# Next Generation Sequencing (NGS) Feasibility and Guidance Study for Forensic DNA

Award Number: 2014-DN-BX-K023

Contract Number: CON00019796

DUNS:

Contract Period: January 1, 2015 to March 31, 2017

Special Report: Draft Final Summary Overview (DFSO)

Client National Institute of Justice (NIJ)

Address 810 Seventh Street, NW
Washington, D.C. 20531

Date February 9, 2019

# Table of Contents

Page

# 1. Purpose of the Project

The purpose of this study was to assess NGS applications for use in forensic DNA laboratories across the United States. The study included participation from key forensic DNA laboratories representing city, county, state, federal, academic, and research institutions with collaboration, education and validation of NGS technology as primary objectives. The result of research conducted under this grant is an objective assessment of the NGS technology for forensic applications to include strengths, vulnerabilities, and opportunities for improvement.

# 2. Project Design, Methods, and Data Analysis

The study conducted under this grant consisted of two phases: Phase 1 Performance Testing and Phase 2 Inter-laboratory Testing. During Phase 1, Battelle identified and assessed operations for each NGS workflow, optimized processes for each workflow, locked down the methods, and delivered standard operating procedures (SOPs) to the respective partner laboratories. Phase 1 was initiated by on-going discussions with the partner laboratories to assess the commercially available NGS products for use within this study. Performance testing was conducted in Phase 1 by the Battelle Team using DNA samples provided by the National Institute of Standards and Technology (NIST). Sequencing was performed across Illumina's MiSeq FGx, MiSeq RUO and the ThermoFisher Scientific Ion PGM™ platforms. In addition, commercially available DNA reagent kits, collectively encompassing marker sets that included short tandem repeat (STR), X-STR, Y-STR and single nucleotide polymorphism (SNP) spanning identity (iiSNP), and ancestral (aiSNP) panels, were obtained from Illumina (ForenSeq), Promega (PowerSeq) and ThermoFisher Scientific (Ampliseq) for utilization within the study. The resulting data was analyzed using bioinformatics software provided by Illumina (Universal

Analysis Software™), ThermoFisher Scientific (Ion Torrent Suite™), or Battelle (ExactID™, respectively.

Phase 2 consisted of three main activities: 1) SOP optimization, 2) external studies, and 3) final report submission, dissemination, and education of the forensic community. Upon receipt of partner laboratory data obtained from runs of the study samples, Battelle commenced with a preliminary review and analysis of the data beginning in June 2016. Battelle organized Box online sharing sites for each partner laboratory for this purpose. Data was uploaded to these sites by the partner laboratories upon completion of each NGS run and subsequently downloaded by Battelle for data analysis and review. Sequencing runs from all workflows in the study provided several quality metrics to evaluate the overall performance of the run, many of which are displayed by the instrument software. Selected quality metrics were also evaluated across partner laboratories to assess reproducibility. Battelle scientists assessed the initial metrics and analyzed the data using the appropriate commercial software. Statistical analyses were conducted by Battelle staff using the R statistical software package.

# 3. Project Findings

## 3.1 Quality Metrics

### 3.1.1. Illumina MiSeq

Assessments of quality for the Illumina MiSeq included metrics of cluster density, percentage of clusters passing filter, and phasing/prephasing. Across the four MiSeq runs in this study, which included both the FGx and RUO sequencers, the recorded metrics were within the defined ranges of acceptance for both the ForenSeq (Illumina) and PowerSeq (Promega) workflows. No failed or low-quality runs were observed on the Illumina MiSeq by Battelle or any

participating laboratory. Further, the resulting quality metrics for all MiSeq runs were similar to each other, both within and between participating laboratories.

#### 3.1.2 ThermoFisher AmpliSeq on Ion PGM (Ion PGM)

The percentage of Ion Sphere Particle (ISP) loading (or bead loading) is the number of wells on the sample chip determined to be "positive" for the presence of an ISP and indicates the efficiency of the Ion Chef at sample chip loading. The percentages were calculated by dividing the number of final library ISPs by the total number of library ISPs. The amount of usable sequence is based on the success of the library preparation and ISP enrichment procedures before sample loading on the chip. In this study, these values were within the expected ranges as defined by ThermoFisher Scientific.

### 3.2 Reproducibility, Sensitivity, Mixtures, and Casework Samples

#### 3.2.1 Ion PGM

All replicates yielded concordant genotype data across the entire study, except in 12 instances across two laboratories at the same locus within the Ion PGM workflow. Initial data analysis included concordance checks on all known samples at recommended sample input (used across all laboratories), which included replicates within the reproducibility study and included two replicate samples provided by NIST. These discordant calls are listed in Table 1.

**Table 1. Discordant SNP Calls from the Ion PGM Panel**

| SNP Locus | Sample Replicate | Detected Genotype | Truth Genotype | Reason for Discordance |
|---|---|---|---|---|
| rs2899826 (ancestry) | W21-P1-1 | GG | CG | Major Allele Frequency (imbalance) |
| | W21-P1-2 | | | |
| | W21-P1-3 | | | |
| | W21-P1-4 | | | |
| | W15-P1-1 | | | |
| | W15-P1-2 | | | |
| | W15-P1-3 | | | |
| | W15-P1-4 | | | |
| | W15-P1-3 | | | |
| | W15-P1-4 | | | |
| | W21-P1-1 | | | |
| | W21-P1-2 | | | |

The reason for the discordant calls within this reproducibility study was major allele frequency imbalance, as shown in the table. W15 and W21 are coded identifiers for the two participating laboratories in this portion of the study.

The default settings for filters were implemented to determine the percentage of library ISPs that pass polyclonal, low quality and primer-dimer filters. A polyclonal ISP is defined as a bead carrying indexed library products from two or more templates. Polyclonality measurements were generally within useable ranges in this study as defined by ThermoFisher Scientific, but showed high variances. Average percentages of useable reads were all within expected values, but displayed high run-to-run variances. Relatively high variances were observed in the reproducibility study in the Ion PGM workflow as discussed in the Quality Metrics section. This is thought to be due to a combination of factors, including the low number of laboratories (2) participating in this workflow. In the sensitivity study, higher levels of inter-laboratory variance were noted, especially at the lowest two input concentrations. There was also a degree of inter-locus variation in the identity panel, which at one locus (rs7520386) was very high, showing

imbalance at all input concentrations with two standardized dilution sets, also provided by NIST. These results are best explained by sub-optimal kit design. On the other hand, results that could be explained by reference to the genotype of sample S1 were found in the ancestry panel at two loci: rs7722455 and rs7520386, where imbalance was observed at every input amount, but not with sample S2.

### 3.2.2 Illumina ForenSeq

All genotyping results in the ForenSeq reproducibility study across the laboratories were consistent with truth data. However, autosomal STRs in this kit showed relatively high levels of allele imbalance across the study as compared to this metric in the PowerSeq kit. The sensitivity study revealed intermediate levels of allele and locus drop-out, again as assessed in comparison to the PowerSeq kit. Figure 1 shows the heat map graphic depicting the results obtained for the autosomal markers in the ForenSeq kit. Complete profiles obtained across laboratories are shown in green, while imbalanced results as defined by Illumina's default parameters are depicted in yellow, and interpretational, allele, and locus drop-out as similarly defined using the corresponding threshold divisors for each of these parameters are shown in brown, orange, and red, respectively.

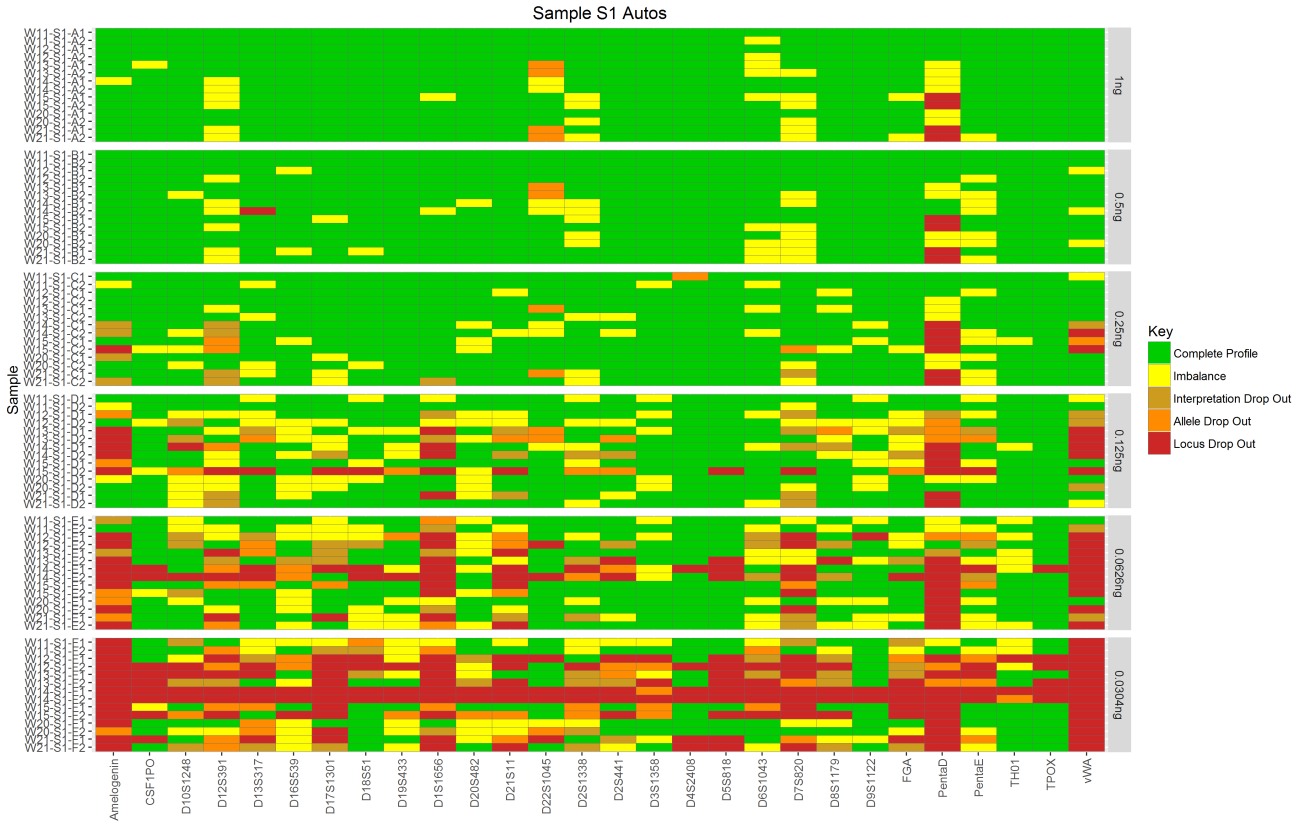


*Figure 1. Sensitivity Heat Map: Sample S1, Illumina autosomal STRs*

As can be seen in Figure 1, variation in sensitivity can be observed between the autosomal STR markers in the kit. For instance, the amelogenin, D1S1656, D7S820, Penta D, Penta E, and vWA loci appear to be more prone to imbalance as reflected by the accompanying default metrics in Sample S1. X-STR and Y-STR markers in the kit showed similar imbalance issues as revealed by these metrics. One X-STR locus in the kit, DXS10103, performed poorly in the sensitivity study as compared to the remaining X-STR loci, while another, DXS10074, was clearly less prone to imbalance and drop-out as defined by these parameters across the results from the participating laboratories.

An expected, decline in balance as the input amount of DNA was reduced was also observed in the identity SNPs within the ForenSeq panel. At or below an input amount of 125 picograms of



This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

DNA, the results across laboratories generally showed relatively high levels of imbalance and drop-out. All possible components of variation (between loci, laboratories, input amounts, and replicates within laboratories) were observed.

### 3.2.3 Promega PowerSeq

Across the reproducibility studies within the PowerSeq workflow, all genotyping calls were internally concordant as well as internally consistent with CE-based genotype expectations. Some limited additional genetic resolution was noted due to sequence variation within iso-alleles (alleles of the same size showing different sequences) using this kit. Sensitivity studies revealed the expected patterns of increased allele and locus drop-out as DNA input amounts were reduced.

The design of the Promega PowerSeq kit and the use of ExactID software allowed Battelle to examine stutter patterns (including prevalence and forms) at the STR loci contained within the kit. Figure 2 depicts both locus balance across the kit as well as the relative amounts of all forms of stutter (shown in orange). Unassigned reads that, upon initial review, could not be assigned as alleles or stutter, are shown in gray.

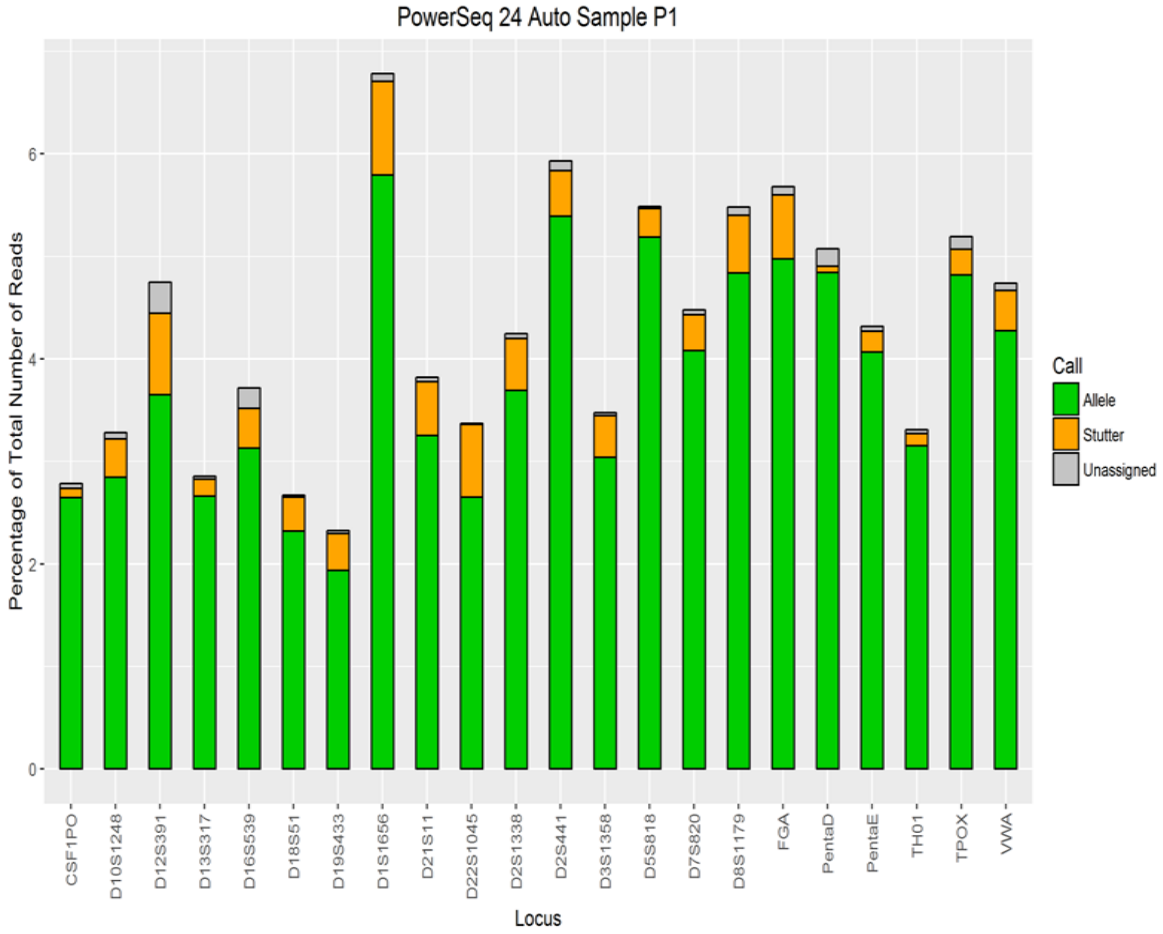


*Figure 2. PowerSeq reproducibility averages. Sample variances are omitted for clarity.*

In general, the ForenSeq and Ion PGM workflows showed a higher susceptibility to allele and locus drop-out compared to the PowerSeq workflow at the same input DNA concentrations. An example is shown in Figure 3, where the results of the first sensitivity run from the PowerSeq autosomal STRs are displayed.



This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

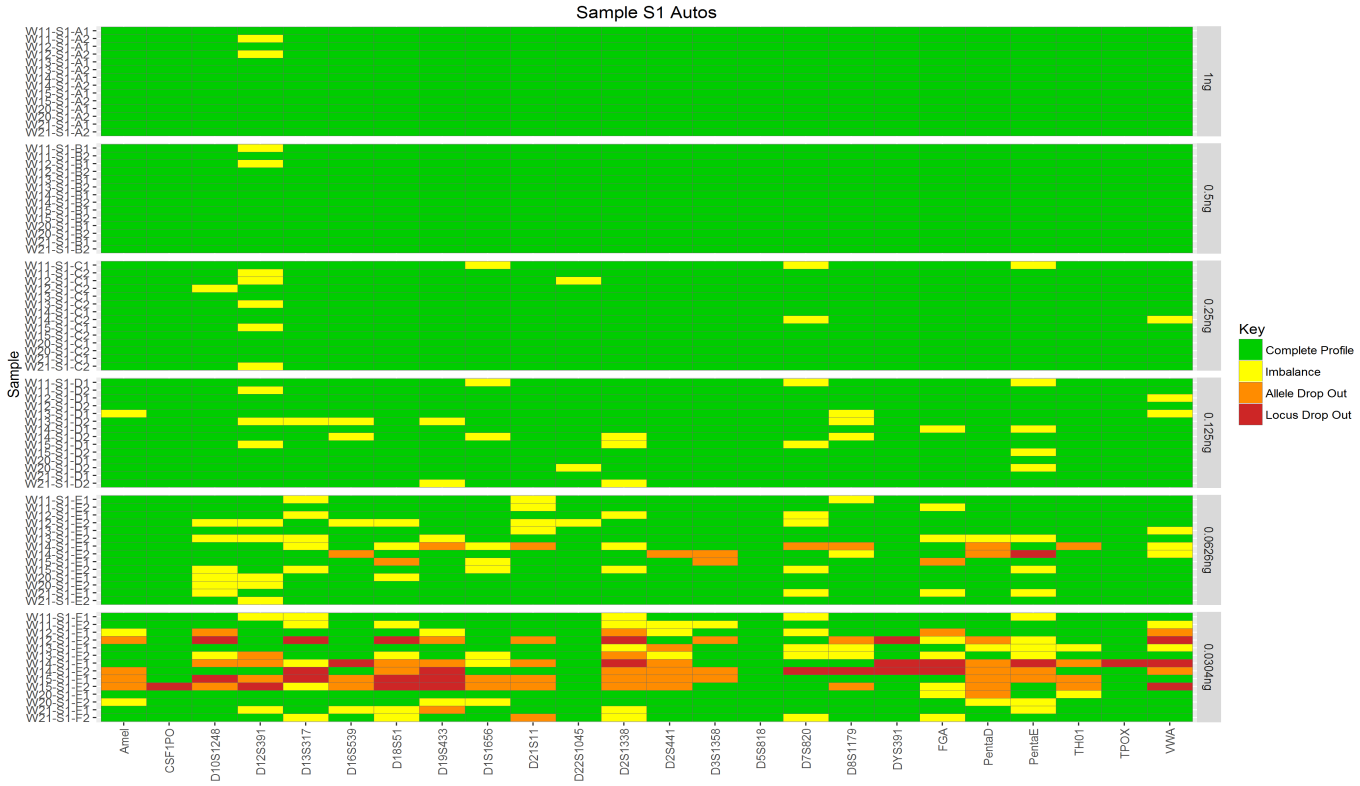


*Figure 3. PowerSeq Sensitivity 1 results*

Locus-specific patterns were observed in both the reproducibility and sensitivity studies, and were consistent across laboratories, revealing that the variances in these instances may be explained by reference to each individual kit design rather than individual laboratory performance. Analysis of the mixed samples within the PowerSeq workflow revealed the expected presence of additional alleles of the minor contributor to the mixture in all cases.

# 4. Implications for criminal justice policy and practice in the United States

From the study's objective and on a more technical level, with the exception of the discordances observed in the Ion PGM kit, the Phase 2 testing results demonstrated that NGS-based genotyping across different laboratories is consistent, accurate and reproducible. This is a

critical observation from a technology transfer perspective considering the study's data were generated from a range of laboratories varying in both size and mission, to include those gaining their first exposure to the NGS technology through this study to those with considerable experience. Further, discussions with partner laboratories have identified a consensus of key forensic applications for which NGS would provide near-term operational gains, as well as those for which near-term application appears more limited. A more formalized gap assessment will examine such information in greater detail. Such discussions, augmented by experiences gained within this study, have also indicated areas in which commercial NGS product service and support presently departs from forensic laboratory expectations, as based upon legacy capillary electrophoresis experiences. Finally, an appreciation for the richness and depth of the resulting data has stimulated discussions within Battelle and across partner laboratories for consideration of novel approaches to interpret and apply such information within the relevant forensic context. Specifically, a commitment to using all the MPS data is warranted, as detailed analysis has revealed patterns in the data that are consistent with current models related to the enzymatic amplification process of repeat sequence targets.


This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.