Plexity considerations for the ForenSeq MainstAY Kit

Lucio Avellaneda, BS; Damani Johnson, BS; Rachel Houston, PhD
Department of Forensic Science, Sam Houston State University, Huntsville, Texas, USA

MEMBER THE TEXAS STATE UNIVERSITY SYSTEM

WASHINGTON, DC ISFG 2022 INTERNATIONAL SOCIETY FOR FORENSIC GENETICS

INTRODUCTION

Short Tandem Repeat (STR) kits coupled with capillary electrophoresis are still the gold standard for human identification in forensic laboratories. The discriminatory power of these kits is effective for most cases with high quality DNA. However, there continue to be examples where recovered DNA may not be sufficient to produce STR profiles due to low quantity, low quality, degradation, or mixtures[1].

Sensitivity of the MiSeq FGx™ (Verogen, San Diego, CA) system has been previously demonstrated with degraded or damaged DNA[2]. Challenging samples collected from skeletal remains, environmentally aged samples, sexual assault kits, or historical cases can benefit from Next Generation Sequencing (NGS).

The ForenSeq MainstAY Kit (Verogen) is an NGS option that places forensic laboratories at the focus of its development. This new NGS assay focuses on identification and contains only autosomal and Y- STRs (53 total) (**Figure 1).** In comparison, the ForenSeq DNA Signature Prep Kit (Verogen) has more amplicons and can be used for intelligence purposes for phenotypic and ancestry estimations[3-4].

Prior to implementation, NGS has additional considerations which must be taken into account. Plexity is the focus of this poster. The results obtained from a run are affected by the quantity and quality of the libraries[5]. Three factors that affect sample representation will be investigated. First, the size of the flow cell can be a standard or micro flow cell, which can obtain approximately 12.5 or 5 million paired read, respectively. Second, the number of samples on a flow cell affects sample representation. The limited space on the cell is needed to attach and generate clusters; more samples means less representation per sample. Third, the quality and concentration of the samples themselves must be considered. For example, a lower quality or concentration sample cannot be represented as effectively when batched with many other pristine samples.

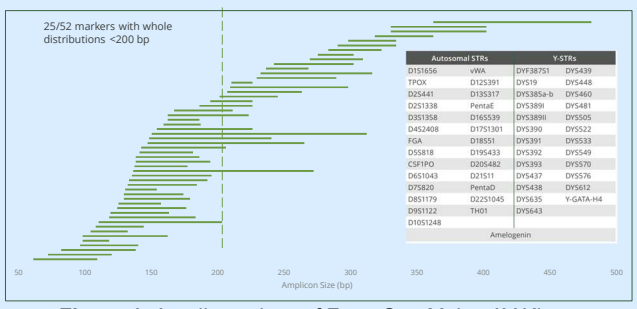

Figure 1. Amplicon sizes of ForenSeq MainstAY Kit

In this study, we evaluated the ForenSeq MainstAY Kit. The performance of the kit with respect to repeatability, sensitivity, and challenging samples was evaluated. Previously collected capillary electrophoresis (CE) data for the challenging samples were used for comparison purposes. Additionally, the challenging samples products run with the MainstAY chemistry were rerun using two flow cell models. The purpose of this poster is to consider the differences in sample number, sample quality, and library setup for the most effective use in laboratories.

MATERIALS & METHODS

Sample Acquisition: This research was conducted in accordance with the guidance of the SHSU Institutional Review Board for the Protection of Human Subjects. Controlled pristine extracts for the baseline and sensitivity testing part of the study were obtained from Verogen (San Diego, CA). Challenging samples (n=30) were obtained in collaboration with Southeast Texas Applied Forensic Science Facility (STAFS). This included 8 bone extracts, 2 burned or cremated bone extracts, 4 formaldehyde treated bone extracts, 3 embalmed tissue extracts, 4 blood extracts, 2 saliva extracts, 2 touch extracts, and 2 hair extracts.

Challenging Samples and DNA Extraction: Skeletal samples (n=14) extracted using Prepfiler® BTA Forensic DNA Extraction Kit, Embalmed tissue (n=3) extracted using QIAamp® DNA FFPE Tissue Kit, decomposing muscle tissue (n=3) extracted using QIAamp® DNA Investigator Kit, aged saliva (n=2); blood (n=4);rooted hairs (n=2) extracted using EZ1 DNA Investigator Kit, Touch DNA collected from handled rifle magazines using nylon FLOQSwabs™ (n=2) and extracted using the QIAamp DNA Investigator kit.

DNA Quantification: The DNA extracts were quantified using the Quantifiler™ Trio DNA Quantification Kit (ThermoFisher Scientific, Waltham, MA) on the 7500 Real-Time PCR® System (ThermoFisher Scientific) according to the manufacturer's protocol. Data was accepted with R^2 values greater than 0.99.

STR Analysis: DNA extracts were amplified with the Investigator® 24plex QS PCR Amplification Kit (QIAGEN) or GlobalFiler™ PCR Amplification Kit (Thermo Fisher Scientific). Amplified products were separated and detected on a 3500 Genetic Analyzer. Data were analyzed on GeneMapper ID-X v.1.4.

NGS Analysis: DNA extracts were amplified with the ForenSeq Mainstay Kit (Verogen)[6]. Amplified products were detected on a MiSeq FGx system. Data was analyzed using ForenSeq Universal Analysis Software v2.4

Experimental Setup:

Controlled Repeatability
- Four repetitions of 12 pristine DNA extracts and controls were prepared for a 64-sample run on a standard flow cell

Challenging Samples
- 30 samples were tested with two controls for a 32-sample run using the MainstAY chemistry on a standard flow cell

Sensitivity
- **Sensitivity 1**: Dilutions from (370-1pg) 2800M DNA (Promega)
- **Sensitivity 2**: Serial Dilutions (1ng – 8pg) 2800M DNA (Promega)
- Prepared by two examiners

Sample and Flow cell comparison
- 30 challenging sampels were rerun using the same normalized library and reran with the same 30 samples amplified with ePCR buffer on a Micro flow cell for direct comparison

RESULTS & DISCUSSION

- Controlled repeated samples resulted in 100% recovery with read counts ~ 150,000 (**Fig 2**)
- DNA recovery began to decrease at 37pg and below for both sensitivity tests (**Fig 3)**
  - Overall representation decreased as input decreased after 500-370pg
  - Sensitivity 1 had higher read count values Sens 2 at 125pg, 37pg, 16pg
  - Approximately 50-60% of loci were still recovered at 8pg
- Challenging samples decrease in count and recovery with the use of a 64-sample micro flow cell run (**Fig 4-5**)
- ePCR buffer on the micro flow cell displays an increase in recovery with inhibited samples (**Fig 5**)
- The 64 sample standard flow cell had higher uniform distribution of intensity for each sample (**Fig 6**)
  - This differs in both the sensitivity and in the challenging sample runs where some samples are over or under represented in the run results

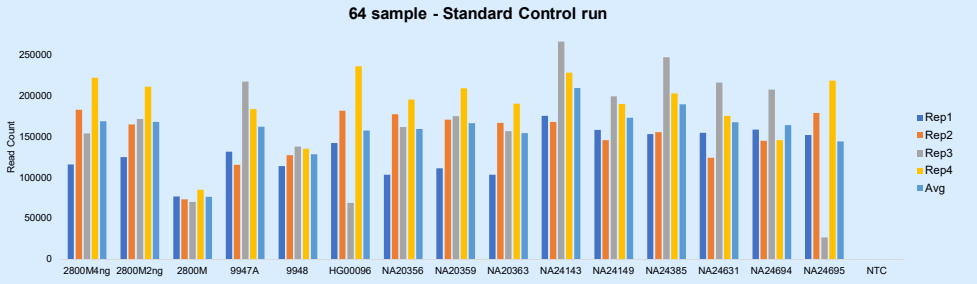

Figure 2: 12 samples plus 2800M control. Read counts between 100000 to 250000 were obtained in most cases. All obtained 100% recovery.

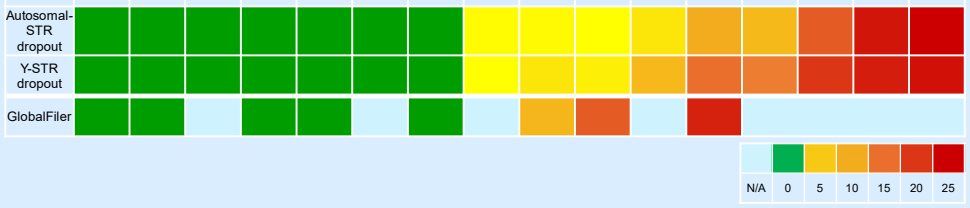

Figure 3: Autosomal and Y-STR dropout displayed for both runs as DNA input decreases. GlobalFiler dilution series tested for comparison.

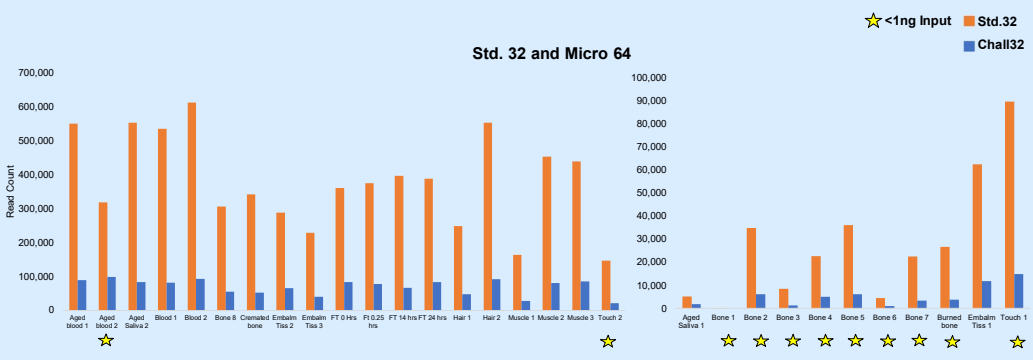

Figure 4a-b: Challenging sample Read count comparison between the standard flow cell and micro flow cell. b) figure separated due to low read count.

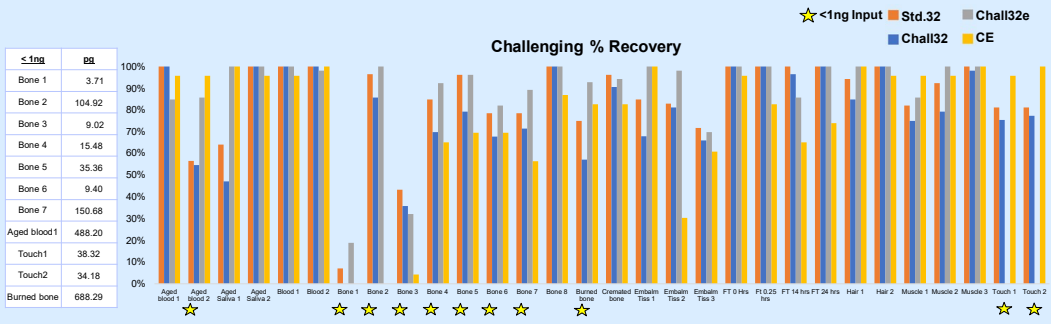

Figure 5: Challenging sample percent recovery comparison for flow cell and capillary electrophoresis. Table on left shown lists DNA inputs which were below the recommended 1ng input for the kit. (labeled with stars in the figures)

Runs	Total Read count Intensity	Max Per Loci Intensity	Cluster Density k/mm²	% Recovery considerations	DNA input differences
Standard 64 Controlled	9,740,165	2872	1383	All 100% recovery	1ng inputs x64 samples
Standard 32 Sens(1)	2,705,283	1595	1068	Decrease in recovery after 37pg	370ng-1pg x4rep
Standard 32 Sens(2) x2	4,358,969	2570	1102	Decrease in recovery after 31pg	1ng-8pg x4rep
Standard 32 Challenging	7,975,764	4703	1298	Recovery ranges from near 0 to 100%	1ng-3pg
Micro 64 Challenging	3,120,126	920	1470	Decrease from std 32: -2% to -17%	1ng-3pg x2

Table 1: Run setup comparison table and overview. Read count intensity and cluster density increase with each other. Max per loci intensity = Intensity / n / loci(53)

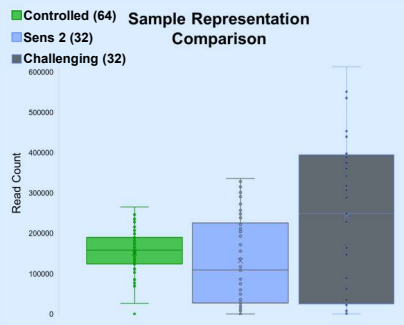

Figure 6: Sample representation distribution between samples of similar quality and those of varied quality

CONCLUSIONS

- Quality testing with MainstAY demonstrated the chemistry to be highly sensitive and reproducible, with consistent results observed in four repetitions of 12 samples, including controls (**Fig 2**)
- Normalization purpose to keep sample representation consistent varied with degraded or low input DNA decreasing the intensity read count for that sample
- Due to the use of fewer large input (near 1ng) samples in Sens 1, there is a increase in intensity compared to similar input samples used in Sens 2
- Dropout increases from standard to micro flow cell use as sample representation decreases (**Table 1**)
- ePCR buffer is a valuable tool to increase recovery of inhibited samples even on a smaller flow cell
- Required read count threshold for allele calls must be considered before library preparation as suboptimal batching and flow cell use may reduce certainty (**Fig 4-5**)
- When samples are of similar quality the read count is more evenly dispersed, but when samples are varied, disproportionate representation can cause pristine samples to obtain higher representation in comparison while challenging samples are lower (**Fig 6**)

ACKNOWLEDGEMENTS

The authors would like to thank Verogen for providing reagents for this research. In addition, the authors would like to acknowledge the support of the Department of Forensic Science at Sam Houston State University (SHSU) and the Southeast Texas Applied Forensic Science Facility (STAFS).

REFERENCES

[1] S.R. Hughes-Stamm, K.J. Ashton, A. Van Daal, Assessment of DNA degradation and the genotyping success of highly degraded samples, Int. J. Legal Med. 125 (2011) 341–348. https://doi.org/10.1007/s00414-010-0455-3.

[2] A.C. Jäger, M.L. Alvarez, C.P. Davis, E. Guzmán, Y. Han, L. Way, P. Walichiewicz, D. Silva, N. Pham, G. Caves, J. Bruand, F. Schlesinger, S.J.K. Pond, J. Varlaro, K.M. Stephens, C.L. Holt, Developmental validation of the MiSeq FGx Forensic Genomics System for Targeted Next Generation Sequencing in Forensic DNA Casework and Database Laboratories, Forensic Sci. Int. Genet. 28 (2017) 52–70. https://doi.org/10.1016/J.FSIGEN.2017.01.011.

[3] C. Xavier, W. Parson, Evaluation of the Illumina ForenSeqTM DNA Signature Prep Kit – MPS forensic application for the MiSeq FGxTM benchtop sequencer, Forensic Sci. Int. Genet. 28 (2017) 188–194. https://doi.org/10.1016/j.fsigen.2017.02.018.

[4] V. Sharma, K. Jani, P. Khosla, E. Butler, D. Siegel, E. Wurmbach, Evaluation of ForenSeqTM Signature Prep Kit B on predicting eye and hair coloration as well as biogeographical ancestry by using Universal Analysis Software (UAS) and available web-tools, Electrophoresis. 40 (2019) 1353–1364. https://doi.org/10.1002/elps.201800344.

[5] A. Senst, A. Caliebe, E. Scheurer, I. Schulz, Validation and beyond: Next generation sequencing of forensic casework samples including challenging tissue samples from altered human corpses using the <scp>MiSeq FGx</scp> system, J. Forensic Sci. 67 (2022) 1382–1398. https://doi.org/10.1111/1556-4029.15028.

[6] Verogen, ForenSeq MainstAY Product Line Reference Guide, (2022).