

### 5.2.1 Summary

The interpretation of DNA signal is complex and is encumbered by factors that affect the quality of the signal. Forensic DNA evidence is, typically, processed utilizing the following forensic pipeline: 1) the sample is collected from the environment, submitted for testing and the DNA is extracted; 2) the extract is quantified; 3) a portion of the extract is amplified; 4) the amplified fragments are separated and detected using capillary electrophoresis; and 5) the electropherogram (EPG) data is interpreted. While data from high-copy, low-contributor samples results in EPG signal that is relatively simple to interpret, signal from low-copy, high-contributor samples is difficult to interpret because of allele dropout and allele overlap. Moreover, stutter ratios are elevated in low template samples, making it hard to distinguish between stutter peaks and allelic peaks. In response to the aforementioned issues, a probabilistic procedure and corresponding software tool to infer the number of contributors in a forensic DNA sample was developed. The procedure and software tool is called NOCI.

Forensic EPG interpretation of short tandem repeats (STRs) typically applies the following procedure: 1) an analytical threshold (AT) is applied to distinguish peak from noise; 2) the number of contributors is assessed; 3) the genotypes from the item of evidence are deconvolved; 4) the inferred genotype obtained from the evidentiary profile is compared to the genotype from a known, standard or suspect; and lastly 5) a conclusion and ‘match-statistic’ is reported. The number of contributors (NOC) is an important assumption in the DNA interpretation process since it has a direct impact upon the way in which the evidentiary profile is interpreted. Further, it may be necessary to determine the most likely number of contributors for investigative purposes when a suspect is not available.

The traditional method by which the number of contributors is assessed is by the utilization of a binary method termed the ‘Maximum Allele Count’ (MAC) technique. This method relies upon the application of an AT and stutter filter. All peaks below the AT are considered indistinguishable from noise, while all peaks above the AT are included in the interpretation. Application of an AT to a low-copy profile is not optimal because of the increased risk of labeling a noise peak as an allelic peak or an allelic peak as a noise peak. Furthermore, an allelic peak could be masked by stutter; that is, any peak in the stutter position could be a stutter peak, an allelic peak or a combination of the two. MAC is a widely used method and gives the minimum number of contributors that explain the signal. In short, the maximum number of obligate alleles that exceed AT is divided by two and rounded up; this gives the minimum number of contributors that explain the profile.

Effects of the NOC assumption on the LR have been explored, but the studies have been limited to examining the effects of overestimation on small or simulated sample-sets [1, 2]. A few studies that examine underestimations are available, but the test sets were small [3] or employed semi-continuous systems which do not utilize all of the information contained within

the profile [4]. In all cases, an AT was applied to the data. It is for these reasons that NOCIIt was developed. NOCIIt fits into a full interpretation pipeline, wherein artefactual peaks are filtered, and the LR is computed based on the NOC assessment. Software programs other than CEESIt or CleanIt can be utilized for LR and artifact filtering, respectively.

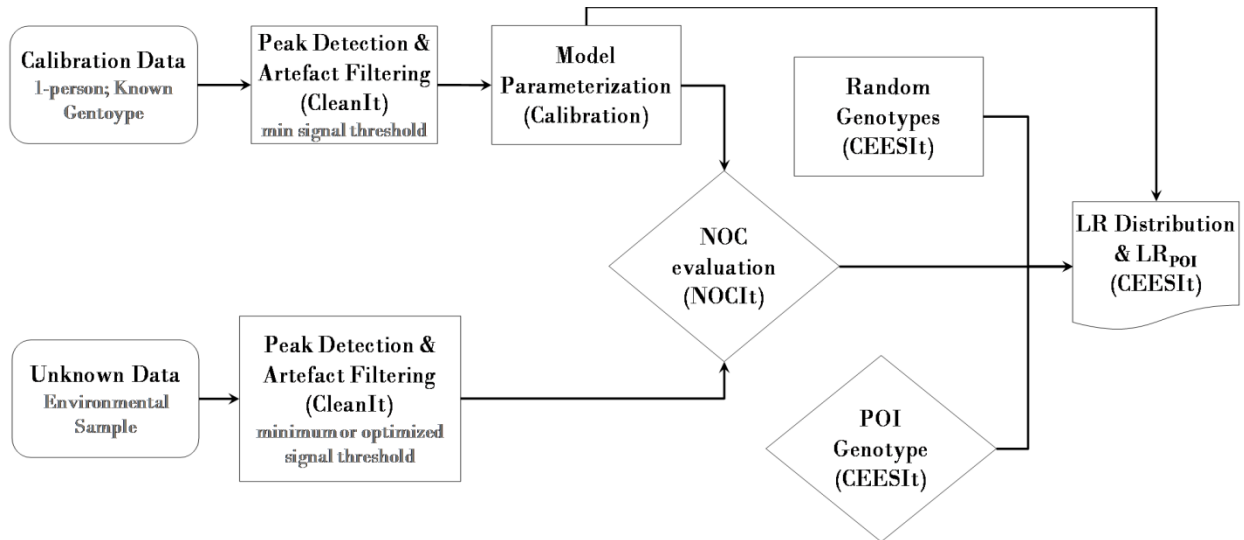


Figure 1. An exemplar interpretation pipeline that includes NOCIIt. All data are analyzed using a peak detection software of choice. The calibration data is garnered from single-source profiles of known genotype analyzed using the lowest possible signal threshold setting. Well-characterized artifacts, such as pull-up and minus A, are filtered manually or with CleanIt and user-defined criteria. These calibration data then parameterize the models utilized by NOCIIt and CEESIt. NOCIIt determines the APP (a posteriori probability) distribution on the NOC from data acquired from an unknown sample, which may contain any number of contributors in any proportion. As with the calibration data, the STR data acquired from the environmental sample will undergo pre-processing steps in which peak detection and artefact filtering are completed. Unlike the calibration data, however, a validated high pass signal threshold, also known as an analytical threshold (AT), may be applied. The NOCIIt output informs the analyst as to the NOC assumption(s). CEESIt, or alternative probabilistic software can be used to compute the LR (likelihood ratio). For this pipeline, data from calibration and unknown samples are expected to be acquired using the the same DNA laboratory processing protocols (i.e., same STR assays, cycle numbers, and electrophoresis settings).

In contrast to the MAC method, NOCIIt works upon the entire signal obtained – the background frequencies of the alleles observed as well as the peak heights - while calculating the number of contributors. To calibrate the software, single source samples from at least 50 donors with known genotypes are required. We generated these samples using the AmpF/str<sup>®</sup> Identifiler<sup>®</sup> Plus and GlobalFiler<sup>®</sup> kit (Life Technologies, Foster City, California). During creation of the sample profiles, 3 injection times (5, 10 and 20s for the 3130 Genetic Analyzer) or 5, 15 and 25s for the 3500 Genetic Analyzer) were used. At each injection time, samples were amplified from at least 7 template DNA amounts (0.008 – 0.25ng). The resultant EPG peaks were separated into one of three categories: True peaks (all peaks representing the alleles in the sample); Stutter peaks (all peaks in the stutter position of True peaks); and Noise peaks (all other peaks in the signal). Both forward and reverse stutter were considered. Calibration parameters (namely the mean and the variance) for the three categories of peaks were computed for each

DNA amount at the three injection times for every locus as a function of peak height. Dropout rates were also computed at each DNA amount at the three injection times for every locus. Models to estimate the rise in baseline noise and allele heights were developed. Both the baseline noise and allele peak heights increased linearly with peak amplitude. The exponential function was used to model the frequency of allele dropout. The stutter peak heights were modeled on a per locus basis as a function of parent peak height. The Gaussian distribution class was used in all cases.

A Monte Carlo approach is used by NOCI to compute the likelihood of the evidence given  $n$  contributors. In every iteration of the Monte Carlo process, genotypes for the  $n$  contributors are chosen based on the frequencies of the alleles in the frequency table. Only alleles present in the frequency table are sampled. A mixture ratio is chosen at random and all mixture ratios are assumed to occur with equal probability. Allele dropout is simulated by Bernoulli trial. Two assumptions are made with regard to dropout: a) Dropout of one allele of a contributor is independent of dropout of the contributor's other allele; and b) Dropout of an allele from a contributor is independent of dropout of the same allele from another contributor. Based on the evidence observed, the likelihood of observing the heights of the peaks given the genotypes of the contributors, the mixture ratio, the amount of DNA amplified and the time of injection is computed using the calibration data. This is repeated numerous times. The average is the likelihood of observing the evidence at a locus, given  $n$  contributors. The  $n$  that results in the highest likelihood is taken to be the number of contributors most supported by the evidence as calculated by NOCI.

The performance of NOCI was tested on 1-, 2-, 3-, 4- and 5- person mixtures for samples amplified with the Identifiler Plus and GlobalFiler Amplification Kits. Like the calibration data, these samples were also generated using three injection times. Samples were amplified using seven DNA target masses (0.08-0.25 ng). The performance of NOCI was assessed by evaluating accuracy and repeatability. In addition we compared the NOCI results to those obtained when allele counting or maximum likelihood estimator (MLE) methods are employed. NOCI does not depend upon an AT and works on the entire electropherogram obtained; thus, no AT was applied to these data when utilizing NOCI. Counting and MLE methods, however, rely upon analytical thresholds; as such an AT of 50 or 100 RFU was used for data generated using the 3130 and 3500 Genetic Analyzers, respectively.

Results demonstrate that NOCI can accurately detect up to five-contributors for samples containing at least 0.03 ng of DNA for each contributor. Comparisons between MAC, MLE and NOCI demonstrate that NOCI outperforms the other methods in both accuracy and the degree to which it underestimates the NOC under all conditions.

NOCI provides an APP distribution on the NOC, providing the analyst with a means to determine the likely NOC and to eliminate unlikely NOC values from consideration.

## 5.2.2 Introduction

A number of state-of-the-art methods, algorithms and software packages that present ‘match-strengths’ [5-12] exist and all have nuances associated with their algorithms [13]. Conclusions are typically presented as a likelihood ratio (LR) which is a statistic that compares the probability of two competing hypotheses. In the forensic DNA context the LR depicts the probability of the evidence given the person-of-interest contributed to the evidence versus the probability of the evidence if they did not contribute. Though the algorithms differ in the ways in which they model stutter [12, 14], allele drop-out [15, 16] and/or baseline noise [17], they all require the user to input an NOC (number of contributors) assumption. Despite the advances in probabilistic tools and DNA mixture interpretation over the last decade, a method to determine the NOC for complex DNA signal has yet to be developed. In this context we define complex DNA signal as EPG signal garnered from DNA extracts containing DNA from more than two contributors or from minor DNA components constituting less than 20% of the mixture. The EPG signal may also be complicated if the DNA is degraded or damaged, or if the PCR was inhibited in some way. Currently, the Maximum Allele Count (MAC) method is the common approach by which to determine the NOC. This is a counting method which counts the alleles that contain peak heights greater than the analytical, or signal, threshold. In order to obtain the minimum NOC using this method, the number of obligate alleles that exceed the AT at locus,  $l$ , is divided by two and rounded up. A number of issues with this approach exist. Mainly, utilization of the NOC obtained from MAC in order to compute the LR is not accurate unless the minimum NOC is the same as the actual NOC. The MAC approach has a tendency to underestimate the NOC. First, MAC does not take into account the frequency of the alleles and the propensity for allelic overlap. Therefore, as the number of actual contributors increases, the probability that the actual NOC equals the minimum NOC decreases. For example, in simulation studies using the SGM+<sup>TM</sup> multiplex loci, it was shown that ~66% of four-person mixtures would present six or fewer alleles at all loci. As a result, these complex mixture samples would likely be incorrectly interpreted as a mixture from three or fewer people [18]. Paoletti et al. report that ~97% of the mixtures had at least one locus with five or six alleles present in the three-person mixtures, and thus, most of the mixtures in this study would be recognized as mixtures of DNA from at least three contributors [19]. Using the information in Tables 2, 3 and 4 from Paoletti et al., one can calculate that 14-15% of the profiles have only a single locus with 5 or 6 alleles. Thus, if the locus loses one or two alleles due to allelic drop-out or if the entire locus is not amplified due to degradation or low-template levels, the mixture would not be recognized as a three-person mixture based on the allele count. Of the profiles with a maximum of 6 alleles per locus, 28-29% have only one locus with 6 alleles. In this case, if one allele is lost from the locus with 6 alleles, the profile still contains one, or more than one locus, with 5 alleles. However, the obvious concern is that if there are multiple loci with allele drop out, the conclusion regarding the number of contributors would be compromised. Similarly Paoletti et al. showed that ~76-77% of the simulated four-person mixtures would not have been recognized as four-person mixtures based on the maximum allele count as they had only five or six alleles (or fewer) at all 13 CODIS loci [19]. Based on the aforementioned studies, it is unlikely that the calculation of the minimum number of contributors for the majority of mixtures containing DNA from three or more contributors will accurately reflect the *true* number of contributors. In fact, the minimum calculation could be off by one, two or more individuals even in the unlikely event that all alleles from each contributor are above the analytical threshold.

Recall, probabilistic-based interpretation systems typically require an assumption on the number of contributors (NOC) to a sample [1, 20]. As a result, work on NOC estimation and studies that examine the effects of the NOC assumption [4, 21-23] on statements of evidential strength have catalyzed the development of methods that manage this limitation [24-27].

To illustrate the effects of allele drop-out on the ability to accurately infer the NOC, consider the allelic peaks 13 and 16 obtained from a typical DNA electropherogram at the D8S1179 locus. We examine the probabilities that one random individual gave rise to the observed peaks in positions 13 and 16. If one person gave rise to the stain, then there is only one way that a single person could have resulted in a D8S1179 signal at allele positions 13 and 16: That person must have a genotype  $G = 13,16$ , and no drop-out could have occurred. Therefore, (population substructure is not considered here for exposition) the probability of the evidence given one contributor is

$$Pr(E|n = 1) = 2f_{13}f_{16}(1 - Pr(DO))^2. \quad (\text{Equation 1})$$

Here,  $n$  is taken to be the NOC,  $f_{13}$  is the frequency of observing the allele 13 within the population,  $f_{16}$  is the frequency of observing allele 16 and  $Pr(DO)$  is the probability of allele drop-out.

Table 1. Possible genotype combinations that explain how two random individuals could give rise to the evidence D8S1179 allele signal 13, 16.

Row	Possible Genotypes Person 1	Possible Genotypes Person 2
1	13,13	13,16 or 16,16
2	13,16	13,16 or 13,13 or 16,16
3	16,16	13,13 or 13,16
4	13,13	16,O
5	13,16	13,O or 16,O
6	16,16	13,O
7	13,O	13,16 or 16,16
8	16,O	13,13 or 13,16
9	13,16	O,O
10	13,O	16,O
11	16,O	13,O
12	O,O	13,16

To calculate the possibility that two unrelated individuals gave rise to signal at positions 13 and 16, three scenarios must be considered: First, the profile could have originated from two random individuals, where neither of the individuals alleles dropped-out (Rows 1-3 Table 1); the second scenario is that two random individuals' DNA gave rise to the peaks, but one allele dropped out (Rows 4-8, Table 1); and the last scenario is that two random individuals gave rise to the signal, but two alleles dropped out (Rows 9-11, Table 2). If drop-out is considered (and the frequency of the allele that dropped out is taken to be 1) then the probability that two random individuals gave rise to the evidence observed at D8S1179 is

$$Pr(E|n=2) = (4f_{13}f_{16}^3 + 6f_{13}^2f_{16}^2 + 4f_{13}^3f_{16}) \cdot (1 - Pr(DO))^4 + 4Pr(DO) \cdot (3f_{13}f_{16}^2 + 3f_{13}^2f_{16}) \cdot (1 - Pr(DO))^3 + 6Pr(DO)^2 \cdot (2f_{13}f_{16}) \cdot (1 - Pr(DO))^2. \quad (\text{Equation 2})$$

Therefore, we can see that the  $Pr(DO)$  has a significant effect on the ability to accurately assess the NOC and neither the MAC nor MLE methods are suitable for low-template samples since they assume all alleles are observed and do not consider all possible genotype combinations that could have

given rise to the evidence (i.e. they exclude the possibilities of rows 4-12, Table 1).

One way to minimize the loss of allele information is to interpret the entire signal. This may be partially accomplished by the application of an AT that is well suited to complex mixture interpretation [28]. Some have suggested that ATs below the commonly applied AT=50 RFU (3130 Genetic Analyzer) may warrant further validation and investigation [29], while others examine the noise signal directly [30, 31]. Figure 2 depicts the impact of ATs on the ability to assess the NOC by allele counting when mixtures increase in complexity and when the probability of drop-out increases. Specifically, Figure 2 depicts the inferred NOC utilizing MAC for a set of 21 autosomal STRs, consistent with the GlobalFiler™ Amplification kit. It demonstrates that as the AT and Pr(DO) increases so does the misclassification rate.

Signal interference from stutter also inhibits accurate interpretation of complex, low-level mixtures, particularly when counting methods are utilized. Current practice dictates that when analyzing single source samples, if the %stutter is below the cutoff value it is assumed that the peak is derived from stutter. If the stutter percent value is above the stutter cutoff value, the peak is considered an allele. This determination becomes error-prone when a mixture of DNAs is present. Using a binary method to classify peaks as stutter can have significant effects on the interpretation of samples with major/minor components. Further, previous work has shown that low-level samples are prone to produce stutters which are more varied and, in general, larger than samples amplified from high-template masses [32]. This is due to: 1) stuttering early in PCR cycling for low-template samples results in large stutter ratios; and 2) noise is not negligible when examining signal acquired from samples in the low-copy regime. When 0.25ng of DNA is amplified, stutter is always present and linearly increases with longest uninterrupted repeat [14, 33]; however at the low-template regime, stutter will not necessarily be detected. This is evidence that a robust model of stutter is needed to correctly interpret stutter/allele likelihoods when attempting low-level DNA interpretation. We utilize NOCI to evaluate the ability to apply robust models to compute the probability on the NOC and determine, through accuracy assessments, if stutter is effectively modeled by NOCI.

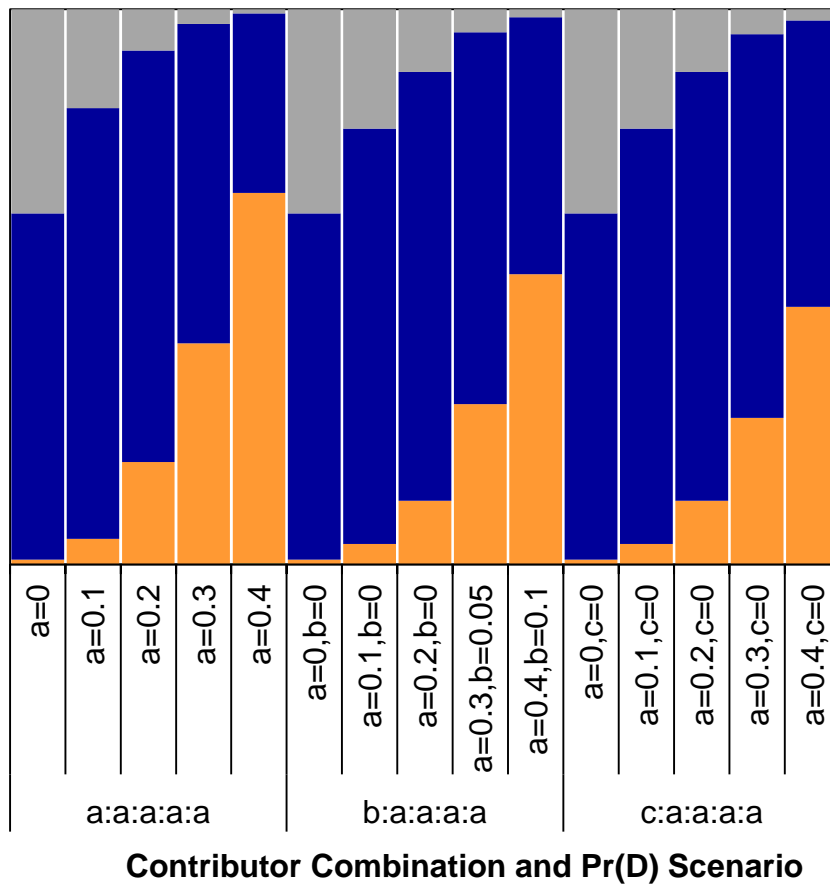


Figure 2. The probability determining the NOC is (■)3; (■)4; and (■)5 for 150,000 simulated true NOC=5 samples with increasing levels of drop-out using MAC. The level of drop-out is indicated by a, b and c. These are the probabilities of drop-out for different contributors within the mixtures. As a, b or c increase, so does the level and rate of NOC misclassification.

Further complicating DNA interpretation is the high likelihood that samples submitted to the forensic laboratory are compromised (i.e. degraded/inhibited). Examples include skeletal remains buried in soil, decomposed bodies, paraffin-embedded tissue or shed telogen hairs. Upon the death of an organism, internal nucleases contained within the cells cause chromosomal DNA to degrade into increasingly smaller fragments over time. The mechanism inducing DNA damage can include strand breakage, formation of pyrimidine dimers, chemical oxidation and hydrolysis. The rate of decay of DNA depends on contamination with microorganisms, temperature, geochemical properties of soil, and effects of the surrounding environment. Highly degraded DNA can induce extensive problems in STR DNA typing, such as poor amplification efficiency, peak imbalance, poor signal to noise ratio, non-specific amplification, and allele drop-out, especially in the large molecular weight range. At worst, DNA can be degraded to such an extent that it is no longer possible to obtain an STR profile. In situations where DNA is highly degraded, poor amplification of the larger sized loci (300 - 500 base pairs) in standard multiplex typing kits is common. To solve this problem, redesigned primer sets were developed in which the primers were positioned as close as possible to the ends of the repeat to reduce the amplified

product size. However, if the level of degradation is severe, the ‘slope-effect’ or ‘decay profile’ can still be observed with mini-STRs.

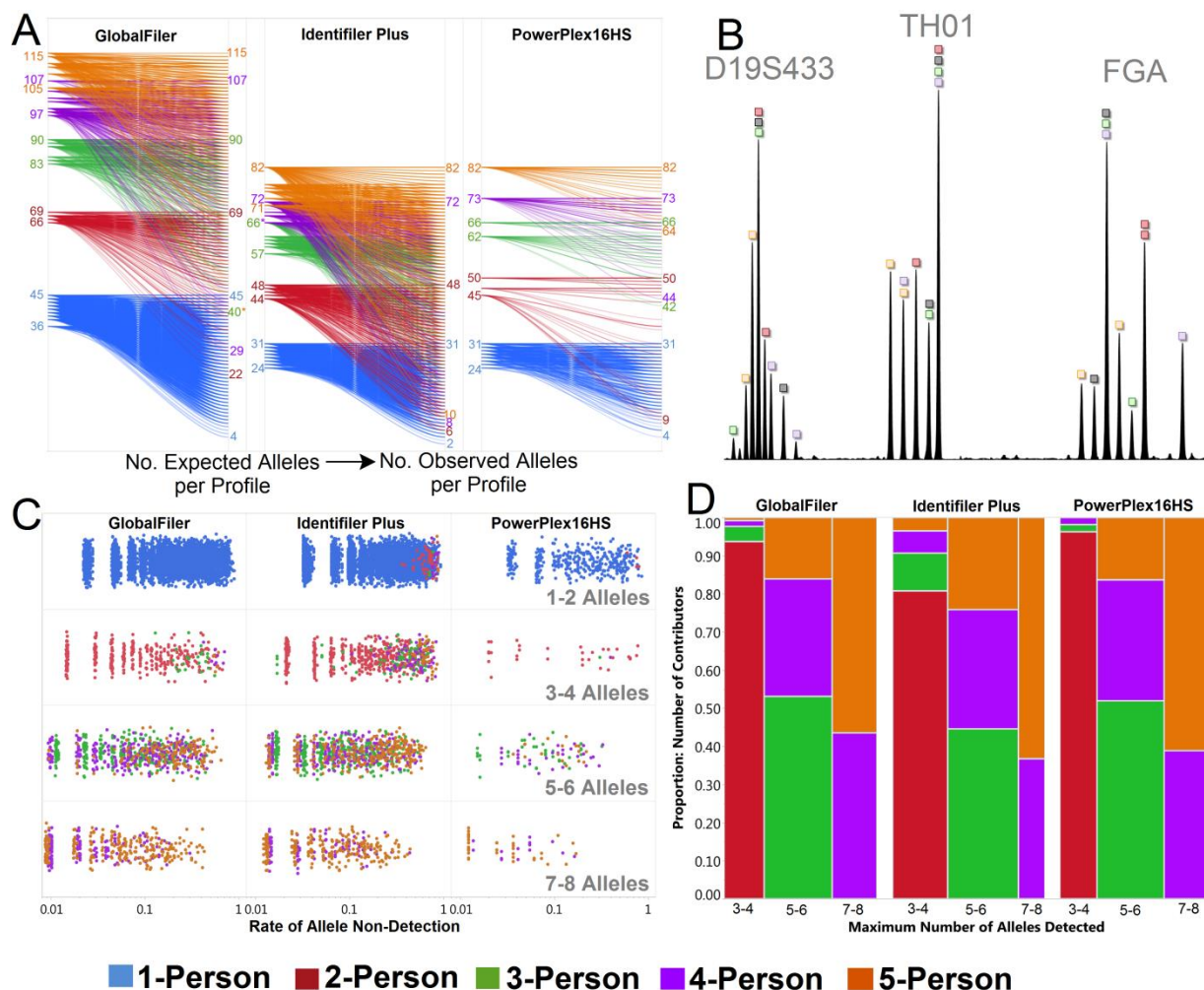


Figure 3. Allele stacking and allele dropout. In plots (A), (C), and (D) the data are color-coded by the known number of contributors and separated according to the PCR amplification kit utilized. (A) The expected number of alleles per profile, given the genotypes of the contributors in the mixture (left axis), connected to the number of times signal was detected in allele positions (right axis). (B) An EPG of three representative loci obtained for a five-person mixture containing equal parts from each person: (■) Contributor 1; (■) Contributor 2; (■) Contributor 3; (■) Contributor 4; and (■) Contributor 5. (C) The maximum number of detected alleles at a locus for each sample against the rate of non-detection. The magnitude of the y-axis is not significant as the points were jittered for visualization purposes. (D) The proportion of samples originating from the known number of contributors versus the maximum number of alleles detected at a locus for all samples, which represents the upper bound on the minimum number of contributors that would explain the profile. In no instance are greater than eight alleles at a locus observed, despite the presence of five-person genotype combinations in the database.

We used the PROVEDIt database (a database of ca. 25,000 profiles) generated over 4-years, using 144 distinct laboratory conditions to supplement studies which demonstrate the

limitation associated with allele counting methods. For each profile the genotypes of the contributors are known. Given the known genotypes of each contributor within the sample, Figure 3A plots the number of expected alleles versus the number of peaks observed in allele positions. Notably, there is large overlap between the number of alleles observed in the three-, four- and five-person mixtures, which suggests that factors associated with allele loss affect the determination of NOC [34, 35] if only the number of detected STR alleles is considered.

An upper bound on the minimum number of contributors can be determined by summing the number of alleles at the locus exhibiting the highest number of detections and dividing the total by two (i.e., upper bound of MAC). Figure 3B represents the EPG signal from three representative loci obtained from a five-person mixture containing equal parts from each contributor. As no more than seven alleles are detected at any locus, the minimum NOC of 4, rather than the actual number, may be inferred and utilized as the NOC assumption [36] during interpretation. Further, manual interpretation of the peak heights for this sample is unlikely to yield consistent results between analysts. This illustrates the deleterious effects of allele overlap, or stacking, on NOC determinations. Allele drop-out is expected to compound the issue. Figure 3C depicts the samples that exhibit a maximum of two, four, six, or eight detected peaks in allele positions at any single STR locus as it relates to the rate of non-detection for the PROVEDIt samples.

As the rate of allele non-detection surpasses 0.1, many three- and four-person mixtures do not exhibit more than four alleles at any one locus. Similarly, at high rates of allele drop-out, some two-person mixtures do not contain loci with greater than two alleles. Profiles that exhibited greater than a maximum of four to five alleles at a locus originated from three-, four- and five-person samples. Similarly, profiles that exhibited a maximum of seven to eight alleles at a locus could have originated from four- or five-person mixtures. The mosaic plot of Figure 3D summarizes the data for the two- to five-person mixture profiles and shows the proportion of samples from the known number of contributors versus the minimum number of contributors that explain the EPG. Strikingly, 53%, 45%, and 52% of samples that exhibit five to six alleles at a locus originated from three-person mixtures for the GlobalFiler<sup>®</sup>, Identifiler<sup>®</sup> Plus, and PowerPlex<sup>®</sup> 16 HS multiplexes, respectively. The remainder originated from either four- or five-person mixtures. Similarly, 44%, 37%, and 39% of samples that exhibit an allele count of seven or eight were from four-person mixtures, with the remainder originating from five-person samples. Despite the variety of five-person genotype combinations represented, no five-person mixture displayed more than eight alleles at any one locus.

Given the complex nature of low-template, compromised, multiple-contributor DNA interpretation, methods to assess NOC based on signal thresholds are not sufficient and are prone to error. In response to these issues, we have developed a non-threshold-based computational approach that determines the likelihood that a certain NOC (number of contributors) gave rise to a biological stain. We highlight that though NOCI has been built to work with all signal, including noise, based on the requirements set for by operational laboratories, it has also been engineered to work with data that has had an AT applied to it. ATs have historically been chosen to minimize the risk of false noise detection. With the introduction of probabilistic models, the likelihood that a peak of height,  $H$ , is noise, allele or stutter can be evaluated. This is a reasonable method for examining allele signal in the presence of noise and ensures all EPG data is analyzed. Further, by examining profiles in this way, it ensures that the  $\text{Pr}(\text{DO})$  is kept to a minimum. In the sections below we use NOCI to evaluate the impact of the AT on the NOC distribution.

## 5.2.2 Methods, Assumptions and Procedures

### 5.2.2.1 NOCI algorithm

NOCI is a software platform that significantly enhances the range of DNA mixture samples that can be analyzed by determining an *a posteriori* probability (APP) distribution on the NOC. NOCI provides quantitative measures and statistical evaluations of evidence and takes into account all known confounding factors related to PCR and instrument interferences. It calculates the *a posteriori* probability (APP) on the number of contributors  $N$  given a particular evidence sample (electrophoresis profile)  $E$ . That is, it calculates  $\Pr(N = n | E)$  for  $n = 1, 2, 3, \dots$ . We assume that *a priori*  $N$  is uniformly distributed between 1 and  $n_{\max}$ , the maximum possible number of contributors. Thus, by Bayes' rule, we obtain

$$\Pr(N = n | E) \propto \Pr(E | N = n), \quad (\text{Equation 3})$$

for  $n = 1, \dots, n_{\max}$ . Let  $\Theta_i$  be the fraction of the total sample contributor  $i$  and the degradation of contributor  $i$ 's DNA relative to that of contributor 1, and let  $\Theta$  be the vector with components  $\Theta_i, i = 1, \dots, n_{\max}$ . We have

$$\Pr(E | N = n) = \sum_{\theta \in \mathcal{T}_n} \Pr(E | N = n, \Theta = \theta) \Pr(\Theta = \theta | N = n) \quad (\text{Equation 4})$$

where  $\mathcal{T}_n$  is the set of all discretized possibilities for  $\Theta$  that are consistent with  $N = n$ . We assume that  $\Theta$  is uniform over  $\mathcal{T}_n$ , so we have

$$\Pr(E | N = n) = \frac{1}{|\mathcal{T}_n|} \sum_{\theta \in \mathcal{T}_n} \Pr(E | N = n, \Theta = \theta). \quad (\text{Equation 5})$$

Now, because the genotype at each locus is independent of the genotype at other loci, we have

$$\Pr(E | N = n, \Theta = \theta) = \prod_{l \in \mathcal{L}} \Pr(E_l | N = n, \Theta = \theta) \quad (\text{Equation 6})$$

where  $\mathcal{L}$  is the set of loci in the kit. Moreover, we have

$$\begin{aligned} \Pr(E_l | N = n, \Theta = \theta) &= \sum_{\mathbf{g}_l \in \mathcal{G}_l} \sum_{\mathbf{d}_l \in \mathcal{D}_l} \Pr(E_l, \mathbf{G}_l = \mathbf{g}_l, \mathbf{D}_l = \mathbf{d}_l | N = n, \Theta = \theta) \\ &= \sum_{\mathbf{g}_l \in \mathcal{G}_l} \sum_{\mathbf{d}_l \in \mathcal{D}_l} \Pr(E_l | N = n, \Theta = \theta, \mathbf{G}_l = \mathbf{g}_l, \mathbf{D}_l = \mathbf{d}_l) \\ &\quad \cdot \Pr(\mathbf{G}_l = \mathbf{g}_l) \Pr(\mathbf{D}_l = \mathbf{d}_l | \theta = \theta) \end{aligned} \quad (\text{Equation 7})$$

where  $\mathbf{G}_l$  and  $\mathbf{D}_l$  are vectors of genotypes and dropouts, respectively, at locus  $l$  for all  $n$  contributors, and  $\mathcal{G}_l$  and  $\mathcal{D}_l$  are sets of all possibilities for  $\mathbf{G}_l$  and  $\mathbf{D}_l$ , respectively.

Because the sets  $\mathcal{G}_l$  and  $\mathcal{D}_l$  become very large as  $n$  increases, we implement NOCI using a Monte-Carlo sampling algorithm. We generate random samples of  $\mathbf{g}_l$  and  $\mathbf{d}_l$  and, for each sample, we compute  $\Pr(E_l|N = n, \boldsymbol{\theta} = \boldsymbol{\theta}, \mathbf{G}_l = \mathbf{g}_l, \mathbf{D}_l = \mathbf{d}_l)$ . After many samples, we average all the computed values, appropriately weighted for the sampling distribution, to obtain an approximation of  $\Pr(E|N = n)$  using Equations 4 to 7. We then calculate the APP according to

$$\Pr(N = n|E) = \frac{\Pr(E|N=n)}{\sum_{n=1}^{n_{\max}} \Pr(E|N=n)}. \quad (\text{Equation 8})$$

To determine  $\Pr(E_l|N = n, \boldsymbol{\theta} = \boldsymbol{\theta}, \mathbf{G}_l = \mathbf{g}_l, \mathbf{D}_l = \mathbf{d}_l)$ , the baseline noise, stutter proportions, drop-out rates and allele heights/areas are all considered and modeled. These models are calculated during calibration using profiles obtained from single-source samples with known genotypes.

#### 5.2.2.2 Experimental Sample Preparation

Fifty whole bloods were diluted by a factor of 10, 100 and 1000 and then extracted using the QiaAmp Investigator Kit using the manufacturer's recommended protocol [37]. The samples were quantified using the Quantifiler<sup>TM</sup> Trio Kit and using a single calibrator as described in [38]. The extract that resulted in the concentration closest to the target mass was the extract utilized for the amplification. The DNA target levels were 0.0078, 0.0156, 0.0313, 0.0625, 0.125 and 0.25 ng. Each sample was amplified twice: once with the Identifiler Plus Amplification Kit and once with the GlobalFiler Amplification kit. Further, an additional sample, which utilized the maximum volume of the extract with the lowest concentration, was created. A total amplification volume of 25  $\mu\text{L}$  was employed during the course of this study. One  $\mu\text{L}$  of amplified work product was mixed with 0.3  $\mu\text{L}$  of LIZ600 v2.0 internal lane standard and 8.7  $\mu\text{L}$  of Hi-Di formamide. The Identifiler Plus work products were injected onto a 3130 Genetic Analyzer using a 5, 10, or 20 s/ 3kV injection protocol. The GlobalFiler products were injected on a 3500 Genetic Analyzer for 5, 15, or 25 s and 1.2 kV. All samples were analyzed at 1RFU with GeneMapper IDx v1.4.

Each of the samples was also subjected to enzymatic degradation, UV-Vis light, humic acid addition or sonication. The sonicated samples were generated by taking 200  $\mu\text{L}$  of the extract and subjecting it to 30 s on-off cycles of sonication for 2, 10 or 30 cycles with a sonicator probe. UV damage was induced by aliquoting whole bloods onto a clean microscope slide and bombarding the whole bloods for 15, 60 or 105 min with UV light. The whole bloods were swabbed, extracted, amplified and run using the aforementioned protocol. Enzymatic degradation was accomplished by adding either DNase I or Fragmentase to extracted DNA. In the case of DNase I degradation, the DNA-free<sup>TM</sup> kit (Invitrogen) was used. Briefly, 0.002, 0.006, 0.012 and 0.024 Units of rDNase I was added to the extract. rDNase I was inactivated by adding 5  $\mu\text{L}$  of DNase Inactivation Reagent and incubating for 2 min at room temperature. The supernatant was removed and stored until further processing. Fragmentase-based enzymatic degradation was accomplished by adding 4  $\mu\text{L}$  of Reaction Buffer v2 and 4  $\mu\text{L}$  incubating Fragmentase<sup>®</sup> (New England Biolabs) to 30  $\mu\text{L}$  of DNA extract obtained from whole bloods. Incubation times were 15, 30 and 45 minutes. Inactivation was accomplished by adding 10  $\mu\text{L}$  of 0.5 M EDTA. Excess EDTA was removed by a re-purification step using the QIAmp Investigator extraction kit. Lastly, the samples were subjected to inhibitory agents. A volume of 15, 22 and 35  $\mu\text{L}$  of 2 mg/mL of Humic Acid was added to each sample. The sample was extracted as previously described. All samples were amplified and analyzed as previously described.

A subset of whole bloods was used to create the mixtures. The mixtures were created by adding appropriate volumes of whole bloods. The required volumes needed to meet the target mixture ratios were approximated by evaluating the amount of DNA obtained per  $\mu\text{L}$  of blood for each individual. If whole-blood dilutions were necessary, then it was the diluted blood that was mixed. The blood mixtures were extracted, quantified, amplified and analyzed as previously described. Figure 4 depicts a density plot, pie chart and general laboratory scheme of the 2-, 3-, 4- and 5- person mixtures.

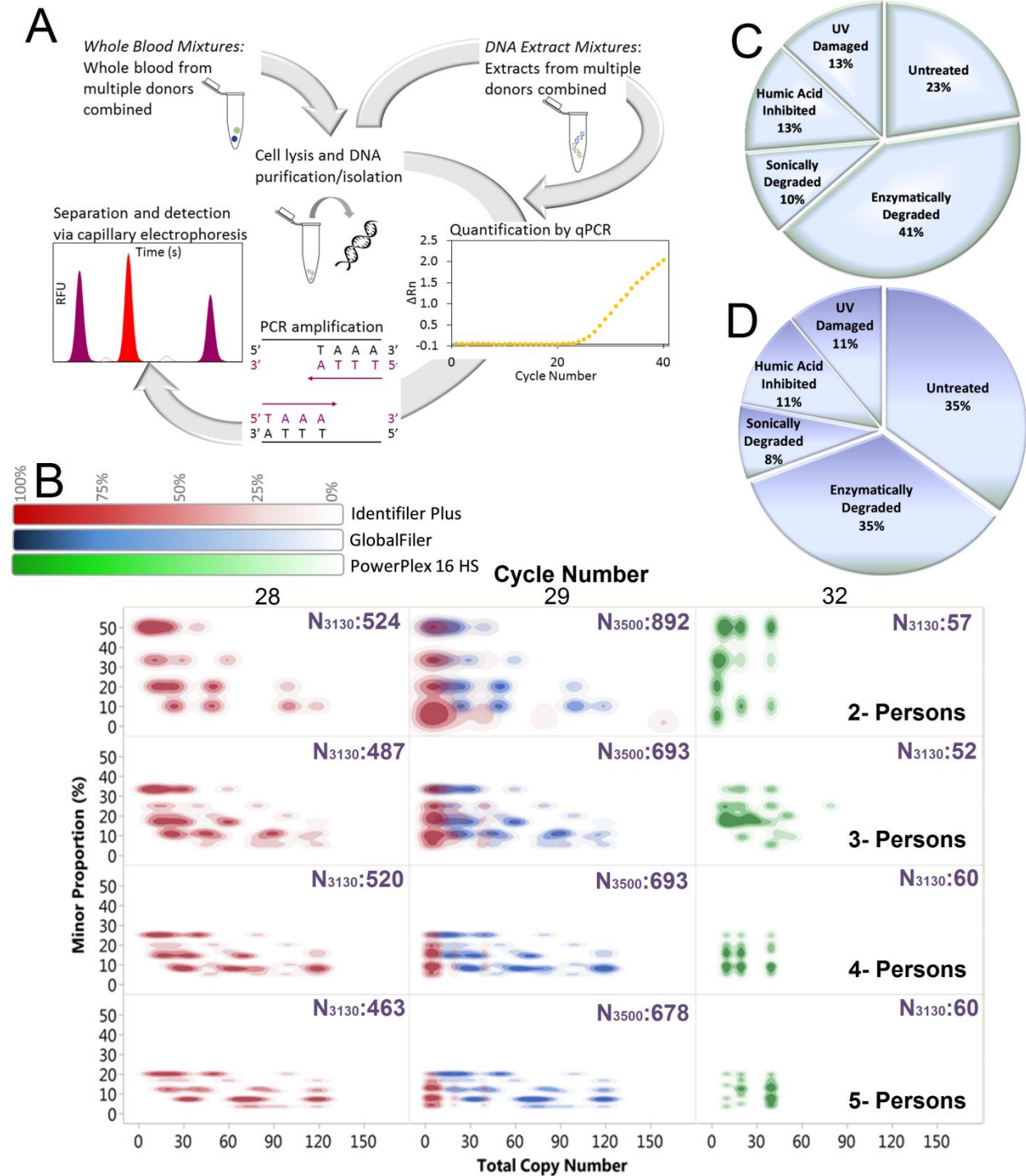


Figure 4. Schematic depicting the general workflow and a summary of the sample types. (A) Profiles were created using two methods. First, 4,236 DEM (DNA Extract Mixtures) STR profiles were generated by extracting DNA from 69 one-person samples using standard organic extraction procedures. Absolute DNA quantification was performed using the 7500 Real-Time PCR System and the Quantifiler<sup>®</sup> Duo DNA Quantification Kit according to manufacturer's recommendations and an external calibrator. Second, 20,928 STR profiles garnered from the extraction of whole blood mixtures (WBM) were generated. Developing a process that included

the dilution and mixing of whole bloods, rather than DNA extracts, allowed for the acquisition of DNA profiles unaffected by dilutions effects, more closely resembling low-copy environmental samples. To generate these profiles, the quantity of DNA per volume ( $\mu\text{L}$ ) of whole blood was determined for 50 single-source donors via qPCR. These concentrations were utilized to determine the quantity of whole bloods to be diluted, mixed (if necessary), and then extracted. To generate compromised profiles, a subset of samples were enzymatically degraded, sonicated, UV irradiated, or treated with humic acid, a known PCR inhibitor. Samples were extracted using the standard silica-based methodologies. Absolute DNA quantification was performed using the 7500 Real-Time PCR System and the Quantifiler<sup>®</sup> Trio DNA Quantification Kit according to the recommendations of the manufacturer and an external calibrator. For both the DEM and WBM processing schemes, the specified target copy numbers of DNA were amplified using the desired STR multiplex, and samples were then electrophoresed on the 3130 or 3500 Genetic Analyzers. (B) A density plot of the proportion of the minor contributor versus the total copy number, as per qPCR, for two- to five-person mixtures amplified with common human identity STR multiplexes, separated by the total number of PCR cycles utilized. The number of samples and the Genetic Analyzer type are also presented. (C) The percentage of samples processed on the 3130 Genetic Analyzer. (D) The percentage of samples processed on the 3500 Genetic Analyzer.

We note that it is the whole bloods, rather than extracts, that were diluted and mixed. To obtain single-source validation samples with low-template levels, it is not uncommon that an extract containing ample copies of DNA may be serially diluted in order to reach low-template levels [16, 39, 40]. Therefore, it was of interest to study the effects of serially diluting concentrated DNA stocks on the EPG and to examine if peak height variability obtained from diluted samples would be representative of the variability seen in low-template samples that do not undergo such processing. To examine this 2000 undiluted and 2000 diluted heterozygous pairs were simulated using the model described in [32]. Briefly, through simulation, we were able to show that the RFU of the less intense peak versus the RFU of the higher RFU peak using heterozygous pairs were different for low-copy extracts and low-copy extracts created through dilution. The tallest heterozygous peak with a sister allele that did not survive the pre-PCR sampling process for undiluted samples was 101 RFU, while the tallest surviving sister peak height for diluted samples was 147 RFU. We further explore the impact of diluting concentrated extracts to reach the low-template regime by examining the heterozygous balance within a locus, defined as,

$$Hb = H_{a1}/H_{a2}, \quad (\text{Equation 9})$$

where  $H_{a1}$  is the peak height of the first allele at D8S1179, and  $H_{a2}$  is the height of the second allele.

Further we explore peak height balance between these two simulations. Peak height balance has been previously studied [14, 41, 42], and it has been suggested that variability in  $Hb$  proportionally decreases with APH as,

$$\text{var}(Hb) = \sigma_{PH}^2 / APH. \quad (\text{Equation 10})$$

In the case of these low-template sample simulations, the APH for the undiluted and diluted low-template samples were 37 and 39 RFU respectively. Using the estimated  $\text{var}(Hb)$  from the simulations in conjunction with Equation 10, it was determined that  $\sigma_{PH}^2$  increased from 37.8 to 43.3 for undiluted and diluted extracts, respectively. Taken together these results suggest a small

but real effect associated with diluting samples containing large copy numbers in order to manufacture samples that contain few copies of DNA.

For well-mixed solutions, sampling due to dilution is accurately described by binomial probabilities, making it particularly amenable to mathematical analysis. For example, mathematically, the outcome of serial dilution is identical to a single dilution with a sampling probability that is the product of sampling volumes. Moreover, the probability of drop-out can be readily computed and so can be considered over a wide range of dilution factors. Striving to obtain a sample containing  $7.8 \times 10^{-4}$  ng/ $\mu$ L of DNA, such that 10  $\mu$ L of this solution will result in approximately 1 copy of DNA to undergo PCR, we can compute the probability that 0 copies survived the pre-PCR sampling process for a given dilution factor. So, in this example  $[(7.8 \times 10^{-4})(48 \times 10^3)/63] = [5.94] = 5$ . Note that 48  $\mu$ L stems from the knowledge that typically 2 of 50  $\mu$ L of the extract is utilized to quantify the DNA with qPCR [43]. Therefore the probability of obtaining 0 copies for a given dilution factor,  $r$ , is

$$P(\text{Binomial}([5.94r], 10/(48r)) = 0) = \left(1 - 10/(48r)\right)^{[5.94r]} \quad (\text{Equation 11})$$

When  $r = 1$ , there is no dilution. As  $r$  becomes larger, the starting density is higher and the sampling probability is proportionally lower, ensuring the same average outcome. The rate of allele drop-out as a function of the dilution factor,  $r$ , is presented in Figure 5, which illustrates that the likelihood an allele survives the pre-PCR sampling process is dependent upon whether or not the extract was diluted prior to amplification. Indeed, Figure 5 demonstrates there is an increasing likelihood of drop-out as  $r$  increases, which converges to a limit. While one can assess this mathematically as described, one of the merits of the model is that this effect can also be seen directly from simulations without recourse to mathematics. When we examine the allele drop-out rate in the simulations, for example, we find that the number of alleles that dropped out increased from 485 (24.25%  $\pm$  0.9% SEM) to 575 (28.75%  $\pm$  1% SEM) for the undiluted and diluted samples, corresponding well to the values determined mathematically.

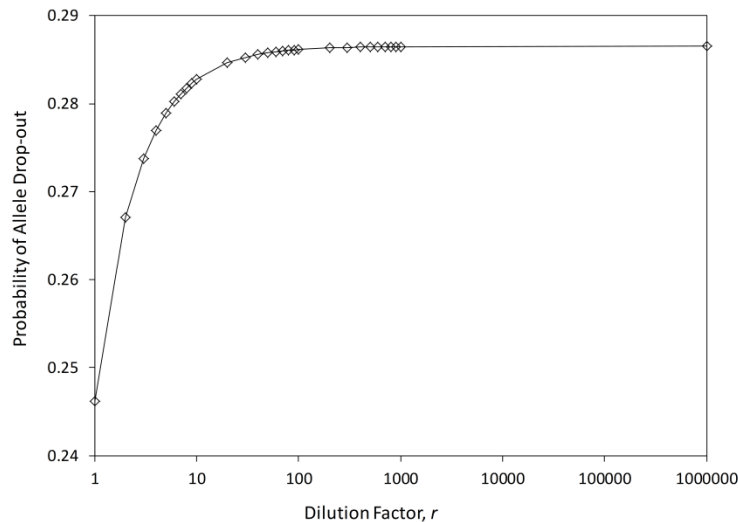


Figure 5. The probability of allele drop-out with respect to dilution factor,  $r$ , for a sample containing  $[5.94r]$  copies of DNA at the start of sampling. A dilution factor of 1 denotes no dilution takes place between quantification and amplification.

From a validation and calibration perspective these results suggest that allele detection rates of low-template DNA samples garnered from diluting extracts with large quantities of DNA do not equal the detection rates obtained of low-template samples that do not undergo dilution. Thus, producing a large validation dataset generated using representative and typical laboratory processes is necessary. In this way, the limitations of a method and, subsequently, the peak height ratio expectations, stochastic effects and single-to-noise of the assay could effectively be established. Appendix A contains additional information regarding the number, type and target recommendations to generate calibration samples.

Differentially degraded samples were created by mixing extracts of that have endured varying levels of degradation. A total of 50 Identifiler Plus and 50 GlobaFiler EPGs of differentially degraded samples were generated.

### 5.2.2.3 NOCI Validation

NOCI Validation consisted of two types of tests: 1) accuracy; and 2) repeatability.

**NOCI Settings:** There are a number of settings that may impact the final result. In particular, NOCI contains a module that allows the analyst to filter pull-up, raised baseline or – A peaks. Further there are set of NOCI run settings that may be modified by the user. Figure 6 depicts the NOCI settings utilized.

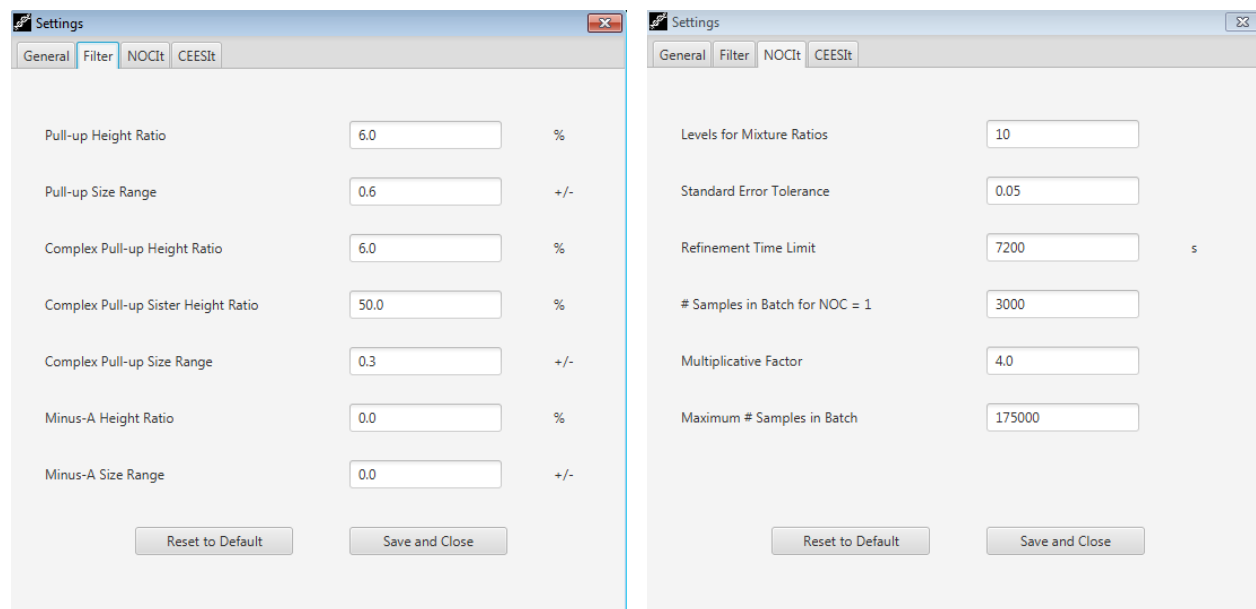


Figure 6. NOCI “Filter” and “NOCI” tab settings utilized to test accuracy and repeatability.

**Accuracy:** Since NOCI outputs the a posteriori probability (APP) distribution on the NOC, any  $APP > 0.01$  was taken to indicate a significant result; thus, if the  $APP(NOC=trueNOC) \geq 0.01$ , then the result was classified as accurate. We note that the probability 0.01 can be modified based on the needs set forth by the laboratory. Further, we assess effects of the accuracy as the minor components decrease.

We note that in order to assess the performance of NOCI, we utilize samples wherein the signal from the minor contributor is from at least 5 cell’s worth of DNA. We explore additional, low-template samples when we compare between methods of ascertaining the NOC, below.

Figure 7 is an example of a sample progressing through the interpretation pipeline depicted in Figure 1. Figure 7B indicates that for this sample, the probable NOC is 2 or 3 with an APP of 0.8 and 0.2, respectively. The maximum a posteriori probability (MAP) of 0.8 suggests that 2-persons explain the mixture, though an APP of 0.2, suggests that 3-persons are also likely to have generated this mixture. Thus, we compute the LR for each suspect ( $s_1$  and  $s_2$ ) using NOC of 2 and 3 for two person-of-interests (POI).

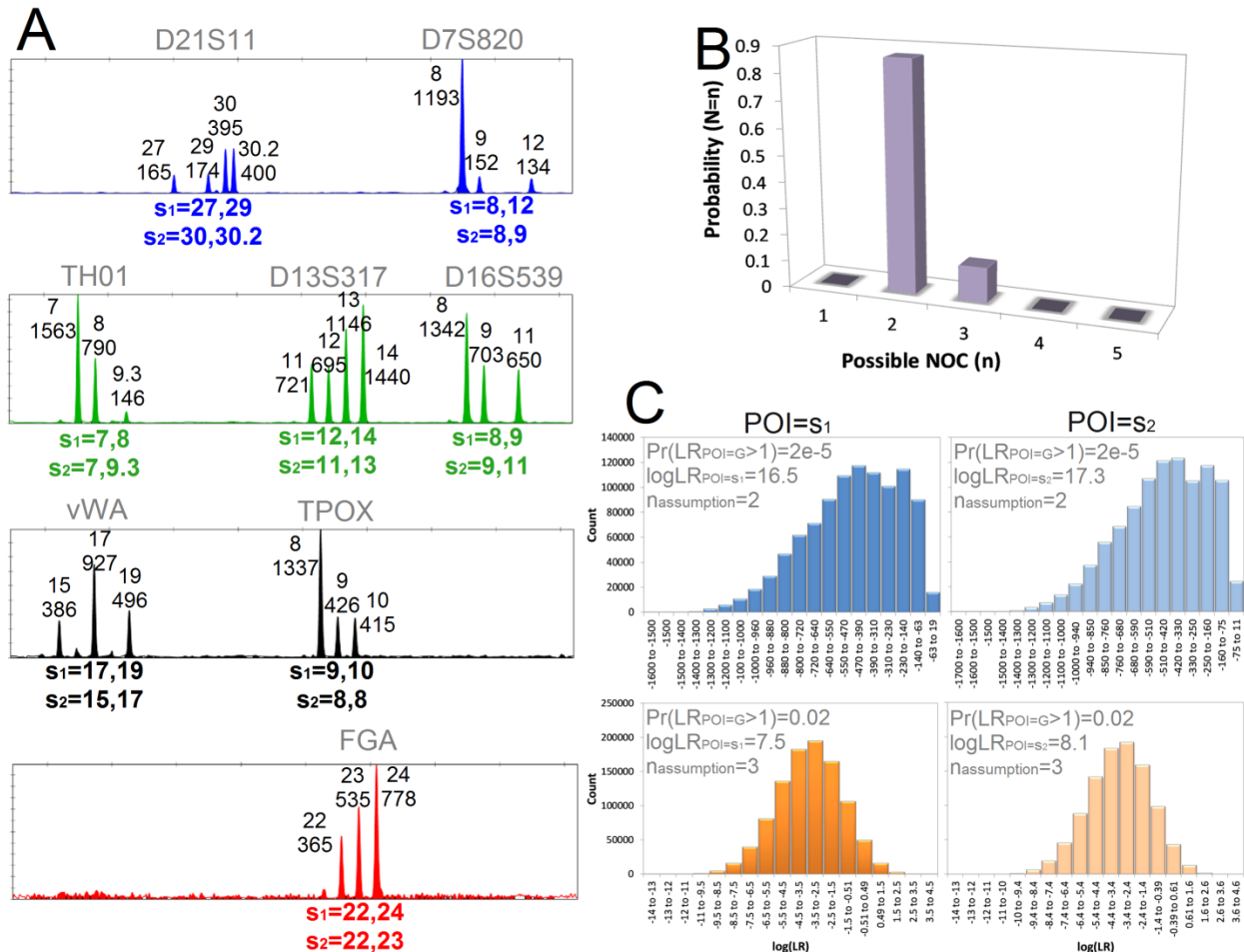


Figure 7. Example of a sample progressing through the interpretation pipeline. (A) The GeneMapper<sup>®</sup> ID-X v1.4 electropherogram of nine representative STR loci of a two-person mixture containing equal parts from each contributor wherein no more than four peaks were in obligate allele positions. The total target mass was 0.047 ng. Each obligate allele peak is labelled with the allele designation and peak height above the peak, while the known genotypes for each contributor,  $s_1$  and  $s_2$ , are listed below. The data were exported from the peak detection software and filtered with CleanIt. (B) In the forensic setting, neither the number of contributors nor the contributor genotypes are known a priori. The filtered data are imported into NOCI, and the a posteriori distribution on the NOC output is shown. NOCI results suggest that this sample could have arisen from two or three contributors. (C) Both assumptions are used to compute the evidentiary summary statistics in CEESIt for contributors,  $s_1$  and  $s_2$ , as potential donors to the mixture. When the smaller, and in this case correct, NOC is assumed (i.e., two, top two panels), the LR distribution shifts left and the  $LR_{POI}$  increases. When a larger number is assumed (i.e.,

three, bottom two panels), the  $LR_{POI}$  decreases for both  $s_1$  and  $s_2$  while the probability that a random person would result in  $LR > 1$  increases.

An example of the full report generated by NOCI is described herein. NOCI Results can be exported in two forms: 1) as a PDF Report, or 2) a CSV summary file.

#### PDF Report

For each *sample* analyzed in NOCI, a PDF report is generated when analysis is complete which provides the computed probability distribution on the number of contributors. In addition to the results, the PDF report summarizes the various input parameters. If one *sample* is analyzed with more than one *Population*, the results of testing with each *Population* will be contained in the same PDF report (i.e., one PDF report is created for each *sample* regardless of the number of *Populations* tested against).

The PDF report summarizes the following elements which pertain to the input parameters:

- The user
- Date and time of the analysis
- A list of the files and settings used during analysis
- Whether the *Filter* function was applied
  - “False” indicates that a *sample* was not filtered; “True” indicates that a *sample* was filtered
- A list of the loci analyzed and the analytical threshold applied at every locus

The PDF report also summarizes the following elements which pertain to the results of analysis:

- The log-likelihood and *a posteriori* probabilities for each number of contributor from 0 to the maximum number of contributors defined by the user (for all *Populations* tested)
- A graphical representation of the *a posteriori* probability versus the number of contributors (for all *Populations* tested)
- A list of errors which may have been encountered during analysis or loci which may have been omitted from the calculation

An example PDF report follows.

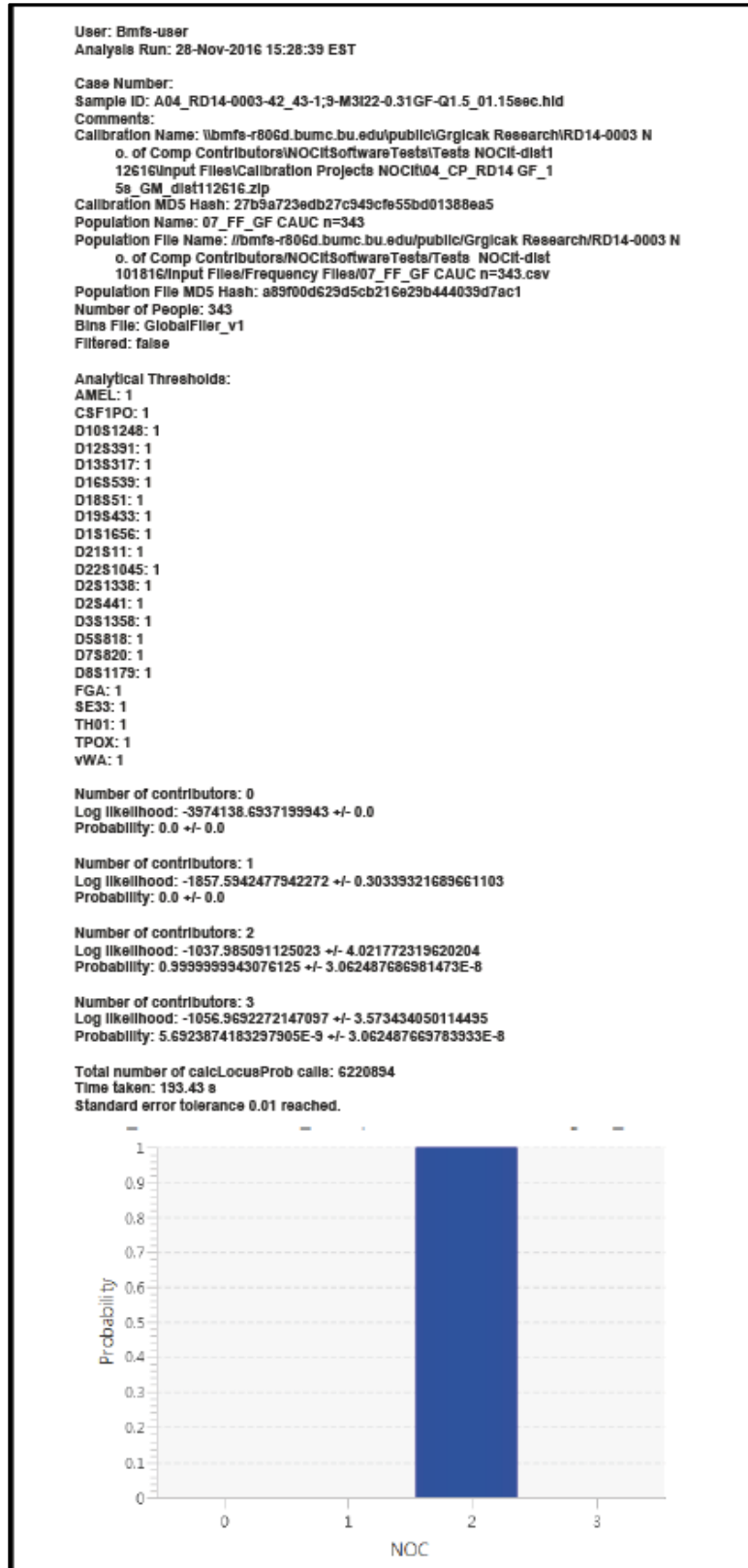


Figure 8. NOCI PDF report, showing file tag and hash values, the log likelihood and the a posteriori probability on the NOC. Also shown is the AT threshold applied to the data.

### Results as CSV

This is an optional output which combines the results of all *Sample Files* analyzed together in one csv file. It also includes all the *General*, *Filter*, and *NOCIt* settings used at the time of analysis. This file is not automatically produced but requires the user to select the *Include Results as CSV* option when adding a batch. Selecting this option will not affect or override the PDF reports which are automatically created for all *samples*.

An example *Results as CSV* file is shown below. Note that this file contains a significant amount of data organized into many columns. Although the probability results are not visible in the image below, they are present in columns to the far right of the spreadsheet.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Output File	Sample ID	Case Num	Comment	Calibratio	Frequenc	Populatio	Number o	Bins File	Filtered	Pull-up Fil	Pull-up Fil	Complex I
2	C:\Users\i	H01-RD12-0002-68-0	Test		\\10.224.1	//10.224.1	IP Caucasi	349	AmpFLSTF	TRUE	5	1	6
3	C:\Users\i	H01-RD12-0002-68-0	Test		\\10.224.1	//10.224.1	IP Caucasi	349	AmpFLSTF	TRUE	5	1	6
4	C:\Users\i	H01-RD12-0002-68-0	Test		\\10.224.1	//10.224.1	IP Caucasi	349	AmpFLSTF	TRUE	5	1	6
5	C:\Users\i	H01-RD12-0002-68-0	Test		\\10.224.1	//10.224.1	IP Caucasi	349	AmpFLSTF	TRUE	5	1	6
6	C:\Users\i	H01-RD12-0002-68-0	Test		\\10.224.1	//10.224.1	IP Caucasi	349	AmpFLSTF	TRUE	5	1	6
7	C:\Users\i	H01-RD12-0002-68-0	Test		\\10.224.1	//10.224.1	IP Caucasi	349	AmpFLSTF	TRUE	5	1	6
8	C:\Users\i	H01-RD12-0002-68-0	Test		\\10.224.1	//10.224.1	IP Caucasi	349	AmpFLSTF	TRUE	5	1	6
9	C:\Users\i	H01-RD12-0002-68-0	Test		\\10.224.1	//10.224.1	IP Caucasi	349	AmpFLSTF	TRUE	5	1	6
10	C:\Users\i	H01-RD12-0002-68-0	Test		\\10.224.1	//10.224.1	IP Caucasi	349	AmpFLSTF	TRUE	5	1	6
11	C:\Users\i	H01-RD12-0002-68-0	Test		\\10.224.1	//10.224.1	IP Caucasi	349	AmpFLSTF	TRUE	5	1	6
12	C:\Users\i	H01-RD12-0002-68-0	Test		\\10.224.1	//10.224.1	IP Caucasi	349	AmpFLSTF	TRUE	5	1	6
13	C:\Users\i	H01-RD12-0002-68-0	Test		\\10.224.1	//10.224.1	IP Caucasi	349	AmpFLSTF	TRUE	5	1	6
14	C:\Users\i	H01-RD12-0002-68-0	Test		\\10.224.1	//10.224.1	IP Caucasi	349	AmpFLSTF	TRUE	5	1	6
15	C:\Users\i	H01-RD12-0002-68-0	Test		\\10.224.1	//10.224.1	IP Caucasi	349	AmpFLSTF	TRUE	5	1	6
16	C:\Users\i	H01-RD12-0002-68-0	Test		\\10.224.1	//10.224.1	IP Caucasi	349	AmpFLSTF	TRUE	5	1	6
17	C:\Users\i	H01-RD12-0002-68-0	Test		\\10.224.1	//10.224.1	IP Caucasi	349	AmpFLSTF	TRUE	5	1	6
18	C:\Users\i	H01-RD12-0002-68-0.0156	IP-00		\\10.224.1	//10.224.1	IP African	357	AmpFLSTF	TRUE	5	1	6

Figure 9. The CSV batch output file exhibiting most of the information outlined in the pdf output file for all samples run within a batch.

**Repeatability:** Repeatability was assessed by inspecting the MAP from a set of three runs. The largest MAP from the set of three was compared to the range of MAP obtained.

All tests were completed for the Identifiler Plus (20 s injection) and GlobalFiler (25 sec injection) datasets where the minor contributor did not constitute less than 0.03 ng of the mixture.

**Comparison:** We compared NOCI results against the MAC and Maximum Likelihood Estimator (MLE) methods [44]. We evaluate overall accuracy for over 1,600 PROVEDIt samples; ca. 800 samples were amplified with Identifiler Plus for 28 cycles and injected for 20 sec on a 3130 Genetic Analyzer while the other ca. 800 samples were amplified with Globalfiler for 29 cycles and injected on a 3500 for 25 sec. We explore the accuracy of each method as it relates to 1%APP and MAP (maximum a posteriori probability), and as it relates to template mass, degradation levels, and the levels to which the samples were over- and under-estimated; that is, we evaluate the accuracy when we only take the maximum a posteriori (MAP) as the inferred NOC and compare it to the true NOC. In addition, we evaluate the number of times NOCI suggested there was a 1% probability that the true NOC was inferred. For example, Figure 7B shows that the sample in Figure 7A resulted in an APP of ca 0.8 and 0.2 for NOC of 2 and 3, respectively. Since the  $NOC_{true}$  was 2 for this sample, we would count the MAP NOCI result as accurate (i.e., the MAP was found at  $NOC=2$ ). In addition we would have classified the

1%APP as correct because the APP at NOC=2 was greater than 1%. We use the same MAP and 1%APP categorizations for the MLE method. MAC has no probability associated with it; thus, there is no APP1% result for MAC.

### 5.2.3 Results and Discussion

**Accuracy:** Figure 10. Depicts the accuracy (%) of NOCIt<sub>APP≥1%</sub> for 221 Identifiler Plus and 248 GlobalFiler 1- to 5- person samples injected on a 3130 Genetic Analyzer for 20 s at 3kV and 3500 Genetic Analyzer for 25s at 1.2 kV, respectively. In all cases the smallest minor component constituted at least 0.03 ng of the mixture. We restricted the minor component in order to evaluate the performance of NOCIt rather than the limits associated with template masses of the minor contributors. The full evaluation of NOCIt across all samples is in the section entitled *Comparison*, below. The samples were analyzed at 1 RFU using GeneMapper IDx v1.4.

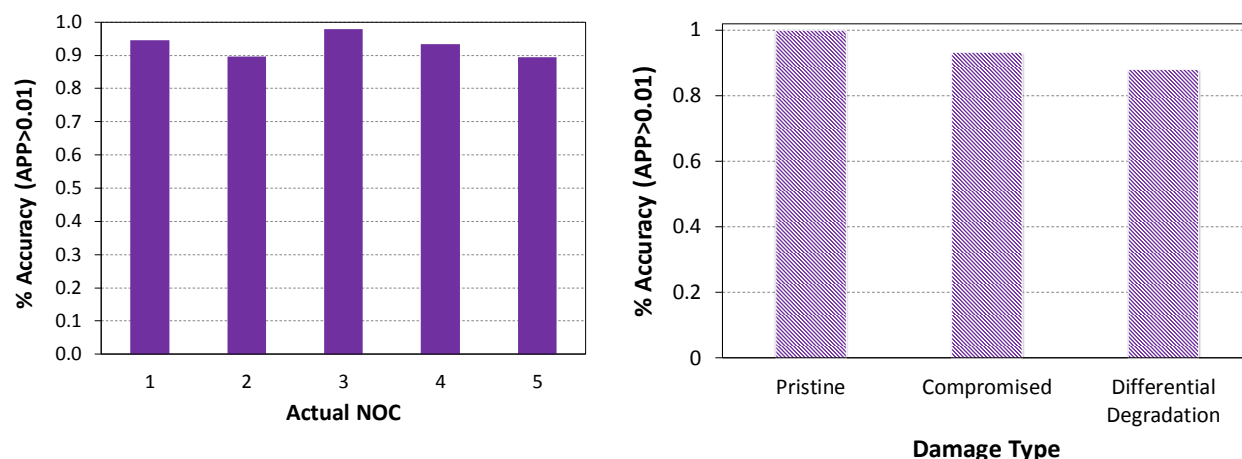


Figure 10. (left panel) The accuracy for 1- to 5-person samples amplified using the Identifiler Plus kit and injected for 20 s at 3kV on a 3130 Genetic Analyzer. (right panel) The accuracy for the same samples as a function of sample type. The term pristine signifies the samples were not subjected to any environmental insults and the signal was of high quality. The compromised category represents all samples that were inhibited, degraded, or UV-Vis damaged. The last category represents differentially degraded samples.

Results indicate that if reasonable levels of signal are detected for each contributor NOCIt was able to accurately detect the true NOC. Specifically, of the 221 Identifiler Plus test samples, 205 resulted in APPs larger than 0.01 for the true NOC. We note that the batch of 221 test samples included samples containing minor/major mixtures, degraded samples as well as differentially degraded samples. If we examine the results as a function of the NOC, we see that the accuracy is not perturbed by increasing the true NOC to the sample. If the sample set is segregated into blocks of non-degraded; degraded; and differentially degraded sample types we observe that 17 of the 17 pristine samples resulted in accurate results. The accuracy decreased to 93% for the 162 compromised samples. The lowest accuracy, i.e., 88%, was obtained for 42 samples that were differentially degraded. Further, we examine the effect of the mass of the minor contributor; Figure 11 suggests that the accuracy remains stable across the actual NOC for samples whose minor contributors constitute > 10 cell's worth of DNA. The accuracies decrease to 74% when the mass of the minor is between 5 and 10 cell's worth of DNA and signal from four other individuals are present in the EPG.

MassMinor (ng) /True NOC	1	2	3	4	5
0.03 to 0.063	100.0%	94.1%	94.4%	89.5%	73.7%
0.063 to 0.124	100.0%	94.1%	95.7%	95.2%	100.0%
0.125 to 0.25	100.0%	84.6%	100.0%	80.0%	100.0%

Figure 11. Heat map and accuracies for samples containing various masses of minor components against the true NOC.

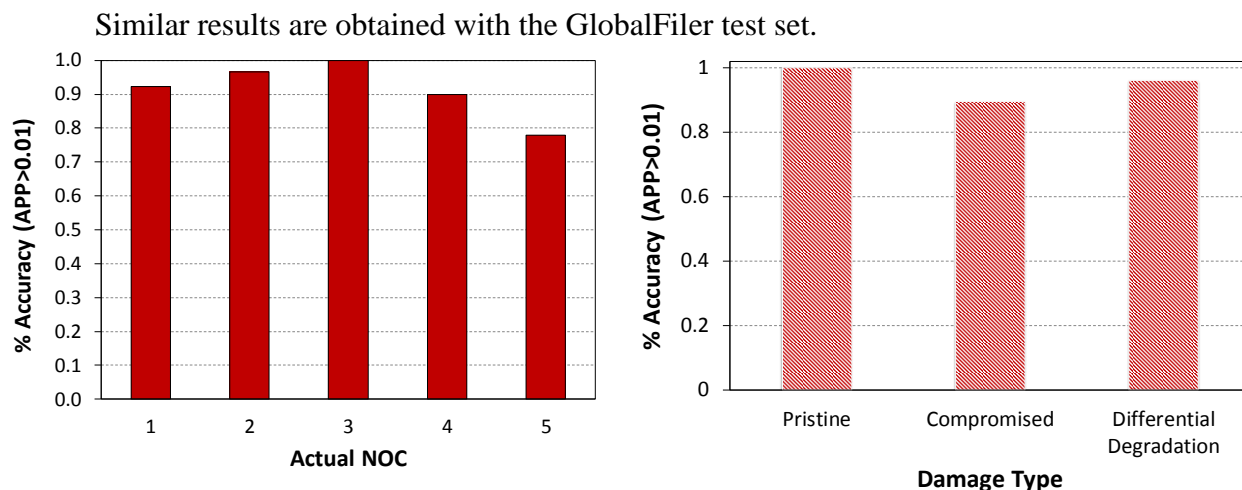


Figure 12. (left panel) The accuracy for 1- to 5- person samples amplified using the GlobalFiler kit and injected for 25 s at 1.2kV on a 3500 Genetic Analyzer. (right panel) The accuracy for the same samples as a function of sample type. The term pristine signifies that the samples were not subjected to any environmental insults and the signal was of high quality. The compromised category represents all samples that were inhibited, degraded, or UV-Vis damaged. The last category represents differentially degraded samples.

Figure 12 shows the accuracy results obtained for the GlobalFiler test set and indicate that the accuracy does not substantially improve when more loci are utilized to evaluate the NOC to a mixture. We note that the SE33 complex stutter was not filtered from this dataset; studies which explore the impact of SE33 complex stutters on the NOC assessment are warranted. Figure 12 also suggests that pristine, degraded and differentially degraded have similar accuracies for this dataset. This demonstrates the utility of degradation models and sampling methods for degraded and differentially degraded samples. In both test sets, the pristine samples resulted in 100% accuracies. This is an indication that the information content and quality of the sample has the largest impact on accurately determining the NOC. Further, Figure 13 depicts the heat map for the GlobalFiler dataset. These results are consistent with those obtained with the Identifiler Plus dataset where the lowest accuracies were obtained when the mass of the minor contributor < 0.063 ng and in the presence of > 4 contributors.

Mass of Minor (ng) / TrueNOC	1	2	3	4	5
0.03 to 0.063	100.0%	100.0%	100.0%	86.4%	70.4%
0.063 to 0.124	75.0%	100.0%	100.0%	95.7%	89.5%
0.125 to 0.25	93.8%	87.5%	100.0%	80.0%	100.0%

Figure 13. Heat map and accuracies for samples containing various masses of minor components against the true NOC.

These results suggest that NOCI’s peak height and degradation based models are effective at detecting the true NOC when as few as 5-cells worth of DNA is present in the sample. Interestingly, the kit type has seemingly little effect on the ability to infer the NOC that comprises the sample.

**Repeatability:** Figure 14 depicts the range of MAP (log-scale) versus the largest MAP obtained when the same sample was run three times on NOCI. The majority of repeatability tests resulted in ranges that were <0.1; however, in two of the 221 Identifiler Plus cases, the range exceeded 0.5 (both were 5- person mixtures). Specifically, for a five-person mixture the MAP of run1 was 0.986, while the MAP of run3 was 0.395. The second sample resulted in a MAP of 0.989 and 0.018 for two consecutive runs. Similar results are depicted for triplicate runs for the GlobalFiler test set. Notably, the GlobalFiler test set resulted in only one sample with the largest MAP < 0.9. Of the 248 triplicate runs, 7 of 248 resulted in MAP ranges exceeding 0.5; these results were obtained when the mixtures contained 4- or 5-contributors. In summary, 206 (out of 221) Identifiler Plus and 236 (out of 248) GlobalFiler runs resulted in MAP ranges that were smaller than 0.1.

Repeatability is a function of NOCI parameter settings. These results were obtained with the parameters depicted in Figure 6.

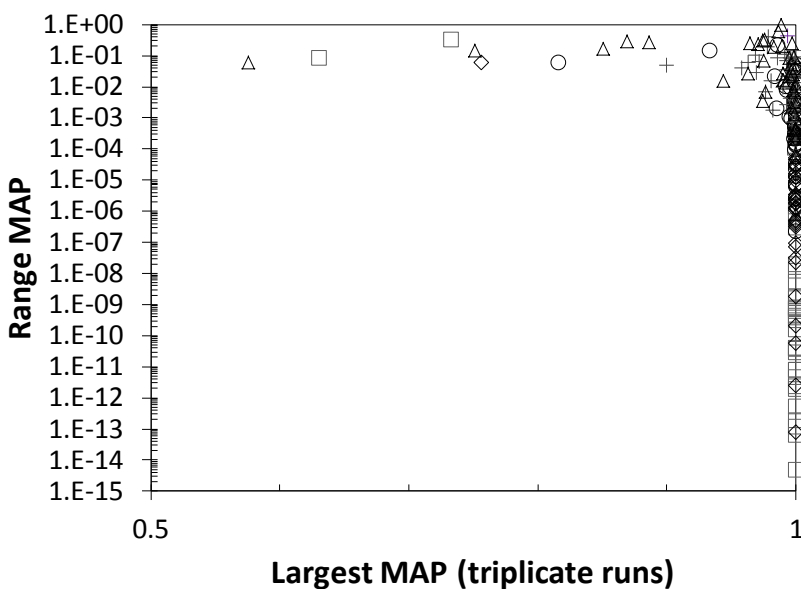


Figure 14. The largest MAP minus the smallest MAP for a set of triplicate runs on NOCI for the Identifiler Plus test set.

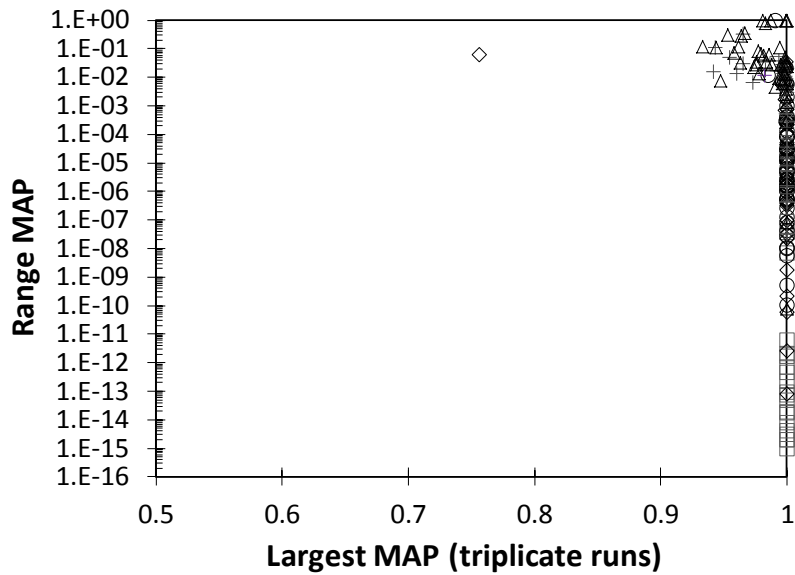


Figure 15. The largest MAP minus the smallest MAP for a set of triplicate runs on NOCI for the Globalfiler test set.

**Comparison:** We tested NOCI performance against the MAC and MLE [44] methods. First, we explore the performance of each method as a function of the minor component, in ng, to the 1- to 5-person mixtures. Both MAC and MLE require an AT >1 RFU; as such, we utilized an AT of 50 and 100 RFU for the samples run with Identifiler/3130 Genetic Analyzer and GlobalFiler/3500 Genetic Analyzer combinations, respectively.

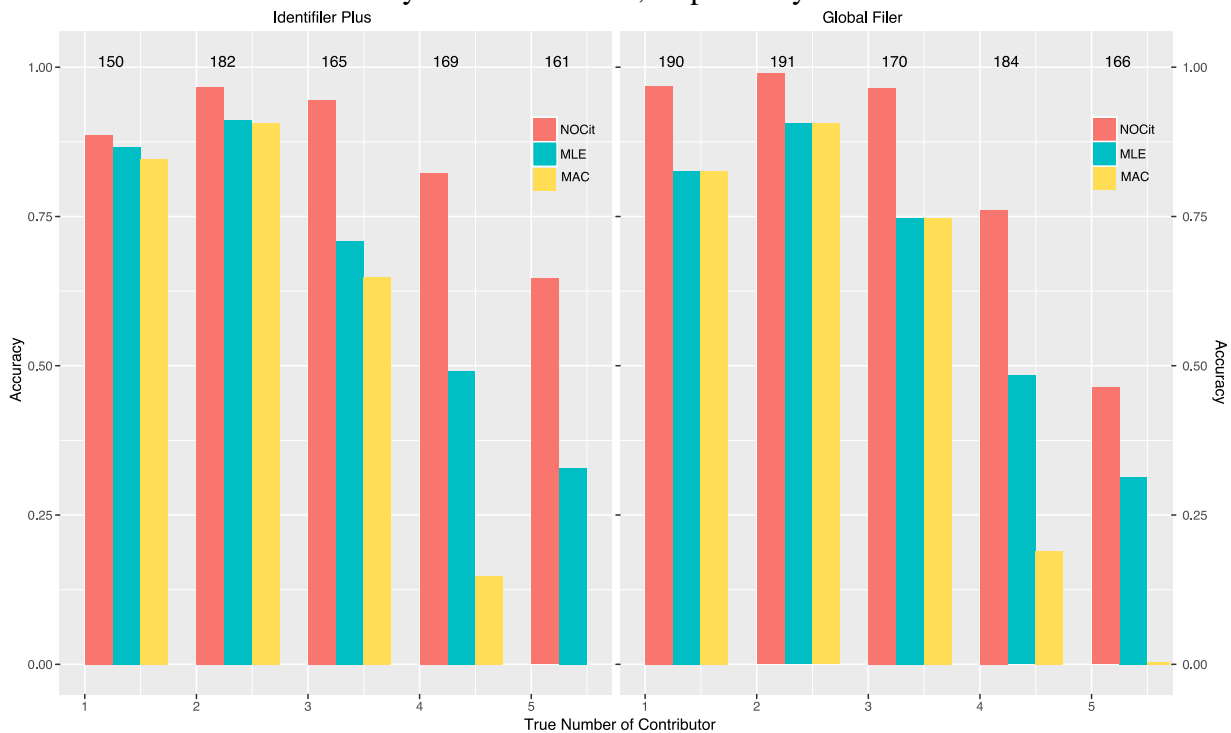


Figure 16. Accuracy on ca. 1,600 samples, wherein the results are classified as accurate if  $APP(NOCI_{True}) \geq 1\%$  for NOCI and MLE. The accuracy results for the MAC method are

classified as ‘accurate’ if the minimum NOC equals the actual NOC. Numbers of samples for each NOC are on top of the bars and the x-axis depicts the true NOC that form the mixture.

When we examine the performance of all methods across the wide variety of samples within the PROVEDIt dataset (i.e., these results are not conditioned on the minor contributor containing greater than 0.03 ng of DNA), we observe that NOCIt outperforms both methods across all  $NOC_{True}$  values. As the  $NOC_{True}$  increases the accuracy of all methods decreases (Figure 16).

Next, we examine the NOCIt results as a function of template mass of the minor (Figure 17).

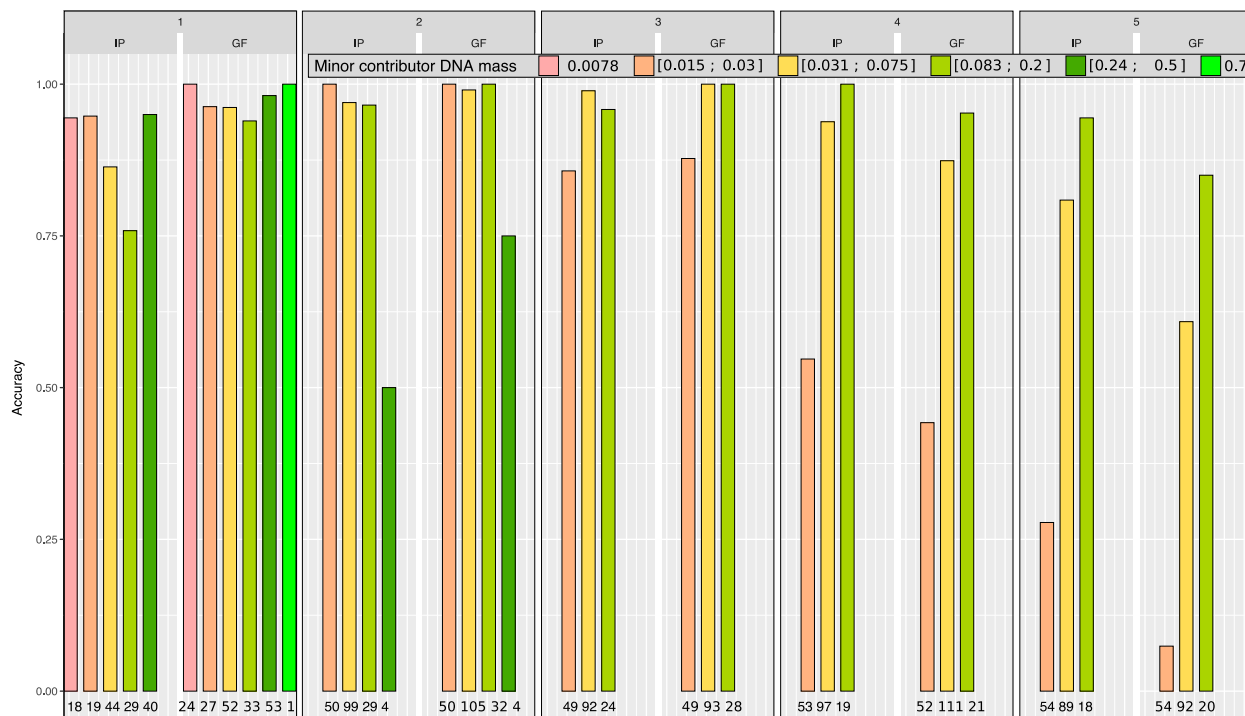


Figure 17. Accuracy of NOCIt on the entire dataset as  $APP(True\ NOC) \geq 1\%$ , where samples are gathered by minor contributor DNA mass (in ng) and separated by  $NOC_{True}$  from 1 to 5 (left to right). IP= Samples analyzed with Identifiler Plus kit, GF = samples analyzed with Global Filer kit. The values on the bottom of the bars represent the number of samples in that category.

Figure 17 demonstrates that, as expected, NOCIt accuracy is affected by the mass of the minor contributor for high-order mixtures containing four or greater contributors. In particular, when the mass of the minor contributor represents less than 0.03 ng of DNA (i.e., ca. 5 cells’ worth) the accuracy substantially decreases. When the EPG contains signal from 3 or fewer contributors, NOCIt retains accuracy rates that exceed 85%, regardless of quantity.

In a similar vein, we examined the outputs of each of the methods with respect to the EPG sloping value. The degree of degradation or inhibition can be assessed using the contour of the STR signal, which was modeled as exponential decay in fluorescence as a function of molecular weight:

$$H_l = Ae^{B\bar{s}_l} \quad (\text{Equation 12})$$

where  $H_l$  is the weighted peak heights of locus  $l$ ,  $\bar{s}$  is the weighted average base pair size of the STR alleles at locus  $l$ , and  $A$  and  $B$  are the exponential parameters. In extreme cases of decay, the highest molecular weight peaks may not reach detectable levels. If high molecular weight markers exhibit low peak heights due to degradation or inhibition of the PCR reaction,  $B$  will take a large negative value. In contrast, if there is good signal balance across all loci, indicating efficient PCR and high quality template DNA,  $B$  will be near zero. As illustrations of the different conditions of DNA and  $B$  values that may be encountered, Figures 18A and 18B depict EPGs obtained from single-source samples that were untreated and severely damaged through sonication, respectively. Figure 18C presents the QI value versus  $B$  calculated for each sample, separated by the amplification chemistry and the treatment protocol utilized. As expected, all untreated samples resulted in QI values near one and  $B$  parameters near zero. Samples that were subjected to conditions that degraded the DNA show that the QI and  $B$  parameter are correlated, suggesting that the QI metric can be used to predict the STR sloping pattern for these sample types. Interestingly, samples subjected to conditions intended to induce PCR inhibition show that while the qPCR QI metric demonstrates that PCR inhibitors affect the amplification of the large autosomal fragment, the STR profile presents only minor signs of inhibition. Spearman's  $\rho$  for QI and  $B$  was computed, with correlation being strong for enzymatically degraded, sonicated, and UV-damaged samples ( $\rho$  between  $-0.8656$  and  $-0.6324$ ), but less so for untreated and inhibited samples ( $\rho$  between  $-0.1382$  and  $-0.0143$ ).

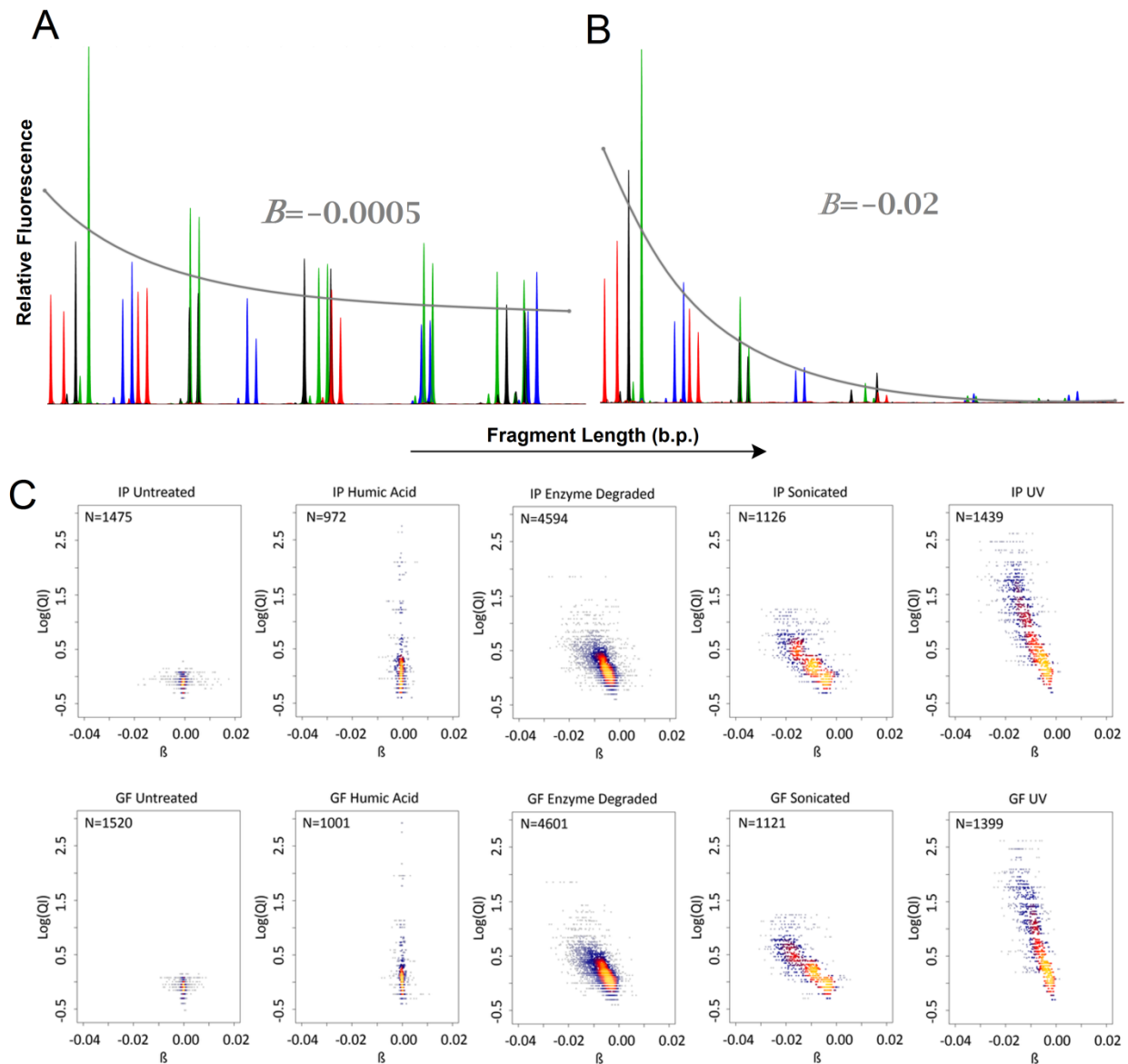
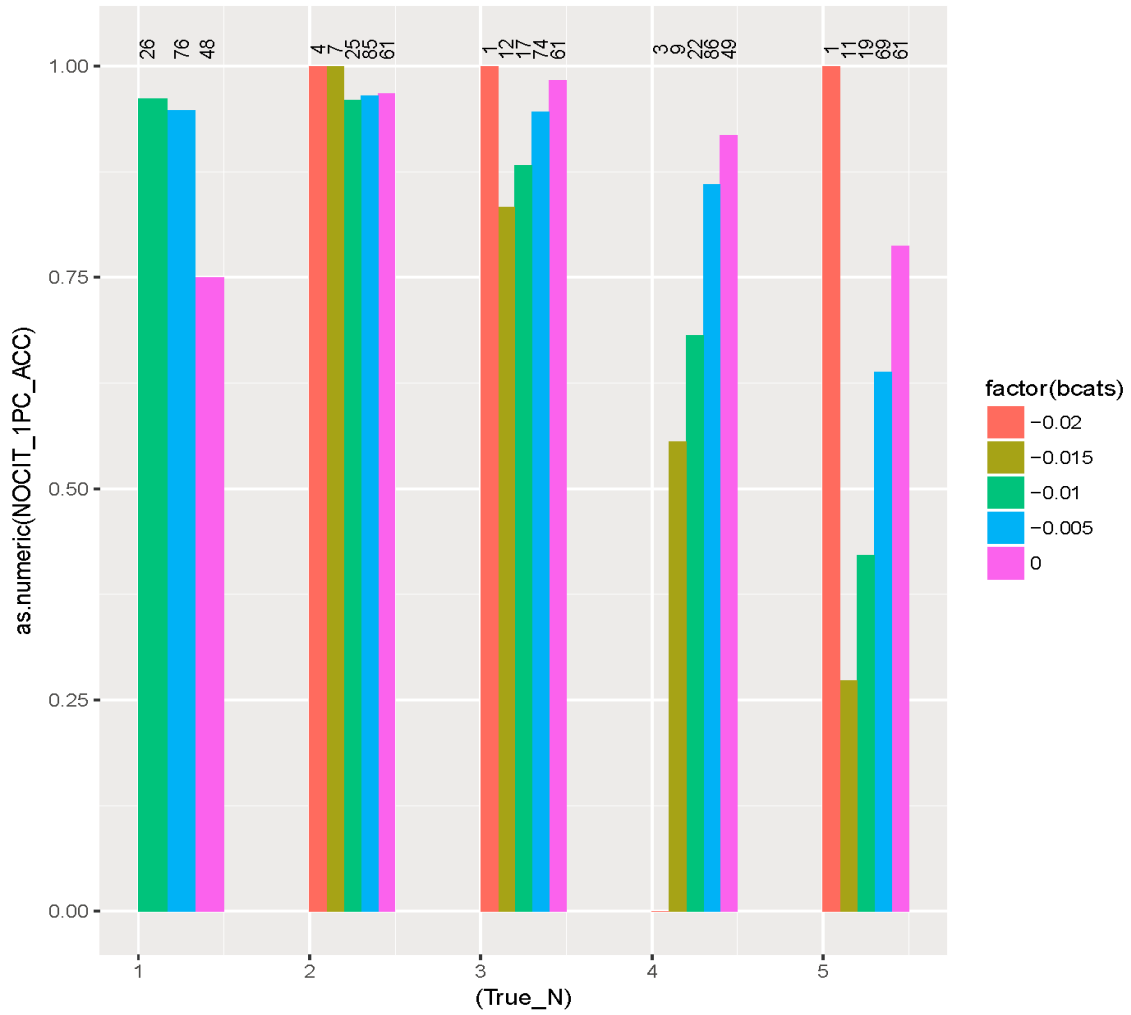


Figure 18. The degree of sloping observed in the STR profile. The exponent in the decay in fluorescence as a function of molecular weight,  $B$ , ranged from 0.02 to -0.04.  $B$  values significantly below zero correspond to compromised samples and indicate reduction in RFU signal as the length of the amplicons increase. (A) An electropherogram obtained from an untreated sample amplified at 0.25 ng (~40 copies). There is good intra-locus peak height balance across all heterozygous loci, and the total RFU signal is approximately equivalent across all loci labeled with the same dye, which is represented in the  $B$  value (-0.0005). (B) An electropherogram obtained from a sample treated with 30 sonication cycles amplified at 0.25 ng. The decrease in peak height as the fragment length increases is apparent and characteristic of the “sloping effect” observed in degraded profiles; this is represented by the highly negative  $B$  term obtained (-0.02). (C) The correlation between QI (Quality Index factor, per qPCR) and  $B$ . The treatment protocol, PCR amplification kit, and number of profiles are noted for each plot. The

plots are color-coded by density, where yellow and purple represent areas of highest and lowest sample density, respectively.

As expected, effectively inferring the NOC is affected by the degree of decay seen in the profile (Figure 19).



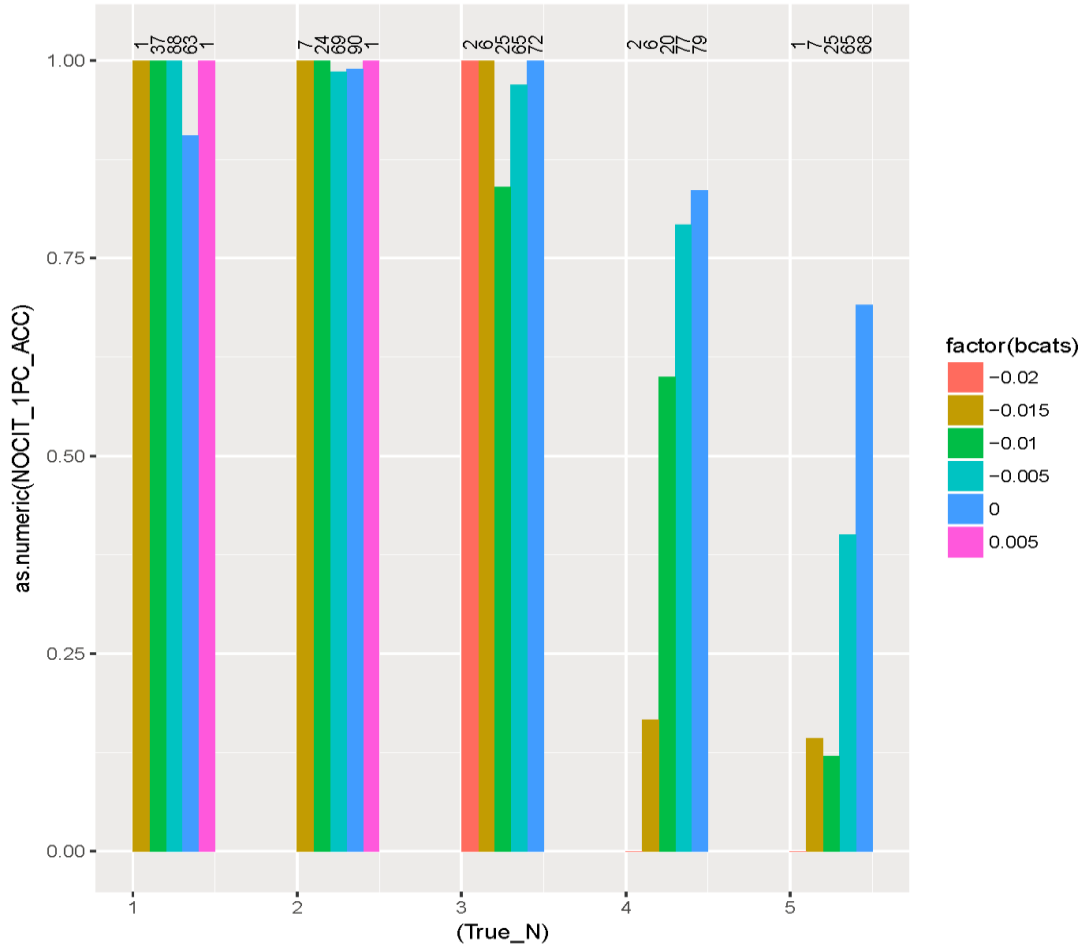


Figure 19. The accuracy of NOCIt as a function of degradation index,  $B$ , and the actual NOC (x-axis). The values on top of the figure signify the sample count that fell into that category. Top panel: the samples amplified with Identifiler Plus and run using a 3130 Genetic Analyzer (20s). Bottom panel: all samples amplified with GlobalFiler and run using a 3500 Genetic Analyzer (25s).

In the absence of spurious alleles or exceedingly large stutter events, MAC will always tend toward underestimation of the NOC as depicted in Figure 3D, wherein we plot the number of expected alleles versus the number of peaks observed in those allele positions. In comparison, NOCIt<sub>MAP</sub> results indicate that when we evaluate the NOC<sub>MAP</sub>, underestimation still occurs; however, it is less frequent and the difference between the actual and inferred NOC is smaller.

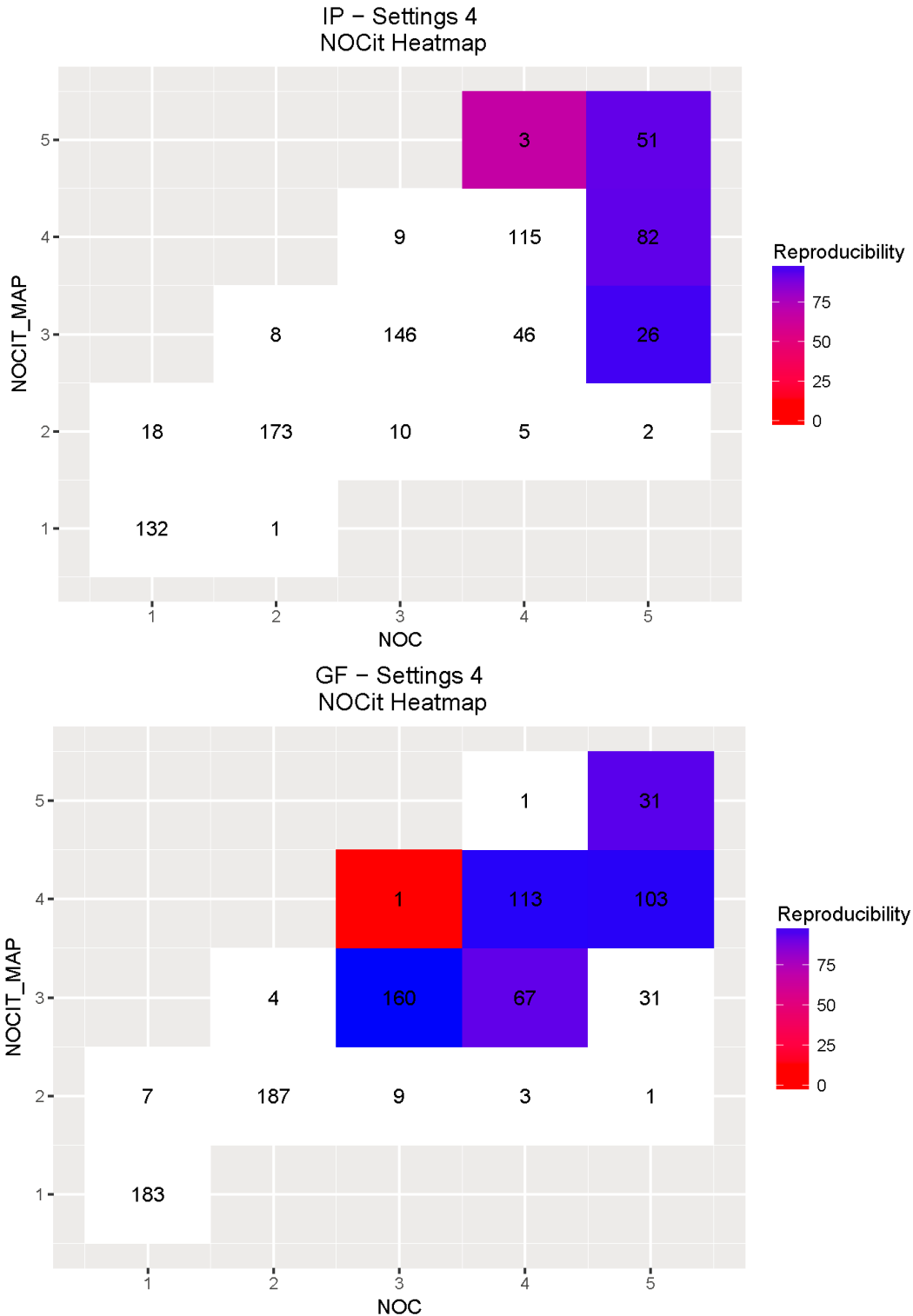


Figure 20. Top panel: the number of samples categorized as originating from 1, 2, 3, 4 or 5 contributors according to NOCIt<sub>MAP</sub>. The colors indicate the percentage of samples - out of 5 NOCIt replicate runs - within that category that resulted in the same NOC results. White coloration demonstrates the results were repeatable for all samples. The x-axis represents the true NOC. Top panel: Identifiler Plus. Bottom panel: GlobalFiler.

### 5.2.3.3 Discussion

Utilizing peak heights to determine the APP distribution is a valuable means by which to evaluate the possible number of contributors that could explain the EPG signal. In contrast to MAC based methods, we show that peak height based methods are able to accurately indicate the presence of the true NOC that comprise the EPG for signal originating from at least 0.03 ng of DNA for samples containing up to 5 contributors. Further, previous work has shown that MAC is prone to underestimating the true NOC when the actual NOC exceeds 3 because it relies upon the assumption that all alleles are detected and allele sharing is negligibly affecting the result. In contrast to the MAC method, NOCIt accuracy remains stable as the NOC contributors to the samples reach 5- persons. The greatest effect on the ability to ascertain the NOC to the evidence stems from the signal quality itself; in particular, low signal from minor contributors garnered from complex samples containing compromised DNA from multiple contributors has the largest effect.

### References

- [1] Perlin MW, Hornyak JM, Sugimoto G, Miller KWP. TrueAllele® Genotype Identification on DNA Mixtures Containing up to Five Unknown Contributors. *Journal of Forensic Sciences*. 2015;60:857-68.
- [2] Bright J-A, Curran JM, Buckleton JS. The effect of the uncertainty in the number of contributors to mixed DNA profiles on profile interpretation. *Forensic Sci Int Genetics*. 2014;12:208-14.
- [3] Moretti TR, Just RS, Kehl SC, Willis LE, Buckleton JS, Bright J-A, et al. Internal validation of STRmix™ for the interpretation of single source and mixed DNA profiles. *Forensic Sci Int Genetics*. 2017;29:126-44.
- [4] Benschop CCG, Haned H, Jeurissen L, Gill PD, Sijen T. The effect of varying the number of contributors on likelihood ratios for complex DNA mixtures. *Forensic Sci Int Genetics*. 2015;19:92-9.
- [5] Gill P, Buckleton J. A Universal Strategy to Interpret DNA profiles that does not require a definition of low-copy-number. *Forensic Science International - Genetics*. 2010;4:221-7.
- [6] Gill P, Kirkham A, Curran J. LoComatioN: A software tool for the analysis of low copy number DNA profiles. *Forensic Science international*. 2007;166:128-38.
- [7] Gill P, L. Gusmao, Haned H, Mayr WR, Morling N, Parson W, et al. 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. *Forensic Science International - Genetics*. 2012;6:679-88.
- [8] Perlin MW, Szabady B. Linear mixture analysis: A mathematical approach to resolving mixed DNA samples. *Journal of Forensic Sciences*. 2001;46:1372-8.
- [9] Wang T, Xue N, Birdwell JD. Least-Square Deconvolution: A Framework for Interpreting Short Tandem Repeat Mixtures. *Journal of Forensic Sciences*. 2006;51:1284-97.
- [10] Curran JM, Gill P, Bill MR. Interpretation of repeat measurement DNA evidence allowing for multiple contributors and population substructure. *Forensic Science International*. 2005;148:47-53.
- [11] Curran JM. A MCMC method for resolving two person mixtures. *Science and Justice*. 2008;48:168-77.
- [12] Puch-Solis R, Rodgers L, Mazumbder A, Pope S, Evett I, Curran J, et al. Evaluating forensic DNA profiles using peak heights, allowing for multiple donors, allelic dropout and stutters. *Forensic Science International: Genetics*. 2013;7:555-63.
- [13] Steele CD, Greenhalgh M, Balding DJ. Evaluation of low-template DNA profiles using peak heights. *Stats Appl Genet Mol Biol*. 2016;15:431-45.
- [14] Bright J-A, Taylor D, Curran JM, Buckleton JS. Developing allelic and stutter peak height models for a continuous method of DNA interpretation. *Forensic Science International: Genetics*. 2013;7:296-304.
- [15] Haned H, Egeland T, Pontier D, Pene L, Gill P. Estimating drop-out probabilities in forensic DNA samples: A simulation approach to evaluate different models. *Forensic Sci Int Genetics*. 2011;5:525-31.
- [16] Tvedebrink T, Eriksen PS, Mogensen HS, Morling N. Estimating the probability of allelic drop-out of STR alleles in forensic genetics. *Forensic Science International - Genetics*. 2009;3:222-6.
- [17] Perlin MW, Sinelnikov A. An information gap in DNA evidence interpretation. *Plos One*. 2009;4:e8327.
- [18] Buckleton JS, Curran JM, Gill P. Towards understanding the effect of uncertainty in the number of contributors to DNA stains. *Forensic Science International - Genetics*. 2007;1:20-8.

- [19] Paoletti DR, Krane DE, Raymer DL, Doom TE. Inferring the Number of Contributors to Mixed DNA Profiles. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*. 2012;9.
- [20] Taylor D, Bright J-A, Buckleton J. The interpretation of single source and mixed DNA profiles. *Forensic Sci Int Genet*. 2013;7.
- [21] Evett IW, Pope S. Is it to the advantage of a defendant to infer a greater number of contributors to a questioned sample than is necessary to explain the observed DNA profile? *Science & Justice*. 2014;54:373-4.
- [22] Brenner CH. Fairness in evaluating DNA mixtures. *Forensic Sci Int Genetics*. 2017;27:186.
- [23] Presciuttini S, Egeland T. About the number of contributors to a forensic sample. *Forensic Sci Int Genetics*. 2016;25:e18-e9.
- [24] Swaminathan H, Grgicak CM, Medard M, Lun DS. NOCI: A computational method to infer the number of contributors to DNA samples analyzed by STR genotyping. *Forensic Sci Int Genetics*. 16:172-80.
- [25] Marciano MA, Adelman JD. PACE: Probabilistic Assessment for Contributor Estimation— A machine learning-based assessment of the number of contributors in DNA mixtures. *Forensic Sci Int Genetics*. 2017;27:82-91.
- [26] Taylor D, Bright J-A, Buckleton J. Interpreting forensic DNA profiling evidence without specifying the number of contributors. *Forensic Sci Int Genetics*. 2014;13:269-80.
- [27] Slooten K. Accurate assessment of the weight of evidence for DNA mixtures by integrating the likelihood ratio. *Forensic Sci Int Genetics*. 27:1-16.
- [28] Rakay CA, Bregu J, Grgicak CM. Maximizing allele detection: Effects of analytical threshold and DNA levels on rates of allele and locus drop-out. *Forensic Science International - Genetics*. 2012;6:723-8.
- [29] Taylor D, Buckleton J, Bright J-A. Does the use of probabilistic genotyping change the way we should view sub-threshold data? *Australian Journal of Forensic Sciences*. 2015:1-15.
- [30] Swaminathan H, Garg A, Grgicak CM, Medard M, Lun DS. CEESIt: A computational tool for the interpretation of STR mixtures. *Forensic Sci Int Genetics*. 22:149-60.
- [31] Perlman MW, Sinelnikov A. An information gap in DNA evidence interpretation. *PloS One*. 2009;4.
- [32] Duffy KR, Gurrum N, Peters KC, Wellner G, Grgicak CM. Exploring STR signal in the single- and multicopy number regimes: Deductions from an in silico model of the entire DNA laboratory process. *ELECTROPHORESIS*. 2017:n/a-n/a.
- [33] Brookes C, Bright J-A, Harbison S, Buckleton J. Characterising stutter in forensic STR multiplexes. *Forensic Science International: Genetics*. 2012;6:58-63.
- [34] Perez J, Mitchell AA, Ducasse N, Tamariz J, Caragine T. Estimating the number of contributors to two-, three-, and four-person mixtures containing DNA. *Croatian Medical Journal*. 2011;52:314-26.
- [35] Tvedebrink T. On the exact distribution of the numbers of alleles in DNA mixtures. *Int J Legal Med*. 2014;128:427-37.
- [36] Bleka Ø, Storvik G, Gill P. EuroForMix: An open source software based on a continuous model to evaluate STR DNA profiles from a mixture of contributors with artefacts. *Forensic Sci Int Genetics*. 2016;21:35-44.
- [37] Qiagen Inc. QIAamp<sup>(R)</sup> DNA Investigator Handbook. Venlo 2010.
- [38] Grgicak CM, Urban ZM, Cotton RW. Investigation of reproducibility and error associated with qPCR methods using Quantifiler Duo DNA Quantification Kit. *Journal of Forensic Sciences*. 2010;55:1331-9.
- [39] Timken M, Klein SB, Buoncristiani MR. Stochastic sampling effects in STR typing: Implications for analysis and interpretation. *Forensic Science International - Genetics*. 2014;11:195-204.
- [40] Rutledge RG, Stewart D. Assessing the Performance Capabilities of LRE-Based Assays for Absolute Quantitative Real-Time PCR. *PloS One*. 2010;5:1-10.
- [41] Kelly H, Bright J-A, Curran JM, Buckleton J. Modelling heterozygote balance in forensic DNA profiles. *Forensic Science International: Genetics*. 2012;6:729-34.
- [42] Gilder JR, Inman K, Shields W, Krane DE. Magnitude-dependent variation in peak height balance at heterozygous STR loci. *Int J Legal Med*. 2011;125:87-94.
- [43] Life Technologies Corporation. Quantifiler<sup>(R)</sup> Duo DNA Quantification Kit User Guide. Foster City 2008.
- [44] Haned H, Pene L, Lobry JR, Dufour A, Pontier D. Estimating the number of contributors to forensic DNA mixtures: Does maximum likelihood perform better than maximum allele count? *Journal of Forensic Sciences*. 2011;56:23-8.

## Appendix

### Supplemental Calibration Information

To establish the functions and models, a set of calibration data is generated by the laboratory. This calibration data may be produced by preparing 500 single source amplified work products from at least 50 sources. These single source samples will be amplified with the amplification kit of interest, and run using the Genetic Analyzer of interest using the manufacturers (or laboratory's) recommended voltage and injection times. Prior to amplification, the samples shall be degraded/damaged to varying degrees. The concentration of the extracts can be obtained using the quantification kit where the amplicon size in the quant kit is ca. the same size as the smallest STR locus. For example, in the case of Identifiler Plus and GlobalFiler amplifications, the Quantifiler® Duo DNA Quantification Kit is an appropriate choice. Please note that human specific quantification (i.e., using qPCR) is preferred. The damaged series can be prepared by numerous methods, including but not limited to;

- 1) Digesting the extracted DNA with different concentrations of DNase I for some length of time. The extracted DNA is aliquoted to new microcentrifuge tubes and DNase I (various concentrations such as 0, 1, 2, 5, 10, 15, 20, 30 Unit/ml) will be added directly upon the start of the digestion reaction. The digestion mixture is incubated for a specified amount of time. To stop the degradation reaction and deactivate the enzyme, EDTA or other reaction inhibitor is added into each tube and the tube is incubated.

- 2) Inhibiting the samples by adding a known inhibitor such as humic acid to the sample.

- 3) UV-damaging the samples by placing the neat sample under a UV-lamp and bombarding the sample for some length of time.

- 4) Sonicating the sample using a sonication probe.

Figure A2. is a summary of the sample preparation tree that describes the number of samples that will be prepared. Note that 500 amplicons is the recommended number based on our results shown in Figure A1., which plots the log(likelihood) between the expected and true peak heights of a set of test samples versus the number of calibration samples used to parameterize the peak height model. We see that the likelihood levels off for the loci at ~500. Further, since we calibrate only from samples that exhibit heterozygous genotypes at a locus, we recommend that, if possible, as many heterozygous loci/persons, be included in the calibration.

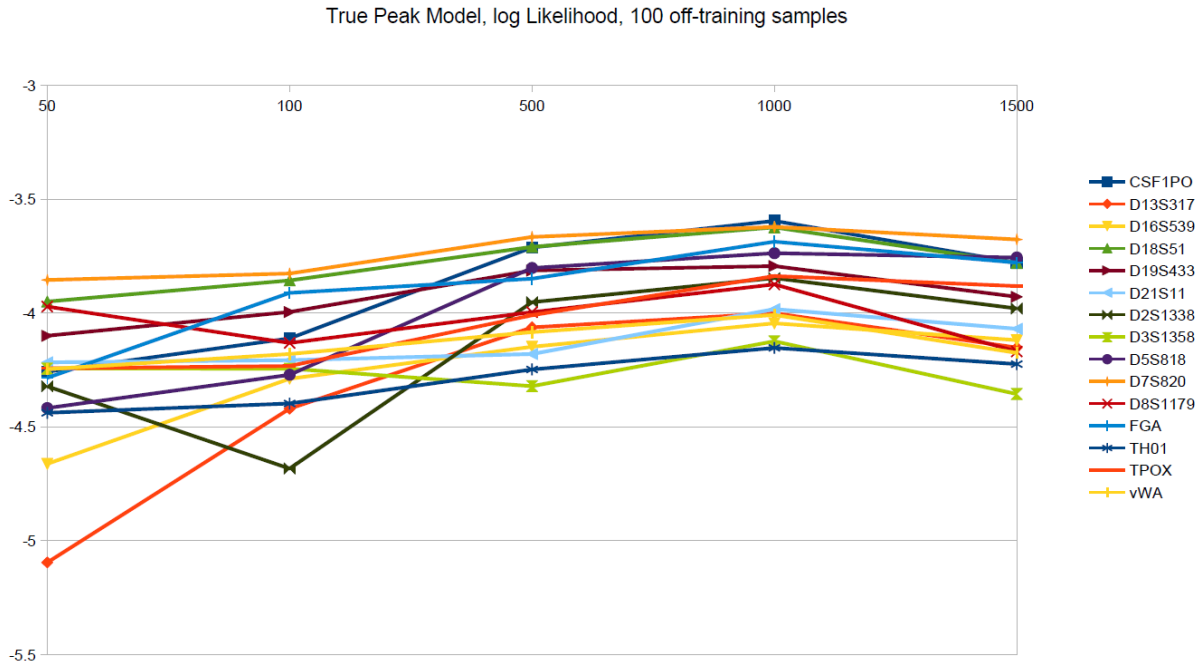


Figure A1. Log Likelihood between true and expected peak height obtained when using 50, 100, 500, 1000 and 1500 calibration samples in the calibration set.

Figure A2. depicts a tree diagram summarizing the number of degradation levels for the calibration set.

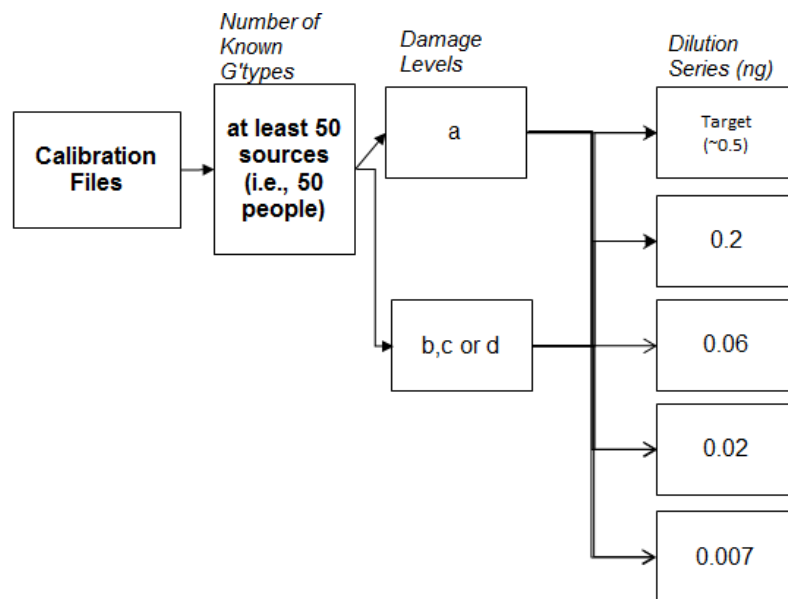


Figure A2. Tree diagram of the calibration flow chart. First, choose 50 single source samples, extract, choose degradation level for each extract (a **and** (b,c or d)), and amplify at the desired target. Note that target amounts are guidelines and the target mass may change based on the laboratory's SOP. Despite some flexibility in creating the calibration data set, it is recommended

that low-level 0.07 and 0.02 ng samples be amplified to ensure the ability to evaluate allele drop-out rates. NB: a= not damaged; b= slight damage; c= moderate damage; d= heavy damage.

Figure A3. demonstrates the severe and slight sloping that may be included in the calibration set. These were obtained using sonication.

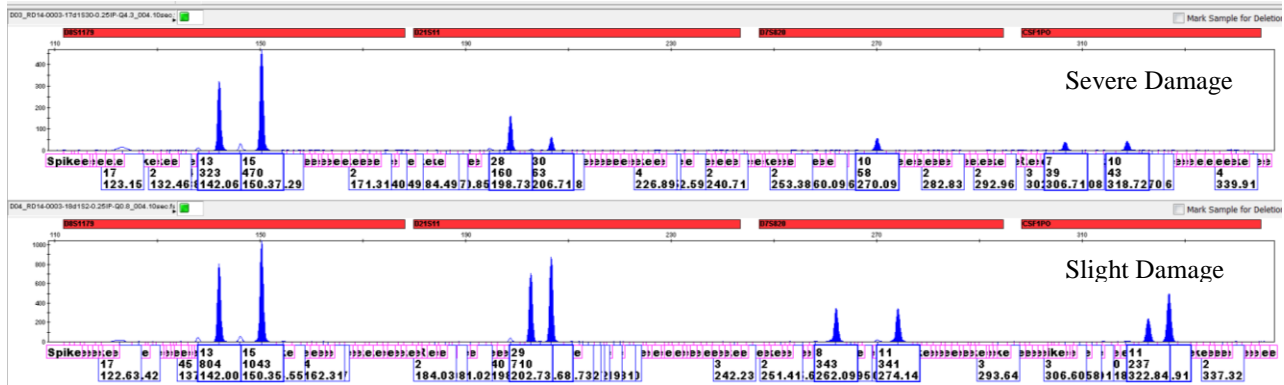


Figure A3. Representative electropherograms of ‘severely’ and ‘slightly’ damaged DNA generated by using probe-sonication.

The mechanism by which the laboratory damages the DNA can vary from laboratory to laboratory. However, results suggest that probe-sonication provides a reliable mechanism by which to develop slight, moderate and severe sloping electropherogram patterns.