



**The author(s) shown below used Federal funding provided by the U.S. Department of Justice to prepare the following resource:**

**Document Title:** Validation of Illumina's MiSeq FGx NGS Platform for Transition to Casework

**Author(s):** Elisa Wurmbach, Ph.D.

**Document Number:** 306168

**Date Received:** March 2023

**Award Number:** 2015-DN-BX-K005

**This resource has not been published by the U.S. Department of Justice. This resource is being made publicly available through the Office of Justice Programs' National Criminal Justice Reference Service.**

**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.**

## **Validation of Illumina's MiSeq FGx NGS Platform for Transition to Casework**

Submitted to the National Institute of Justice

Grant Number **2015-DN-BX-K005**

Elisa Wurmbach, Ph.D., City Research Scientist

[EWurmbach@ocme.nyc.gov](mailto:EWurmbach@ocme.nyc.gov)

DUNS NO.:

GRANTEE IRS/VENDOR NO.:

Recipient Organization:

New York City Office of Chief Medical Examiner

Department of Forensic Biology

421 East 26<sup>th</sup> Street, New York, NY, 10016

### **Final Draft**

Grant Period: 01/01/2016 to 12/31/2017

Final Draft submitted on: Sep 21, 2017

---

Elisa Wurmbach, Ph.D.

City Research Scientist

## **ABSTRACT**

The basis of individual identification is DNA typing of short tandem repeats (**STRs**). This technique has brought a standardized, quantitative method with strong statistical underpinnings to the criminal justice system. While fundamental principles behind STR typing have not changed, new instrumentation and informative markers developed over the past few years have the potential to address limitations of current techniques.

Current DNA analysis techniques are based on capillary electrophoresis (**CE**) separating PCR products by length and fluorescent label. Inherent limitations include: no capabilities for multiplexing, failing to provide sequence specific information that could improve STR analysis, and approaching the maximum number of STRs that can be processed, thus effectively abandoning improvements from new STRs and single nucleotide polymorphisms (**SNPs**) that could improve individual identification, mixture deconvolution and hereditary analysis - at least in a single system.

The forensic community is beginning to evaluate next generation sequencing (**NGS**) as a means to overcome these problems. NGS technology not only adds additional sequencing information, but has a nearly unlimited capacity for additional STRs as well as SNPs, thereby enhancing individual identification. NGS methods also have the potential for significant improvements in throughput at lower costs.

This study evaluated Illumina's MiSeq FGx forensic sequencing platform:

Concordance was verified by using current CE methods.

Read numbers varied tremendously within samples for specific loci, revealing an imbalance of the multiplex reaction kit. Samples with low DNA inputs had lower read numbers leading to less reliable results: more dropouts and greater imbalance of read numbers in heterozygotes. Therefore, the DNA input had the strongest effect on the outcome. However, full STR profiles were occasionally obtained from as little DNA input as 50 pg.

Genotype errors occurred primarily due to dropouts and could only be detected by comparison to repeatedly run samples.

Essentially, no errors were found due to contamination caused by pooling or handling of the samples. Spurious alleles were found in Y-STRs on a few occasions: in four female samples that could be easily edited, and in one male sample, which was only found by comparison and considered as a genotype error.

Using ideal conditions does not assure 100% outcomes neither for the reaction samples nor the positive control due to drop-outs and genotype errors. In addition, typed ADI and stutter were observed in all runs and required manual editing.

Degraded and challenging samples resulted in almost full profiles.

(393 words)

Abstract . . . . .	2
Table of Contents . . . . .	4
Executive Summary . . . . .	7
Final Draft . . . . .	18
1. Introduction . . . . .	18
1.1 Statement of the problem . . . . .	18
1.2 Literature, Citations and Review . . . . .	19
1.3 Purpose of the Project . . . . .	22
2. Methods . . . . .	22
2.1 Sample collection, DNA extraction and quantification . . . . .	22
2.2 CE concordance data: PowerPlex® Fusion and AmpFISTR® Yfiler® . . . . .	22
2.3 Experimental overview and ForenSeq™ DNA Signature Prep library preparation . . . . .	23
2.4 Data analysis . . . . .	23
3. Results . . . . .	24
3.1 Evaluation of Illumina's MiSeq FGx Platform . . . . .	24
3.1.1 <i>Concordance</i> . . . . .	24
3.1.2 <i>Sensitivity</i> . . . . .	24
3.1.3 <i>Flow Cell Capacity</i> . . . . .	25
3.1.4 <i>Allele Coverage Ratio (ACR)</i> . . . . .	26
3.1.5 <i>Stutter</i> . . . . .	26
3.1.6 <i>Number of Reads: Balance</i> . . . . .	27

3.1.7	<i>Effects of Experimental Conditions on Number of Reads</i>	27
3.1.8	<i>Positive and negative controls</i>	29
3.1.9	<i>Sequence Variants</i>	29
3.1.10	<i>PCR and/or Sequence Errors</i>	30
3.1.11	<i>Errors in the STR genotypes</i>	30
3.1.12	<i>Analysis of iSNP genotypes</i>	31
3.2	<i>Mixed DNA Samples</i>	32
3.2.1	<i>Two-Person mixtures</i>	32
3.2.2	<i>Mixtures of more the two persons</i>	33
3.3	<i>Degraded and challenging DNA Samples</i>	33
3.3.1	<i>Degraded DNA</i>	33
3.3.2	<i>Challenging samples: DNA from blood cards and teeth</i>	34
3.4	<i>Costs and Throughput</i>	34
4.	<i>Conclusions</i>	35
4.1	<i>Discussion of Findings</i>	35
4.2	<i>Implications for Further Research</i>	39
4.3	<i>Implications for Policy and Practice</i>	40
5.	<i>References</i>	41
6.	<i>Disseminations of Research Findings</i>	47
7.	<i>Tables and Figures</i>	48
7.1	<i>Tables</i>	48
7.1.1	<i>Table 1: Overview of experimental runs</i>	49

7.1.2 Table 2: Effects of Experimental Conditions on Read Numbers	50
7.1.3 Table 3: Sequence variants	51
7.1.4 Table 4: Mixture Study: two-person mixtures	53
7.1.5 Table 5: Mixture Study: more than two-person mixtures	54
7.1.6 Table 6: Illumina and PP Fusion for one sample that was degraded by boiling	55
7.1.7 Table 7: Illumina and PP Fusion for degraded blood and teeth samples	56
7.1.8 Table 8: Costs and Throughput.	57
7.2 Figures	48
7.2.1 Figure 1: Sensitivity testing	58
7.2.2 Figure 2: Flow Cell Capacity	59
7.2.3 Figure 3: Allele Coverage Ratios (ACRs)	60
7.2.4 Figure 4: Number of Reads - Balance	61

## EXECUTIVE SUMMARY

### 1. Introduction

The foundation of the identification of individuals in modern forensic science is DNA typing by short tandem repeat (**STR**) analysis. STRs are DNA sequences with repeats of the same few (2-5) bases in order. STRs are highly variable in their length across individuals and analyses of various STRs create a unique profile for every individual. The acceptance and the implementation of STR analysis has revolutionized modern forensic science and has brought a standardized, quantitative method with a strong statistical base into the criminal justice system that has dramatically improved just and impartial outcomes. While the fundamental principles behind STR typing have not changed, new instrumentation and more informative biological markers developed over the past few years have the potential to address the limitations of current techniques. Limitations that can/have been improved upon include: mixture interpretation (samples that contain DNA of more than one individual), as well as increasing throughput at lower cost (keeping up with the many crime cases as well as any large number of unexpected samples such as those from a mass disaster).

Current DNA analysis methods for individual identification have technical, cost and throughput limitations. These methods are based on capillary electrophoretic (**CE**) sizing of a selected group of amplicons. CE is a detection method processing only a single sample and does not possess the capabilities for multiplexing samples and is consequently slow. Furthermore, CE a crude analog measure of amplicon length and consequently fails to detect informative sequence specific information that could significantly improve individual identification and aid in mixture deconvolution. Finally, because of the inherent constraints of the CE method itself: the need for sufficient loci separation for adequate resolution along with the limited number of amplicon lengths it is capable of separating, CE is approaching the maximum number of STRs it can process, effectively abandoning improvements from new autosomal, Y, and X STRs as well as single nucleotide polymorphisms (**SNPs**) that can improve individual identification, mixture deconvolution, and hereditary analysis - all in a single system. The New York City Office of Chief Medical Examiner (**NYC OCME**) typically processes about 750 samples per month. The problem is that with an ever increasing number of samples along with the need for more informative genetic identity markers for better mixture



deconvolution, CE may have reached its limit. Consequently, it is necessary to begin validating the next generation of forensic DNA analysis platforms capable of higher throughput, and increased discriminative power.

The use of Next Generation Sequencing (**NGS**) also called “massively parallel sequencing” (**MPS**) in forensic testing offers a variety of advantages. These include: the ability to add new, informative STRs and/or SNPs capable of improving mixture deconvolution and kinship analysis while maintaining legacy STRs; amplicon length is no longer a limiting constraint as multiple informative loci of similar or identical length can be used increasing discrimination; the potential discovery and incorporation of new, multiple, short informative STRs and/or SNPs; the fact that degraded samples that currently produce few STRs may become deducible; and finally, the ability to barcode individual samples, so that hundreds of samples can be multiplexed. These advantages combined create the potential to drastically increase sample throughput and consequently reduce sample costs. The potential for NGS to improve forensic DNA analysis and consequently produce just legal outcomes is clear, and there have been several publications evaluating its use for STR analysis as well its potential in mitochondrial analysis.

The overall goal of this study is to evaluate Illumina's MiSeq FGx next generation forensic sequencing platform for use in routine casework and to detect any limitations or restrictions that should be addressed before implementation.

## 2. Methods

This study was approved by the New York City Department of Health and Mental Hygiene's Institutional Review Board (IRB# 15-125). An Institutional Review Board is an ethical committee assigned to each institution where human subjects are involved in research. The purpose of the IRB is to ensure that the rights and safety of all human subjects and/or personal identifying information is protected. Buccal swabs were obtained with informed consent from 16 volunteers (nine males and seven females). The collected samples were anonymized. DNA was extracted using an M48 BioRobot® using their MagAttract® extraction kit and quantified using Quantifiler® Trio. A negative control

was included with each extraction to detect any contamination; if it tested positive, all samples in the batch were discarded.

Concordance data for capillary electrophoresis was obtained with PowerPlex® Fusion and AmpFISTR® Yfiler® by amplifying the DNA from the 16 individual samples using the PowerPlex® Fusion 5C Kit. In addition, male DNA was amplified with AmpFISTR® Yfiler® PCR Amplification Kit.

This project consisted of 16 experimental runs and tested various parameters including: concordance, repeatability, sensitivity, allele coverage ratio (ACR), stutter, sequence variants, reliability, mixed DNA, degraded samples, and casework samples. Each experimental run contained the reaction samples as well as two controls (one positive and one negative). Experimental runs were performed on Illumina's MiSeq FGx system in the Forensic mode (Illumina) using the MiSeq FGx Reagent Kit. Illumina's default settings were specifically chosen for all runs.

It is important to note that all runs and methods performed in this project were in accordance with the manufacturer's instructions and/or default settings in all cases. This was meticulously designed in order to create standard experimental runs of which analysis and/or results can be easily compared with runs at other laboratories. This was also to test the capability of the MiSeq FGx platform at its recommended settings and create a benchmark which can be used for further improvements in the default settings of the platform.

### 3. Results

#### 3.1 Evaluation of Illumina's MiSeq FGx Platform

The evaluation of Illumina's MiSeq FGx Platform included tests and assessments of single-source samples of high quality for concordance with other platforms testing: sensitivity, flow cell capacity, repeatability, allele coverage ratio (ACR), stutter, and sequence variants.

##### 3.1.1 Concordance

Concordance of a system, method, or platform refers to the concept that this new system will produce the same (true) results as other systems, methods, or platforms previously accepted. In this study, DNA samples of 16 individuals (nine male and seven female) were assessed with both the ForenSeq™ Primer Mix A and PowerPlex® Fusion. In addition to this, the nine male samples were tested with AmpFISTR® Yfiler® as well. Notably, concordance was verified for all loci common between the different tests.

For the additional loci included in the ForenSeq™ DNA Signature Prep Kit and were not assessed by the CE technique (5 aSTRs, 11 Y-STRs, and 7 X-STRs), the loci were evaluated for reproducible outcomes in at least two experimental runs and/or in the various dilutions. Consistency was verified for all 16 individual samples at all STR loci.

### 3.1.2 Sensitivity

Evaluating the sensitivity of a kit for Forensic Casework is important and will allow defining an input range that processes the samples in sufficient quality. This is especially true for non-probative casework samples that typically contain much lower than the 1ng DNA input recommended by the manufacturer.

In this project, the sensitivity of the ForenSeq™ DNA Signature Prep Kit was evaluated for Primer Mix A using various DNA inputs ranging from 800pg (0.8ng) all the way down to 50pg (0.05ng) compared to the 1ng (1,000pg) DNA input recommended by the manufacturer. The outcomes for these lower DNA inputs produced 100% correct outcomes for the STRs at all DNA inputs including at 50pg. However, not every reaction sample at a low concentration resulted in a full profile. In most cases, aSTRs had better outcomes compared to Y and X STRs. This seemed to be exacerbated as the DNA input becomes lower.

### 3.1.3 Flow Cell Capacity

The flow cell is the location where DNA clonal bridge amplification occurs in the MiSeq FGx instrument. Flow cell capacity refers to the phenomenon that there is a limited amount of estate on the flow cell itself resulting in an unintentional competition for surface area to replicate on.

In this study, flow cell capacity was investigated using the varying DNA inputs (1ng, 800pg, 400pg, 200pg, 100pg, and 50pg) however, this time the ForenSeq™ DNA

Signature Prep Kit Primer Mix B was implemented. Kit B has an additional 78 targets (loci) compared to Kit A and thus gives more insight into the consequences of increasing the number of loci being tested on the same flow cell as before. Findings showed that the lowest DNA input amount that resulted in full profiles for the STRs was 200pg (compared to the 50pg of Kit A). Findings also included the fact that for the aSTRs, allele drop-out (**ADO**) started at 400pg DNA input. More drop outs occurred at lower DNA inputs. However, allele drop-ins (**ADIs**) were found for the Y-STRs at the higher DNA inputs of 400pg and 800pg, reaching >20% of the true allele. This can be attributed to sequencing problems and typed stutter, which were not editable. In addition to this, by comparing the average data of the six reaction samples generated from Primer Mixes A and B, differences were notable, particularly for lower DNA inputs (50pg and 100pg), for which the reaction samples tested with Primer Mix B showed more drop outs.

#### 3.1.4 Allele Coverage Ratio (ACR)

The Allele Coverage Ratio (**ACR**) gives insight into the level of balance of the read numbers of heterozygous signals. ACR can be determined for heterozygous loci by dividing the lower number of reads by the higher number of reads. Equal numbers of reads will result in ratios of 1.0 and increasing imbalance will result in smaller ratios. Theoretically, heterozygous signals should lead to similar read numbers. ACRs are determined for each locus. If there is a greater imbalance, then the sample may seem to contain DNA from at least another individual. Moreover, a great imbalance may lead to allele drop outs.

Findings included the fact that most aSTRs at higher DNA inputs (800pg and 400pg) led to ACRs of >0.7, while at lower DNA inputs (100pg and 50pg) led to ACRs of <0.7 showing a relationship between DNA input and ACR. This is concordant with the hypothesis that higher DNA inputs would result in higher read numbers simply because there is more DNA.

#### 3.1.5 Stutter

A STR "stutter" is an artifact of the amplified DNA. In some cases when the DNA polymerase comes into contact with a strand of DNA repetitions of the same units consisting of approximately four nucleotides, it may mistakenly slips one unit leading to a shorter or longer product than the template. Stutter in most cases is one repeat unit

shorter (-1 stutter), however, it can be more than one repeat unit(s) shorter or longer than the true allele.

It is important to understand the frequency of stutter for the various loci in order to characterize single source samples. This may help an analyst or analysis software to be able to recognize when a stutter is present and distinguish it from the true allele of an additional contributor.

Observations from this study included up to three stutters were typed from Illumina;s Primer Mix A for aSTRs at higher DNA inputs (400pg and 800pg). This number went up to six stutters that were typed for the samples with 100pg DNA input and 12 stutters were typed from samples with DNA inputs of 50pg. The Y- and X-STRs showed slightly more typed stutters at higher DNA input, while at lower input the number of typed stutters was comparable to the aSTRs. It was also observed that plus-one stutter was typed less frequently, and minus-two or plus-three stutter occurred rarely. At higher DNA input, more stutters can be found that were just above the threshold. At lower DNA inputs, a greater number of stutters occurred that were closer to the true allele.

### 3.1.6 Number of Reads: Balance

The numbers of reads that are produced for different loci within the same sample offer valuable information regarding the balancing of the chemistry of the kit. It was observed that loci with high read numbers were consistently high and loci with low read numbers were consistently low. Autosomal STRs revealed a 15-fold difference between the locus with the highest average number of reads and the locus with the lowest average number of reads; Y-STRs a 13-fold difference; X-STRs a 47-fold difference; and iSNPs an 80-fold difference. These findings indicate that each locus should be considered independently for stochastic and analytical thresholds.

### 3.1.7 Effects of Experimental Conditions on Number of Reads

Findings included the observation that replicate experimental runs that were run within a short period of time from each other showed good repeatability. However, replicate experimental runs that were performed 11 weeks apart, showed degrading repeatability suggesting that there may be changes in the kit overtime (kit stability/aging). Next, increasing the number of samples in an experimental run from 32 to 96 samples while

keeping DNA input (at the recommended 1ng) and the number of loci (Kit A) constant, resulted in a reduced number of reads by a half in the 96 reaction sample experimental run. In addition to this, read numbers also declined when the number of loci sequenced was increased (Kit A, 153 loci vs. Kit B, 231 loci), while the number of reaction samples and DNA input was kept constant. Moreover, reducing reaction sample DNA input by half, from 1ng to 0.5ng while keeping other variables constant (32 samples and 153 loci, Kit A), resulted in a reduced number of reads by 3.5-fold for the 0.5ng reaction samples. Finally, in samples where DNA was present at: 800, 400, 200, 100, and 50pg. In duplicate runs, the number of reads at 800pg were 7.4 and 4.5 times greater than those at 50pg respectively. In an additional experimental run that used the same range of sample DNA inputs but with Kit B (increased targets), 800pg reaction samples had a 6.8 fold increase in read numbers compared to 50pg samples.

Perhaps most relevant with respect to forensic casework were the changes in read numbers when the amounts of DNA were varied within a run. This is important in casework where multiple contributors may be present at different concentrations within a sample.

### 3.1.8 Positive and Negative Controls

Each experimental run contained one positive control with the exception of Expt. 12, which had 5 positive controls. The positive control that was utilized was the recommended 2800M DNA (Illumina). 1000pg (1ng) of the 2800M DNA was used in every experimental run (as recommended). The positive controls were 100% positive in a few experimental runs. In a number of experimental runs the positive control lacked some genotypes. Consequently, the positive control failed to test all loci within the multiplex kit, meaning it behaved as a reaction sample rather than an "ideal" positive control.

Each experimental run contained one negative control, (including Expt. 12). Using default settings, coverage information of the negative no template controls (**NTCs**) for 15 of the 16 experimental runs showed no reads. Expt. 8 showed 34 reads for A in the iSNP rs1493232. This is important because the library preparation includes a step in which all of the samples are pooled. Thus a negative control showing no reads supports the

notion that no cross-contamination occurred throughout library preparation and sequencing steps.

### 3.1.9 Sequence Variants

A total of 16 individuals were tested and several sequence variants were identified. The type of sequence variants includes so-called "SNPSTRs", which are SNPs within the STR repeats or its flanking regions, or changes in the pattern of the repeats for compound or complex STRs. Sequence variants were found at 14 loci.

### 3.1.10 PCR and/or Sequence Errors

Sequence errors represented only a small number of reads (about 2% of the true allele) that differ from the dominant allele. Most likely, these errors were due to DNA polymerase infidelities causing insertions, deletions, or substitutions of usually one but sometimes more nucleotides. While most of these errors were not typed (i.e. not recognized as "true alleles" by Illumina's software), some were. Samples at higher DNA input showed more sequence errors, up to approximately eight for specific STR loci and more for certain Y- and X-STRs.

### 3.1.11 Errors in the STR Genotypes

All autosomal-, Y- and X-STR sequences, whether typed or untyped by the UAS software were used for data analysis. Because DNA from 16 individuals was repeatedly tested in the nine experimental runs, it was possible to accumulate sufficient data to detect and evaluate sequence inconsistencies such as genotype errors and sequence artifacts. Analysis assessed multiple types of errors including those flagged by Illumina's UAS software as well as those that were not flagged. Errors evaluated included: sequence errors, stutter, additional alleles, and drop-outs.

Analysis of the 314 reaction samples within nine experimental runs revealed 79 STR genotype errors of which 71 were not flagged; meaning that nothing was pointing towards them as being errors and so there was no way for an analyst to detect the errors without manually comparing them to a reference sample (only available for controlled samples and not actual casework samples). Only eight genotype errors were flagged. Most genotype errors occurred at lower DNA inputs and were allele drop-outs (**ADO**) that falsely resulted in typing heterozygotes as homozygotes. Additional genotype errors

occurred as a result of ADO plus typed stutter; locus drop-out (**LDO**) plus typed allele drop-in (**ADI**); or ADI at Y-STRs in female reaction samples. Five ADIs at Y-loci were deemed spurious alleles. The four ADIs in female samples were considered as editable. However, the allele drop-in with simultaneous locus drop-out for a male sample was detected by using reference samples and could not have been detected in an unknown sample.

### 3.1.12 Analysis of iSNP Genotypes

iSNP reports included read numbers for all genotypes detected. All iSNP reads, including low reads that were untyped, were used for evaluating genotypes. Two types of errors were identified: i) those which were flagged but could be interpreted (edited), and ii) those which were not flagged and thus, resulted in genotype miscalling. Some errors appeared locus specific. Locus drop-outs were also observed (both flagged and not flagged).

### 3.2 Mixed DNA Samples

Subsequently after analysis and characterization of single-sourced samples, mixed DNA samples were analyzed. DNA mixtures result in very complex samples due to the fact that all of the factors that played a part in single-sourced samples are now compounded for all of the contributors with the addition of the effects that are due to the amplification and sequencing of a mixed sample at the same time. The DNA of two male and two female samples was used to create the mixtures. The DNA of these individuals was also tested as single-source samples, which were used as reference for the analysis. The mixed DNA samples were prepared in ratios of 1:1 (for which designed mixtures included: M:M, F:F, and M:F) and 1:4 (for which designed mixtures included: MM 4:1, FF 4:1, MF 4:1, and MF 1:4). Data analysis focused on the number of contributors, their sex, and the ratio. Deconvolution was not performed on purpose, since most laboratories are using probabilistic software for their data analysis avoiding deconvolution by hand.

Additional mixtures of three, four, and six persons were performed. The results showed that all outcomes determined the number of contributors, their sex and ratio correctly.

### 3.3 Degraded and Challenging DNA Samples



Degraded DNA and challenging samples were assessed by comparing the results from Illumina's ForenSeq™ Mix A with the results from PowerPlex® Fusion. In order to generate several levels of degradation, DNA of good quality was boiled at 95°C for various times ranging from 5 to 60 min. The higher the degradation, the more loci dropped out. While only three loci dropped out at the highest level of degradation from Illumina's ForenSeq Kit A, 15 loci dropped out with PowerPlex® Fusion, revealing the advantage of the independency of the length of the amplicon of the NGS technology. In other words, the smaller amplicon size used in Illumina's ForenSeq Kit A compared to those in PowerPlex® Fusion created a big advantage for Illumina's ForenSeq Kit A when dealing with degraded samples (such as those from missing persons cases).

Further, challenging samples (DNA from blood cards and teeth which were several years old) were assessed for their level of degradation that was moderate. These four challenging samples were also tested with both Illumina's ForenSeq™ Mix A and PowerPlex® Fusion. As expected, Illumina ForenSeq™ Mix A handled the challenging samples better and showed drastically less dropouts compared to PowerPlex® Fusion.

### 3.4 Costs and Throughput

A comparison of costs and throughput was calculated for Illumina's ForenSeq™ and PowerPlex® Fusion followed by separation and detection on a Genetic Analyzed 3130xl. The price per locus for PowerPlex® Fusion is roughly twice as high as it is for Illumina's ForenSeq™ Kit A and three times higher than it is for Illumina's ForenSeq™ Kit B. In regards to run-time, the CE technique is approximately 4.5 times shorter than the MPS technique.

## 4. Implications for Further Research, Policy, and Practice

The results of this study pointed to the notion that there is an issue presented by the positive control (2800M). In some experimental runs 2800M experienced ADOs (even though the recommended 1ng input was used), meaning 2800M acted as a sample rather than a positive control, which was incorporated to help verifying that all of the primers were working well and that the experimental run as whole went well. Moreover, this study also pointed out a need for a probabilistic program that can handle and

resolve mixed DNA samples before MPS techniques can be implemented for forensic casework. Furthermore, this study suggests re-designing of Illumina's ForenSeq™ DNA Signature Prep Kit Primer Mix A would be beneficial for a more balanced outcome in terms of read numbers.

In the future, MPS will likely replace "current" applications of capillary electrophoresis. MPS of short tandem repeats facilitates a much more powerful discrimination based on sequence variants producing more accurate and just results than ever before. However, before incorporating new technologies into routine laboratory operations they must be rigorously evaluated based on many aspects of their performance. This study evaluated Illumina's FGx Forensic Genomic System and it uncovered some of the strengths and limitations of this system and thereby contributed to the development of a method that has the potential to become eventually the core of modern day forensics.

## 5. Dissemination of Research Findings

The findings from this research project resulted in one accepted manuscript for publication and two large oral presentations at forensic conferences as well as several in-house educational seminars for approximately 150 criminalists at the OCME. In addition to this, two additional publications are planned along with more presentations.

(3,755 words)

## **FINAL DRAFT**

### **1. Introduction**

#### **1.1 Statement of the problem**

The foundation of individual identification in modern forensic science is DNA typing by short tandem repeat (**STR**) analysis. This technique has brought a standardized, quantitative method with strong statistical underpinnings into the criminal justice system that has dramatically improved just and impartial outcomes. While the fundamental principles behind STR typing have not changed, new instrumentation and informative biological markers developed over the past few years have the potential to address the limitations of current techniques (e.g. mixture interpretation), as well as the need for increased throughput at lower costs, a problem that has developed with the continuing expanded use of this method.

Current DNA analysis methods for individual identification have technical, cost and throughput limitations. These methods are based on capillary electrophoretic sizing of a select group amplicons, some of which (e.g. CSF1PO, TPOX and TH01) would not be accepted today due to poor population frequency distributions or PCR amplification problems [1]. Capillary electrophoresis (**CE**) itself has limitations. It is a single reaction detection method without capabilities for multiplexing and is consequently slow. It is also a crude analog measure of amplicon length and consequently fails to detect informative sequence specific information that could significantly improve individual identification and mixture deconvolution. Finally, because of the inherent constraints of the CE method itself: i) the need for sufficient loci separation for adequate resolution [2] combined with ii) "limited band width" [3] (i.e. the limited number of amplicon lengths it is capable of separating), CE (with the expanded core) is approaching the maximum number of STRs it can process, effectively abandoning improvements from new autosomal, Y and X STRs as well as SNPs that can improve individual identification, mixture deconvolution and hereditary analysis - at least in a single system. The New York City Office of Chief Medical Examiner (**NYC OCME**) typically processes about 750 samples per month. The problem is that with ever increasing numbers of samples (especially property crimes) and the need for more informative genetic identity markers (better mixture deconvolution), CE may have reached its limit. Consequently, it is

necessary to begin validating the next generation of forensic DNA analysis platforms capable of higher throughput, and increased discriminative power, while still able to detect legacy STRs.

## 1.2 Literature, Citations and Review

The limitations of the current CE method can be broken down into three essential areas: technical, time and cost.

1) TECHNICAL - There is a need for additional autosomal, Y and X STRs (or SNPs) for better mixture deconvolution and unambiguous detection of male DNA in mixtures with high female DNA content [2], as well as to reduce the likelihood of a random match between unrelated individuals [2] and to improve the likelihood of detecting informative genetic markers in degraded samples through the use of multiple small (and perhaps equally sized) amplicons. Because STR detection by CE is based on amplicon length using a limited number of fluorescent dyes, it is not possible to detect similar sized alleles, nor is there sufficient chromatographic room to add many more alleles [4]. Consequently, forensic improvements requiring additional amplicons, as described above, are limited in CE.

2) TIME - Because of the limited number of dyes available for CE detection, each sample's STR profile must be detected individually, i.e. multiplexing of samples for simultaneous detection is not possible, thus data acquisition occurs in a relatively slow, linear fashion.

3) COST – Because multiplexing is not possible, multiple CE instruments must be purchased for high throughput sample processing.

Following the human genome project, the need for high throughput genomic sequencing for biomedical research and personalized medicine lead to new chemistries and instrumentation. Rather than Sanger sequencing and detection by capillary electrophoresis, massively parallel clonal sequencing was developed by several companies using different chemistries (e.g. sequencing by synthesis, pyrosequencing and ligation sequencing) and detection techniques [5]. These “next generation sequencing” (NGS) methods, also called “massively parallel sequencing” (MPS)

methods, function by clonal amplification of thousands of specific or random amplicons which are then sequenced simultaneously (in parallel). While the chemistries and detection methods of these instruments (Illumina, Roche, Life Technologies) differ, they effectively sequence large numbers of clonal populations generating hundreds to thousands of reads for each clone and consequently ensure sequence accuracy. Because nucleotide additions are detected as they occur, there is no need for electrophoresis, consequently massive amounts of DNA can be sequenced at relatively high speeds and low costs [2].

Use of NGS in forensic testing offers numerous advantages, not least of which is the ability to add new, informative STRs and/or SNPs capable of improving mixture deconvolution [6, 7] and kinship analysis [5] while maintaining legacy STRs. Part of this increased discriminative power is because amplicon length is no longer a limiting constraint as multiple informative loci of similar or identical length can be used [3], and with the discovery and inclusion of new, multiple, short informative STRs and/or SNPs [5], degraded samples that currently produce few STRs may become deducible. Finally, because of the ability to barcode individual samples, hundreds of samples can be multiplexed [6, 7]. This combined with the tremendous sequencing capacity of NGS systems (10 Mb to several Gb within hours or days) have the potential to significantly increase sample throughput and consequently reduce sample costs [2, 6].

The potential for NGS to improve forensic DNA analysis and consequently just legal outcomes is clear, and there have been a number of publications evaluating its use for STR analysis [7-9]. Earlier publications using Life Science's 454 and Illumina's GAIIx demonstrated promising results, although the 454 showed some difficulty with homopolymers stretches >4 – 6 bp, which is specific to 454 chemistry [3, 7]. More recently [7, 9] using the new Illumina MiSeq FGx forensic platform with Illumina's TruSeq library generation kit found correct allele calling, high signal to noise ratios, and low limits of detection for minor contributors in mixed samples despite the fact that their multiplex assay was not optimized for use in NGS [e.g. intra- and inter-locus balance, polymerase stutter (slippage) etc.] [7]. Importantly, Morling and colleagues [10-13] have demonstrated the strength of NGS not only to detect SNPs within STRs and their flanking sequences, but to resolve microvariants and off ladder alleles in complex STRs

that differed in length by as few as one nucleotide, demonstrating that NGS, even using current core loci, has the potential to significantly improve discriminative power.

However, question have been raised about the reliability of NGS for mitochondrial (and by implication genomic, [14]) sequencing [15-17] following reports of significantly higher frequencies of intra-individual mitochondrial heteroplasmy than previous studies showed [18, 19]. While much (but not all) of what appeared to be sequencing errors were subsequently shown to be contaminants introduced during sample preparation [17], questions emerged about the criteria selected for setting minor allele frequencies thresholds. Other potential problems for NGS use in forensics include: sequence quality for long CODIS loci (e.g. PentaE and PentaD, [7]); the need for better analytical and interpretation software, [4, 7, 20]; and the relatively large amount of DNA needed for analysis (1 ng). Equally as important as sequencing and software problems, however, are throughput and costs. Because of the massive amounts of sequencing NGS instruments have brought to biomedical research and personal medicine, as well as their ability to multiplex samples through barcoding (important to forensic), there is an expectation that these economies of scale will transfer seamlessly to forensics. However, at least at present, this may not occur. There are two reasons for this: i) Kit costs for library preparation and sequencing reagents are fixed, i.e. whether one or 384 samples are run, the costs is the same. Consequently, large runs will be required to achieve sample run costs comparable to CE. ii) The depth of sequencing, i.e. the number of times each amplicon sequence is read, determines the quality of that sequence. Currently, for single source samples Illumina recommends a maximum of 96 samples/run. For mixed samples (sexual assaults and non-exemplars which may or may not be single source) Illumina recommends only 32 samples/run. In both cases costs are the same and 3 days are required for clonal amplification and sequencing. However, over the past decade NGS sequencing costs and run times have dropped almost exponentially. If these trends continue, then the inherent advantages of NGS to forensics will be cost effective and meet necessary turnaround times.

### **1.3 Purpose of the Project**

The overall goal of this application is to evaluate Illumina's next generation forensic sequencing platform (**FGx**) for use in routine casework in the Department of Forensic Biology at the NYC OCME. Specific goals include: i) comparing next generation sequencing (**NGS**) detection sensitivity and specificity of core and expanded core loci against current ABI 3130 capillary electrophoresis platform, ii) evaluate NGS's ability for mixed DNA samples, iii) compare the sensitivity of NGS using degraded DNA samples against CE, and iv) compare throughput and costs between the two systems.

## **2. Methods**

### **2.1 Sample collection, DNA extraction and quantification**

This study was approved by the New York City Department of Health and Mental Hygiene's Institutional Review Board (IRB# 15-125). Samples were anonymized. Buccal swabs (Citmed Corporation, Citronelle, AL) were obtained with informed consent from 16 volunteers (nine males and seven females). DNA was extracted using an M48 BioRobot® (Qiagen, Valencia, CA) using their MagAttract® extraction kit following manufacturer's instructions as recently described [21]. Extracted DNA was quantified using Quantifiler® Trio (Thermo Fisher Scientific, Waltham, MA) following the manufacturer's instructions. A negative control was included with each extraction; if it tested positive, all samples in the batch were discarded.

### **2.2 CE concordance data: PowerPlex® Fusion and AmpFISTR® Yfiler®**

DNA from 16 individual samples was amplified using PowerPlex® Fusion 5C (Promega, Madison, WI) following manufacturer's instructions for half volume reactions of 12.5µl containing 7.5µl master mix and 5µl reaction sample, with a DNA input of 500pg. The PCR was performed using 29 cycles. In addition, 500pg of male DNA was amplified with AmpFISTR® Yfiler® PCR Amplification Kit by Life Technologies (Life Technologies Applied Biosystems, Foster City, CA) following manufacturer's instructions.

Separation of PCR products (1µl) was performed on the 3130x/ Genetic Analyzer (Life Technologies Applied Biosystems, Foster City, CA). Parameters for injection were 3kV for 5s and for separation 13kV for 2000s.

Data was analyzed using GeneMarker® (SoftGenetics, State College, PA) applying a 3% Global Filter. Local Southern was chosen for sizing, and the analytical threshold was set to 50 RFUs.

### **2.3 Experimental overview and ForenSeq™ DNA Signature Prep library preparation**

This project consisted of 16 experimental runs for the testing of various parameters: concordance, repeatability, sensitivity, allele coverage ratio, stutter, sequence variants, reliability, mixed DNA, degraded, and casework-type samples. Each experimental run contained reaction samples plus two controls, one positive (Illumina 2800M DNA, always at 1 ng input), and one negative (water). Library preparation was performed using the ForenSeq™ DNA Signature Prep kit (Illumina, San Diego, CA) following the manufacturer's instructions. The library prep workflow included the following steps: amplification of genomic targets, tagging of amplicons, and enrichment of targets, normalization, pooling, denaturing, and dilution of library. Experimental runs were performed on the Illumina MiSeq FGx system in the Forensic mode (Illumina) using the MiSeq FGx Reagent Kit. Illumina default settings were specifically chosen for all runs. Table 1 shows an overview of the preformed experimental runs.

### **2.4 Data analysis**

Data was extracted and analyzed using default conditions set by Illumina. Loci that were flagged “many alleles” include: stutter (referring to repeated units of STRs) or amplification or sequence errors. Some of these “typed” (Illumina nomenclature for called) alleles could be manually edited based on allele count, sequence comparison, and stutter threshold calculations to determine the genotype.



Incorrect STR length or genotype could be due to: Allele Drop Out (**ADO**), Locus Drop Out (**LDO**), Genotype-Error, or Allele Drop In (**ADI**).

Secondary data analyses (tables and charts) were generated in Excel.

### **3. Results**

#### **3.1 Evaluation of Illumina's MiSeq FGx Platform**

The evaluation of Illumina's MiSeq FGx Platform included tests and assessments of single-source samples of pristine quality for concordance with other platforms, sensitivity and flow cell capacity, repeatability, allele coverage ratio, stutter, and sequence variants (specific goal # 1).

##### *3.1.1 Concordance*

To characterize the ForenSeq™ DNA Signature Prep Kit, DNA samples of 16 individuals (nine male and seven female) were assessed with both the ForenSeq™ Primer Mix A and PowerPlex® Fusion. In addition, the nine male samples were tested with AmpFISTR® Yfiler®. In order to have a higher sample number, four of the 16 individual samples were tested with expired ForenSeq™ Signature kits (Table 1: Expt. 3 and 4, three male and one female samples; drop outs were seen for these samples only at PentaD). Notably, concordance was verified for all loci common between the different tests.

The additional loci of the ForenSeq™ DNA Signature Prep Kit, which were not assessed by the CE technique (5 aSTRs, 11 Y-STRs, and 7 X-STRs), were checked for reproducible outcomes in at least two experimental runs and/or in the various dilutions. Consistency was verified for all 16 individual samples at all STR loci, although allele and locus drop outs occurred at the following loci using expired ForenSeq™ kits (Expt. 3 and 4): DXS10103, DXS10135, DYF387S1, DYS385a-b, DYS389II, DYS390, and DYS448.

##### *3.1.2 Sensitivity*

The level of sensitivity of the ForenSeq™ DNA Signature Prep Kit was evaluated by reducing the DNA amount from 800pg to 50pg for six individual samples per run using

Primer Mix A. Based on concordance and reproducibility, samples showed 100% correct outcomes for the STRs at all DNA inputs including 50pg (Expt. 5). However, not each reaction sample resulted in a full profile. Figure 1 shows the average outcome of these six individual samples in percent at varying DNA amounts. For most inputs, the aSTRs showed a better outcome than the Y- and X-STRs, an effect that was more pronounced at lower concentrations. It is notable that the averaged outcome showed >90% of genotypes for aSTRs and Y- and X-STRs for all inputs, even at starting amounts as low as 50pg DNA.

Importantly, repetition of Expt. 5 as a technical replicate with newly made dilutions (Expt. 6) showed very similar outcomes in percent of full STR profiles, including 100% correct outcomes for one sample at 50pg DNA input (Fig. 1, bars marked in red), thus demonstrating repeatability.

### *3.1.3 Flow Cell Capacity*

To investigate the flow cell capacity, the same six samples at the same dilutions as in Expt. 5 and 6 were used with the ForenSeq™ DNA Signature Prep Kit Primer Mix B, thus increasing the multiplex reactions by an additional 78 targets (Table 1, Expt. 7). Figure 2 shows the average outcome of the six reaction samples in percent. The lowest DNA input amount of a reaction sample that resulted in full profiles for the aSTRs and the Y- and X-STRs was 200pg DNA. For the aSTRs, allele drop-out (**ADO**) started at 400pg DNA input (PentaD). More drop outs occurred at lower inputs. However, allele drop-ins (**ADIs**) were found for the Y-STRs at the higher DNA inputs of 400pg and 800pg, reaching >20% of the true allele (DYS390, DYF387S1, and DYS385a-b), due to sequencing problems and typed stutter, which were not editable. By comparing the average data of the six reaction samples generated from Primer Mixes A and B (compare Figures 1 with 2A), differences are notable, particularly for lower DNA inputs (50pg and 100pg), for which the reaction samples tested with Primer Mix B showed more drop outs. The loci that showed most drop outs were DXS10103, PentaD, D1S1656, DYF387S1, and DYS385a-b. These results show that the number of loci tested can have an effect on low DNA input samples.

In addition, the effect of the total amount of DNA in one run was assessed on samples with lower DNA input. Expt. 8 and 9 assessed four reaction samples with lower inputs (2

x 100pg and 2 x 50pg) loaded together with 26 samples, each at 1,000pg DNA, thereby increasing the total amount of DNA from 10,300pg (Expt. 5-7) to 27,300pg (Table 1). Figure 2B shows the average outcome of these two reaction samples in percent. ADO and locus drop-out (**LDO**) occurred more often for these samples than for the same amounts of DNA input (100pg and 50pg) in Expt. 5 and 6. Nevertheless, Expt. 8 and 9 resulted in similar outcomes, showing over 80% of the aSTRs and the Y- and X-STRs. The loci most affected by ADO and LDO were DXS10103, vWA, AMEL, TPOX, CSF1PO, and PentaE.

### 3.1.4 Allele Coverage Ratio (ACR)

The Allele Coverage Ratio (**ACR**), equivalent to CE peak height ratio (PHR), was determined for heterozygous loci by dividing the lower number of reads by the higher number of reads. Equal numbers of reads will result in ratios of 1 and increasing imbalance will result in smaller ratios.

The allele coverage ratios from Expt. 5 at all concentrations for each of the six individual samples were averaged and plotted per locus (Figure 3). Two loci on the Y-chromosome, DYF387S1 and DYS385a-b, are multicopy STR loci leading to two products. Most of the ACRs from the aSTRs were between 0.5 and 0.9, while the ACRs from the Y- and X-STRs scattered more as shown in Fig. 3. Most aSTRs at higher DNA inputs (800pg and 400pg) led to ACRs of >0.7, while at lower DNA inputs (100pg and 50pg) led to ACRs of <0.7. For D22S1045, the ACRs were noticeably low at high DNA inputs. At 800pg DNA input, the ACRs ranged from 0.39 to 0.81 for the six reaction samples. D22S1045 was often flagged as imbalanced and in most cases the longer alleles had fewer reads. In addition, the loci PentaD, D5S818, D1S1656, vWA, AMEL, and DYS385a-b also showed low ACRs.

The ACRs obtained from Expt. 6 were very similar, including the low ACRs for D22S1045, ranging from 0.55-0.91 (data not shown).

### 3.1.5 Stutter

Generally, the frequency of stutter, typed minus-one repeat unit, increased with decreasing DNA inputs, with Expt. 5 and 6 again showing very similar outcomes. For the aSTRs, up to three minus-one stutters were typed for the six individual samples at

higher DNA inputs (400pg and 800pg). For lower DNA inputs, the number of typed stutters increased to six for 100pg and to 12 for 50pg (typed minus-one stutters for Expt. 5 were represented as crosses above the bars in Fig. 3). Most typed stutters occurred at D21S11 ( $\leq 400$ pg DNA input) and at TH01 ( $\leq 200$ pg DNA input). For 50pg DNA input, Expt. 5 showed two stutters at D2S1338 that were  $>25\%$  of the true allele. Expt. 6 also showed two stutters at the same locus, one 22% and the other 42% of the true allele, and an additional one (44%) at D12S391 (data not shown).

Plus-one stutter was typed less frequently, and minus-two or plus-three stutter occurred rarely.

The Y- and X-STRs showed slightly more typed stutters at higher DNA input, while at lower input the number of typed stutters was comparable to the aSTRs. However, for higher DNA input (400pg and 800pg) some minus-four ( $n=4$ ) and minus-six ( $n=1$ ) stutter was typed at DYS505 and DYS576, respectively. At DYS505 (800pg), 18 alleles were listed that differed in length or by sequence, of which two were typed: the true allele and the minus-four stutter. The highest stutters were observed at DYS385a-b, DYF387S1, and DXS10135 (Figure 3).

### 3.1.6 Number of Reads: Balance

The read numbers from Expt. 10 for the true alleles were plotted as an average from the 30 reaction samples for each locus (Figure 4). As shown in Fig. 4, loci with high read numbers were consistently high and loci with low read numbers were consistently low. Average read numbers were determined for each locus. The lowest average read number for aSTRs was 501 (SD 105) for vWA and the highest was 7,628 (SD 1,050) for TH01, revealing a 15-fold difference. The lowest average read number for Y-STRs was 518 (SD 137) for DYS460 and the highest was 6,939 (SD 877) for DYS438, leading to a 13-fold difference. Overall, read numbers were slightly lower for X-STRs; the lowest was 146 (SD 72) for DXS10103 and the highest was 6,800 (SD 3,043) for DXS10074, resulting in a 47-fold difference. The read numbers for iSNPs were noticeably lower; the lowest was 48 (SD 11) for rs1736442 and the highest was 3,847 (SD 706) for rs1109037, resulting in an 80-fold difference. This indicates that each locus should be considered independently for stochastic and analytical thresholds.

### 3.1.7 Effects of Experimental Conditions on Number of Reads

In order to compare different experimental conditions, read numbers of all loci within an experimental run were averaged to an “experiment read number” (**ERN**, Table 1). This is reasonable within controlled experiments using the same samples because, while differences between loci within an experiment can vary widely, the relative relationships of loci to each other were consistent between experimental runs. Thus, a comparison of ERN offers some insight into the consequences of changing experimental conditions between runs. Table 2 compares ERNs for a variety of experimental runs testing different conditions. It should be emphasized that due to costs, some of these runs were performed only once. By comparing duplicate experimental runs (Table 2/Test 1), the fold-change difference of ERNs between replicate runs is 1.3, suggesting relatively good repeatability. This is in agreement with earlier outcomes (Fig. 1). Replicate runs separated by an 11 week interval (Table 2/Test 2) have an ERN fold-change of 2.1. Increasing sample numbers from 32 to 96 (Table 2/Test 3), while keeping DNA input (1 ng) and the number of loci (Kit A) constant, resulted in a reduced number of reads by about half in the 96 reaction sample experiment. Read numbers also declined, when an increased number of loci were sequenced (Kit A, 153 loci vs Kit B 231 loci) by keeping the number of reaction samples and DNA input constant (Table 2/Test 4).

For the experimental conditions within the recommendations of the manufacturer (Illumina) the variation of the average experimental read number varied approximately two fold.

However, reducing reaction sample DNA input by half, from 1 ng to 0.5 ng (Table 2/Test 5), while keeping other variables constant (32 samples and 153 loci, Kit A), resulted in a reduced number of reads by 3.5-fold for the 0.5 ng reaction samples. In Expt. 5 – 7 (Table 2/Test 6) sample DNA was present at: 800, 400, 200, 100, and 50 pg. In duplicate runs 5 and 6 (Kit A), the number of reads at 800 pg were 7.4 and 4.5 times greater than those at 50 pg respectively. In Expt. 7, using the same range of sample DNA inputs but with Kit B (increased targets), 800 pg reaction samples had a 6.8 fold increase in read numbers compared to 50 pg samples.

Perhaps most relevant with respect to forensic casework are changes in read numbers when the amounts of DNA were varied within a run. This is important in casework where multiple contributors may be present at different concentrations within a sample. It was

also observed that female samples within each run had more read numbers at the X-STRs than the male samples.

### 3.1.8 Positive and negative controls

Each experimental run contained one positive control, except for Expt. 12, which had 5 controls. As positive control was 2800M DNA (Illumina) used, always with 1000 pg DNA input. The positive controls were 100% positive in Expt. 2, Expt. 5-7, and Expt. 8. Some iSNPs were missed resulting in a lack of genotypes in Expt. 9 (5/94), Expt. 10 (3/94), Expt. 12 (2/94) and Expt. 16 (1/94) and some STRs were missed in Expt. 3 (STR: 2/59; iSNP: 0/94), and Expt. 14 (STR: 5/59; iSNP: 0/94). Some STRs and iSNPs were missed in Expt. 1 (STR: 1/59; iSNP: 1/94), Expt. 4 (STR: 8/59; iSNP: 8/172), Expt. 11 (STR: 3/59; iSNP: 36/94; pSNPs: 9/24; and aSNPs: 25/56), Expt.13 (STR: 4/59; iSNP: 3/94), and Expt. 15 (STR: 2/59; iSNP: 28/94). Therefore, the positive control failed to test all loci within the multiplex kit, meaning it behaved as a reaction sample. The affected loci were predominantly iSNPs, and Y- and X-STRs with low read numbers.

Each experimental run contained one negative control. Using default settings, coverage information of the negative no template controls (**NTCs**) for 15 of the 16 experimental runs showed no reads and were consequently flagged "low coverage" for all loci. Expt. 8 showed in the iSNP rs1493232 34 reads for A. These outcomes support the notion that no cross-contamination occurred throughout library preparation and sequencing steps.

### 3.1.9 Sequence Variants

A total of 16 individuals were tested and several sequence variants were identified. The type of sequence variants includes so-called "SNPSTRs" [22], which are SNPs within the STR repeats or flanking regions, or changes in the pattern of the repeats. Sequence variants were found at 14 loci, including D3S1358, D5S818, D8S1179, vWA, FGA, D13S317, D21S11, D12S391, D2S1338, D2S441, D9S1122, D4S2408, DYF387S1, and DXS10135 (Table 3). Most of these variations have previously been identified and were found in the NIST STR database (<http://www.cstl.nist.gov/biotech/strbase/>; [23]) or were published [24-26].

Variants found for D5S818 included a "SNPSTR" at the last nucleotide of the sequence: (AGAT)<sub>n</sub> and (AGAT)<sub>n-1</sub> AGAG, and for D12S391 a change of pattern for allele 21:

(AGAT)<sub>12</sub> (AGAC)<sub>9</sub>, (AGAT)<sub>12</sub> (AGAC)<sub>8</sub> (AGAT)<sub>1</sub>, and (AGAT)<sub>4</sub> (AGGT)<sub>1</sub> (AGAT)<sub>9</sub> (AGAC)<sub>6</sub> (AGAT)<sub>1</sub>.

### 3.1.10 PCR and/or Sequence Errors

Sequence variants, as described above, are known sequence differences between alleles. Sequence errors, as discussed here, represent a small number of reads (about 2% of the true allele) that differ from the dominant allele. While most of these errors were not typed (i.e. not recognized as “true alleles” by Illumina’s software), some were. Samples at higher DNA input showed more sequence errors, up to approximately eight for specific STR loci and more for certain Y- and X-STRs. These typed errors, allele drop in (ADI), occurred mostly at D7S820, while the non-typed sequence errors were seen at D19S433, D21S11 and D12S391. Interestingly, other loci showed no sequence errors (CSF1PO, D5S818, PentaE, D17S1301, DYS385a-b, DYS460, DYS522, and HPRT). These findings were consisted between Expt. 5 and 6.

### 3.1.11 Errors in the STR genotypes

All a-, Y- and X-STR sequences, whether typed or untyped by the UAS software were used for data analysis. Because DNA from 15 individuals was repeatedly tested in the nine experimental runs (Expt. 5-7 plus Expt. 10-15), it was possible to accumulate sufficient data to detect and evaluate sequence inconsistencies such as genotype errors and sequence artifacts. Analysis assessed multiple types of errors including those flagged by Illumina’s UAS software as well as those that were not flagged. Errors evaluated included: sequence errors, stutter, additional alleles, and drop-outs.

Analysis of the 314 reaction samples within nine experimental runs (Expt. 5-7 plus Expt. 10-15) revealed 79 STR genotype errors of which 71 were not flagged by the UAS software. Only eight genotype errors were flagged. Most genotype errors occurred at lower DNA inputs and were allele drop-outs (**ADO**) that falsely resulted in typing heterozygotes as homozygotes (n=67). Additional genotype errors occurred as a result of ADO plus typed stutter (n=7); locus drop-out (**LDO**) plus typed allele drop-in (**ADI**, n=1, Expt. 13, 12 reads for allele 7.2 at DYS448); or ADI at Y-STRs in female reaction samples (n=4, read numbers ranged from 13 to 60 at DYS392, DYS505, DYS576, and DYS643). These five ADIs were deemed spurious alleles for the following reasons: (i) none of the males tested in this study had allele 7.2 at DYS448, (ii) allele 16 at DYS576

was not present in any male sample in this experimental run (found in female B5F\_400 at 400pg DNA input in Expt. 6), (iii) the four female samples showed no other contamination with male DNA, and (iv) the two female reaction samples that showed reads at DYS505 and DYS643 had no male reaction samples adjacent on the 96 well-plate during library preparation. In conclusion, the four ADIs in the female samples were considered as editable. However, at DYS448 the 7.2 allele drop-in with simultaneous locus drop-out of allele 22 for the male sample B12M\_2 was detected here by using reference samples and could not have been detected in an unknown sample. Loci that showed the most STR genotype errors included PentaD (n=18), DXS10103 (n=13), DYS385a-b (n=10), and DXS10135 (n=6).

### *3.1.12 Analysis of iSNP genotypes*

iSNP reports include read numbers for all genotypes detected. All iSNP reads, including low reads that were untyped, were used for evaluating genotypes. Two types of errors were identified: i) those which were flagged but could be interpreted (edited), and ii) those which were not flagged and thus, resulted in genotype miscalling. Some errors appeared locus specific. Locus drop-outs were also observed.

In the nine experimental runs (Expt. 5-7 plus Expt. 10-15) testing 314 samples, 73 iSNP errors were found. As for the STRs, ADOs led to falsely typed homozygous genotypes that were not flagged by the UAS software. These errors affected 41 iSNPs (loci) at frequencies of six or less. The iSNPs containing the most genotype errors were rs914165 (n=6), rs6955448 (n=4), rs9905977 (n=4), and rs1493232 (n=4). Interestingly, loci rs914165 and rs6955448 also showed allelic imbalance in heterozygous samples (samples were flagged accordingly). Similarly, rs338882, which revealed two genotype errors, was also prone to allelic imbalance. It is not surprising that loci with allelic imbalance would also be subject to ADO, suggesting a re-evaluation of the use of these iSNPs. As with the STRs, the genotype errors for the iSNPs occurred more often at low DNA input (Expt. 5-7, and Expt. 14).

Overall, iSNP loci showed a broad range of read numbers (Figure 4). As expected, loci with the lowest read numbers had the most LDOs and ADOs, these included rs1736442, rs1031825, rs719366, rs1294331, rs7041158, rs1357617, rs2920816, rs338882, and



rs2342747. Predictably, at low DNA input the occurrence of LDO and ADO increased for these iSNPs (Expt. 5-7, Expt. 13, and Expt. 14).

In this study were all reads assessed (including those below 30 which were excluded by the UAS software) in order to evaluate iSNP data in detail. Genotypes marked with interpretation threshold "it" flags were most frequently typed falsely as homozygotes. However, by considering all reads, heterozygous genotypes could be determined which were in agreement with reference samples. Full read analysis also revealed that inconclusive genotypes (INC), which were flagged low coverage "lc" and showed reads for both alleles (heterozygotes), were always in agreement with reference samples. However, if ADO occurred, a false homozygote would be a consequence in a heterozygous sample. Even for experiments performed at recommended conditions as well as using Illumina's 2800M positive control DNA (Expt. 10, Expt. 15, Expt. 11, and Expt. 12), many genotypes could only be restored by manually editing. However, these were controlled and not unknowns samples.

### **3.2 Mixed DNA Samples**

Following the evaluation of single-source samples mixed DNA samples were assessed: i) two-person mixtures, ii) mixtures of more than two persons (specific goal # 2).

#### *3.2.1 Two-Person mixtures*

The DNA of two male and two female samples was used to create two-person mixtures. The DNA of these individuals was also tested separately as single-source sample, which was used as reference for analysis. The mixed DNA samples were used in ratios of 1:1 and 1:4. For the 1:1 DNA mixtures of two persons, 500 pg input was used from each individual sample and included mixtures of M:M, F:F, and M:F, thereby resulting to a DNA input per mixed sample of 1 ng. For the 1:4 DNA mixtures, 800 pg DNA was used for the major and 200 pg DNA for the minor contributor. The mixtures included MM 4:1, FF 4:1, MF 4:1, and MF 1:4. Data analysis focused on the number of contributors, their sex, and ratio. Deconvolution was not performed, since most laboratories are using a probabilistic software for data analysis thereby avoiding manual deconvolutions.

Table 4 shows the outcomes for the two person mixture study. By using the aSTRs, Y-STRs, X-STRs the outcomes for the 1:1 mixtures were accurate, while the ratios for the 1:4 mixtures varied between 1:3 to 1:6.

### *3.2.2 Mixtures of more than two persons*

Mixtures of more than two persons included three, four, and six persons, whereby all persons contributed equally and the total input was 1 ng DNA.

Table 5 shows the outcome of the mixed DNA study from more than two persons. The numbers of contributors and their sex were correctly determined for all mixtures. The ratios could be determined for mixtures consisting of four persons. For the six person mixture, the ratio could be determined only for the male contributors by using the Y-STRs.

Taken together, testing more loci plus additional Y- and X-STRs when compared to current CE techniques, such as PowerPlex® Fusion makes it easier to determine the number of contributors, their sex, and their contributing ratio.

## **3.3 Degraded and challenging DNA Samples**

Degraded DNA and challenging samples were assessed by comparing Illumina's ForenSeq™ Mix A to PowerPlex® Fusion (specific goal #3).

### *3.3.1 Degraded DNA*

The quality of DNA was determined by using Quantifiler™ Trio DNA Quantification kit that utilizes two PCR amplicons, a long (214 bp) and a short (80 bp) amplicon, to calculate their ratio, an equivalent to the degradation index (**DI**). The DI is a general indicator of whether large DNA fragments may perform more poorly relative to small DNA fragment in STR reactions. A high DI implies more DNA degradation: DI of <1 indicates no degradation; DI between 1 and 10 reveals slightly to moderately degradation, and DI >10 denotes extensively degradation.

DNA of pristine quality (DI < 1) was boiled at 95°C for various time points, 5 to 60 min, in order to generate several levels of degradation. After quantification the amount

recommended by the manufacturer was used with the Illumina ForenSeq™ Mix A (Expt. 16) and with PowerPlex® Fusion. Table 6 shows dropouts of the two kits at various DIs. The higher the DI the more loci dropped out. While only three loci dropped out at the highest DI (42.5) from Illumina's ForenSeq Kit A, it was five times as much with PowerPlex® Fusion, revealing an advantage of the independency of the length of the PCR amplicon.

### *3.3.2 Challenging samples: DNA from blood cards and teeth*

DNA from blood cards (n=2) and teeth (n=2), which were several years old were assessed for their DIs that ranged between 2.5 and 5.3. Again, these four samples were tested with Illumina ForenSeq™ Mix A and PowerPlex® Fusion. Table 7 shows the dropouts of these four samples. The dropouts of more loci at a moderate DI of 2.5 may be explained by the lower DNA input. As expected, Illumina ForenSeq™ Mix A handled the challenging samples better and showed drastically less dropouts compared to PowerPlex® Fusion.

## **3.4 Costs and Throughput**

A comparison of costs and throughput is given at Table 8 (specific goal #4) for Illumina's ForenSeq™ and PowerPlex® Fusion followed by separation and detection on a Genetic Analyzed 3130xl (Applied Biosystems, Thermo Fisher).

The costs per sample are \$ 84.11 for Illumina's ForenSeq™ Kit, which can test either 152 (Primer Mix A) or 228 (Primer Mix B) loci and would calculate to a price per locus of \$ 0.37 to \$ 0.55, depending which Kit would be used. While the price per locus for PowerPlex® Fusion was calculated to \$ 1.08, which is twice as high.

The time for the ForenSeq™ Kit library preparation was calculated to 11 h and running the MiSeq instrument for clonal amplification and sequencing adds up another 30 h, which resulted in 41 h for 32 samples (Table 8), while the CE techniques needed 9 h for 32 samples, which was 4.56 times shorter.

## 4. Conclusions

### 4.1 Discussion of Findings

This evaluation systematically assessed Illumina's ForenSeq™ DNA Signature Prep Kit in 16 experimental runs to gain more insight into this novel MPS platform (Table 1).

Important when studying novel methods is the concordance to currently used methods. This study used 16 individuals (nine males and seven females) to test concordance of Illumina's ForenSeq™ Kit A with PowerPlex® Fusion and with AmpFISTR® Yfiler® PCR Amplification Kit. Concordance was verified for the overlapping aSTRs and Y-STRs. These findings are in agreement with recent studies testing up to 15 samples [24, 27], thus increasing confidence in the platform due to increasing the numbers of samples tested.

Full STR profiles were consistently found for all six individual samples tested at DNA inputs of 400pg and 800pg. Full STR profiles could also be detected but at lower frequency using ForenSeq™ Primer Mix A at 50pg DNA input (consistent with other MPS platforms [28-31]), revealing that sensitivity for some loci was affected at this low DNA input.

Using the same individual samples and dilutions on more targets by testing the ForenSeq™ Primer Mix B, the results showed poorer outcomes (Fig. 2A). The lowest amount of input DNA used for a reaction sample that resulted in full STR profiles at all loci was 200pg DNA. These results are in agreement with those of Churchill et al. [24]. More importantly, differences in sensitivity (i.e. lowest level of detection) were not solely confined to the number of targets (Primer Mix A versus B). When the number of samples at low DNA input were small compared to the number of samples at high DNA input using Primer Mix A, sensitivity of the low DNA input samples declined, as shown for four reaction samples at inputs of 50pg and 100pg DNA that were tested together with 26 samples at 1,000pg DNA (Fig. 2B). This data suggests that the sensitivity of a sample depends on both the total number of target loci and on the relative amount of each sample's total DNA.

The ACRs found for the STRs in this study were similar to those described recently [24]. We showed, similar to CE, that ACRs were more balanced (>0.7) at higher DNA input

and less balanced (<0.7) at lower input [32]. However, ACRs for D22S1045 even at high DNA inputs were very low (0.57 on average for the six individual samples in run A), which was recently described [24].

Library preparation and sequencing include three amplification steps and one sequencing step all utilizing DNA polymerases: 1<sup>st</sup> amplification: targeted loci on the genomic DNA are amplified by PCR, 2<sup>nd</sup> amplification: addition of sample-specific indices by PCR, 3<sup>rd</sup> clonal bridge amplification, and 4<sup>th</sup> sequencing on the MiSeq instrument. These multiple amplifications may explain why uncommon stutters such as minus-six, minus-four, minus-two, and plus-one stutters were typed for some loci. More research will lead to a better understanding of these artifacts [33]. Overall, the most typed stutters were one repeat shorter (minus-one) and were found at D21S11, TH01, DYS385a-b, DYF387S1, and DXS10135.

Within reaction samples, read numbers for the various loci differed tremendously (Fig. 4). The average read number varied between locus-to-locus for aSTRs and Y-STRs over 10-fold, and for X-STRs and iSNPs over 40-fold. Similar data was reported from other laboratories [24, 27, 34]. Illumina's UAS software for most STRs, including all aSTRs, used default settings of >1.5% analytical and >4.5% interpretation thresholds, which can be adjusted for each locus. This is different from the current CE techniques where measured relative fluorescence unit (RFU) intensity does not differ significantly between loci and therefore the same thresholds are often used for all loci [35, 36].

Besides the locus-to-locus variation of read numbers within the multiplex reaction, it was shown that read numbers correlate with the DNA input. Low DNA input led to low read numbers (Table 2) and therefore resulted in higher frequencies of ADO/LDO, low ACRs, and genotype errors. Repeatability was demonstrated with the experimental runs 5 and 6 (Table 2), which were performed within a short period of time. On the other hand, repeated experimental runs performed 11 weeks apart (Expt. 10 and 15, unexpired kit) showed declined activity: lower read numbers, more drop-outs, more genotype errors, and lower ACRs. The QA data for Expt. 15 revealed lower cluster density and a lower Q30. It is unclear why the run performances differed so dramatically. The difference of nearly three months between kit usage might be a possibility (kit aging), but certainly is not confirmed by two experiments. Another confounding factor could be that the MiSeq instrument was serviced (i.e. the fluidics and optics system were calibrated) immediately

before Expt. 15 was conducted, but this was not the case for Expt. 10. These inconsistencies demonstrate the need for additional assessments in order to achieve more reliable outcomes.

This study showed that drop-outs occurred in several runs that were performed according to the manufacturer's instructions (Expt. 11, 12, and 15) as well as in positive controls (2800M). Therefore, the positive control failed to test all loci within the multiplex kit, meaning it behaved as a reaction sample. The affected loci were predominantly iSNPs, and Y- and X-STRs with low read numbers.

MPS has the advantage to provide more information than CE [37]. Sequence variants may aid in mixture deconvolution and stutter recognition. Testing 16 individuals revealed sequence variants at 14 loci. As most of the variants have been confirmed by other studies [23, 25, 38], the frequencies of these variants may suggest an improvement in discrimination despite the small sample size (n=16).

In addition to stutter, amplification or sequencing errors are also caused by DNA polymerases. Indeed, some errors typed by Illumina's software included insertions, deletions or substitutions of nucleotides and were considered as editable if their percentage of the true allele was <4% (D7S820) [21]. The sequence errors that were not typed (D19S433 and D21S11) showed changes of a nucleotide, but their frequencies were similar to the typed errors [21]. It is not clear why some errors were typed by the software, while others were not. Furthermore, some loci were more prone to sequence errors, while others did not have any, suggesting that the error may be sequence specific.

In order to detect genotype errors, 16 individual samples were repeatedly sequenced and used as reference. Within nine experimental runs (Expt. 5-7 and 10-15), 152 genotype errors were found, 79 affected STRs and 73 iSNPs. Most of the STR (67/79, 85%) as well as all of the iSNP genotype errors were due to ADO resulting in falsely typed homozygotes. The remaining 12 STR genotype errors occurred because of ADO plus typed stutter (n=7), LDO plus ADI (n=1), and ADI at Y-STRs in female samples (n=4). The majority of genotype errors occurred at lower DNA input. However, it should be noted that a few were also detected in experimental runs that were performed

following the manufacturer's recommended DNA input of 1 ng (Expt. 11, 12, and 15), as well as in the positive controls (2800M).

All reads were considered for iSNP data analysis and often allowed for correct genotyping after editing. iSNP loci flagged "it" could only be called correctly when sub-threshold reads were analyzed. iSNPs flagged "lc" could also be interpreted correctly if they were heterozygotes. However, when analyzing sub-threshold reads that appear homozygous, false genotypes were called in cases of ADO. As expected, the loci showing the lowest read numbers were the ones that needed the most editing. It should be pointed out first, that no wrong alleles were detected, besides five spurious alleles. Second, this could only be detected by testing reaction samples repeatedly and comparing the resulting genotypes, which would not be possible with unknown samples. And third, "it" and "lc" flags appeared frequently in experimental runs that were performed following the manufacturer's instructions, as well as in the positive controls (2800M). This data suggests a re-evaluation of the iSNPs should be considered before implementation in forensic casework.

Cross-contamination was assessed by the NTCs of all experimental runs, which showed no reads except for Expt. 8, where 34 reads were found in the iSNP ts1493232 for A. Furthermore, only five ADIs that were deemed spurious alleles were detected at Y-STR loci with read numbers ranging from 12 to 60. Four of these five ADIs were in female samples and one was in a male sample (S1 and S2 Tables). Two of these four ADIs were not present in the experimental setup (library preparation), including one that wasn't even detected among the individuals used for this study. Regarding the other two ADIs, the two female samples were not adjacent to male samples during library preparation. Since all four female samples showed no further contamination with male DNA these ADIs were considered editable. Similar apparent ADIs in female samples were also observed in another study [27]. On the other hand, the last ADI, which was found in a male sample, was a genotype error that could not have been detected without a reference sample. Taken together, these findings suggest that observed sequence errors were the result of DNA polymerase infidelity rather than contamination during the pooling that is part of MPS and cross-contamination did not occur.

The additional loci in Illumina's ForenSeq™ Kit compared to PowerPlex® Fusion Kit made it unproblematic to determine the number of contributors, their sex and ratio.

Deconvolutions by hand were not performed since this will also not be done if implemented into casework. A program is needed that can handle and resolve mixed DNA profiles, generated by NGS technology.

Testing of degraded and challenging DNA samples revealed the advantage of the NGS method by the occurrence of fewer dropouts from Illumina's ForenSeq™ Kit compared to PowerPlex® Fusion Kit [3]. This is in agreement with another study [24].

In summary, Illumina's ForenSeq™ platform was evaluated by concordance, sensitivity, allele coverage ratio, stutter, number of reads, cross contamination, sequence error, and genotype error. Concordance was verified. STRs and iSNPs from reaction samples with high DNA inputs had higher read numbers which led to more reliable results: fewer ADO/LDO and higher ACR. Therefore, the DNA input had the strongest effect on the outcome. Consequently, Illumina's ForenSeq DNA Signature Prep Kit Primer Mix A would benefit from a re-designing of primer and amplification conditions to achieve a more balanced outcome in terms of read numbers.

Genotype errors occurred primarily due to ADO/LDO and were only detected since samples were run repeatedly.

Essentially, no errors were found due to contamination caused by pooling or handling of the samples during library preparation. Spurious alleles were found on a few occasions: typed Y-alleles in four female samples (13-60 reads) that could be easily edited as well as, one ADI in a male sample at DYS448 (Expt. 13: 500pg, 12 reads) that was only found by comparing the genotype to repeated runs.

Running Illumina's FGx Forensic Genomic System under ideal conditions does not assure 100% outcomes neither for the reaction samples nor the positive control due to drop-outs and genotype errors. In addition, typed ADI and stutter were observed in all runs and required manual editing.

## **4.2 Implications for Further Research**

This study found that the positive control (2800M) behaved rather like a reaction sample, even with 1 ng DNA input ADO occurred and therefore primers for certain loci could not



be verified of functioning in these experimental runs. Therefore, it might be beneficial to discuss the roles of (positive) controls.

This study pointed out a need for a probabilistic program that can handle and resolve mixed DNA samples. This is important before MPS techniques can be implemented in forensic casework.

A result of this study was that re-designing Illumina's ForenSeq DNA Signature Prep Kit Primer Mix A would be beneficial in order to gain a more balanced outcome in terms of read numbers.

### **4.3 Implications for Policy and Practice**

In forensic testing, MPS will eventually replace current applications of capillary electrophoresis, because MPS of short tandem repeats contributes to a more powerful discrimination based on sequence variants and has the potential to be expanded to include additional targets. However, before incorporating new technologies into routine laboratory operations they must be evaluated for their performance. This study that evaluated Illumina's FGx Forensic Genomic System is important because it pointed to strengths and limitations of the system and thereby contributing to the development of a method that is capable of producing reliable and accurate data.

## 5. References

1. Butler JM. *Advanced Topics in Forensic DNA Typing: Methodology*. New Edition ed: Academic Press; 2012. 704 p.
2. Yang Y, Xie B, Yan J. Application of next-generation sequencing technology in forensic science. *Genomics Proteomics Bioinformatics*. 2014;12(5):190-7. Epub 2014/12/03. doi: S1672-0229(14)00105-3 [pii]  
10.1016/j.gpb.2014.09.001. PubMed PMID: 25462152; PubMed Central PMCID: PMC4411420.
3. Bornman DM, Hester ME, Schuetter JM, Kasoji MD, Minard-Smith A, Barden CA, et al. Short-read, high-throughput sequencing technology for STR genotyping. *Biotech Rapid Dispatches*. 2012;2012:1-6. Epub 2012/04/01. PubMed PMID: 25621315; PubMed Central PMCID: PMC4301848.
4. Budowle B. *Massively Parallel Sequencing and STR Typing: Basics and Forensic Applications* Forensic Magazine. 2014.
5. Berglund EC, Kiialainen A, Syvanen AC. Next-generation sequencing technologies and applications for human genetic history and forensics. *Investig Genet*. 2011;2:23. Epub 2011/11/26. doi: 2041-2223-2-23 [pii]  
10.1186/2041-2223-2-23. PubMed PMID: 22115430; PubMed Central PMCID: PMC3267688.
6. Parson W, Strobl C, Huber G, Zimmermann B, Gomes SM, Souto L, et al. Evaluation of next generation mtGenome sequencing using the Ion Torrent Personal Genome Machine (PGM). *Forensic Sci Int Genet*. 2013;7(5):543-9. Epub 2013/08/21. doi: S1872-4973(13)00137-3 [pii]  
10.1016/j.fsigen.2013.06.003. PubMed PMID: 23948325; PubMed Central PMCID: PMC3757157.
7. Van Neste C, Vandewoestyne M, Van Criekinge W, Deforce D, Van Nieuwerburgh F. My-Forensic-Loci-queries (MyFLq) framework for analysis of forensic STR data generated by massive parallel sequencing. *Forensic Sci Int Genet*. 2014;9:1-8. Epub 2014/02/18. doi: S1872-4973(13)00224-X [pii]  
10.1016/j.fsigen.2013.10.012. PubMed PMID: 24528572.
8. Fordyce SL, Avila-Arcos MC, Rockenbauer E, Borsting C, Frank-Hansen R, Petersen FT, et al. High-throughput sequencing of core STR loci for forensic

- genetic investigations using the Roche Genome Sequencer FLX platform. *BioTechniques*. 2011;51(2):127-33. Epub 2011/08/03. doi: 000113721 [pii] 10.2144/000113721. PubMed PMID: 21806557.
9. Van Neste C, Van Nieuwerburgh F, Van Hoofstat D, Deforce D. Forensic STR analysis using massive parallel sequencing. *Forensic Sci Int Genet*. 2012;6(6):810-8. Epub 2012/04/17. doi: S1872-4973(12)00079-8 [pii] 10.1016/j.fsigen.2012.03.004. PubMed PMID: 22503403.
  10. Dalsgaard S, Rockenbauer E, Buchard A, Mogensen HS, Frank-Hansen R, Borsting C, et al. Non-uniform phenotyping of D12S391 resolved by second generation sequencing. *Forensic Sci Int Genet*. 2014;8(1):195-9. Epub 2013/12/10. doi: S1872-4973(13)00206-8 [pii] 10.1016/j.fsigen.2013.09.008. PubMed PMID: 24315608.
  11. Dalsgaard S, Rockenbauer E, Gelardi C, Børsting C, Fordyce SL, Morling N. Characterization of mutations and sequence variations in complex STR loci by second generation sequencing. *Forensic Science International: Genetics Supplement Series*. 2013;4:e218-e9.
  12. Rockenbauer E, Hansen S, Andersen MR, Heerup EF, Børsting C, Fordyce SL, et al. Characterization of sequence variations in the D21S11 locus in Danes, Somalis and Greenlanders by second generation sequencing. *Forensic Science International: Genetics Supplement Series*. 2013;4:e302-e3.
  13. Rockenbauer E, Hansen S, Mikkelsen M, Borsting C, Morling N. Characterization of mutations and sequence variants in the D21S11 locus by next generation sequencing. *Forensic Sci Int Genet*. 2014;8(1):68-72. Epub 2013/12/10. doi: S1872-4973(13)00145-2 [pii] 10.1016/j.fsigen.2013.06.011. PubMed PMID: 24315591.
  14. Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, et al. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics*. 2012;13:341. Epub 2012/07/26. doi: 1471-2164-13-341 [pii] 10.1186/1471-2164-13-341. PubMed PMID: 22827831; PubMed Central PMCID: PMC3431227.

15. Bandelt HJ, Salas A. Current next generation sequencing technology may not meet forensic standards. *Forensic Sci Int Genet.* 2012;6(1):143-5. Epub 2011/05/14. doi: S1872-4973(11)00078-0 [pii] 10.1016/j.fsigen.2011.04.004. PubMed PMID: 21565569.
16. Irwin J, Just R, Scheible M, Loreille O. Assessing the potential of next generation sequencing technologies for missing persons identification efforts. *Forensic Science International: Genetics Supplement Series.* 2011;3:e447-e8.
17. Just RS, Irwin JA, Parson W. Questioning the prevalence and reliability of human mitochondrial DNA heteroplasmy from massively parallel sequencing data. *PNAS.* 2014;111(43):E4546-E7.
18. He Y, Wu J, Dressman DC, Iacobuzio-Donahue C, Markowitz SD, Velculescu VE, et al. Heteroplasmic mitochondrial DNA mutations in normal and tumour cells. *Nature.* 2010;464(7288):610-4. Epub 2010/03/05. doi: nature08802 [pii] 10.1038/nature08802. PubMed PMID: 20200521; PubMed Central PMCID: PMC3176451.
19. Ye K, Lu J, Ma F, Keinan A, Gu Z. Extensive pathogenicity of mitochondrial heteroplasmy in healthy human individuals. *Proc Natl Acad Sci U S A.* 2014;111(29):10654-9. Epub 2014/07/09. doi: 1403521111 [pii] 10.1073/pnas.1403521111. PubMed PMID: 25002485; PubMed Central PMCID: PMC4115537.
20. Warshauer DH, Lin D, Hari K, Jain R, Davis C, Larue B, et al. STRait Razor: a length-based forensic STR allele-calling tool for use with second generation sequencing data. *Forensic Sci Int Genet.* 2013;7(4):409-17. Epub 2013/06/19. doi: S1872-4973(13)00105-1 [pii] 10.1016/j.fsigen.2013.04.005. PubMed PMID: 23768312.
21. Almalki N, Chow HY, Sharma V, Hart K, Siegel D, Wurmbach E. Systematic assessment of the performance of illumina's MiSeq FGx forensic genomics system. *Electrophoresis.* 2017;38(6):846-54. Epub 2016/12/13. doi: 10.1002/elps.201600511. PubMed PMID: 27943350.
22. Mountain JL, Knight A, Jobin M, Gignoux C, Miller A, Lin AA, et al. SNPSTRs: empirically derived, rapidly typed, autosomal haplotypes for inference of population history and mutational processes. *Genome Res.* 2002;12(11):1766-

72. Epub 2002/11/08. doi: 10.1101/gr.238602. PubMed PMID: 12421764; PubMed Central PMCID: PMC187554.
23. Ruitberg CM, Reeder DJ, Butler JM. STRBase: a short tandem repeat DNA database for the human identity testing community. *Nucleic Acids Res.* 2001;29(1):320-2. Epub 2000/01/11. PubMed PMID: 11125125; PubMed Central PMCID: PMC29767.
24. Churchill JD, Schmedes SE, King JL, Budowle B. Evaluation of the Illumina((R)) Beta Version ForenSeq DNA Signature Prep Kit for use in genetic profiling. *Forensic Sci Int Genet.* 2016;20:20-9. Epub 2015/10/05. doi: S1872-4973(15)30071-5 [pii] 10.1016/j.fsigen.2015.09.009. PubMed PMID: 26433485.
25. Gettings KB, Aponte RA, Vallone PM, Butler JM. STR allele sequence variation: Current knowledge and future issues. *Forensic Sci Int Genet.* 2015;18:118-30. Epub 2015/07/23. doi: S1872-4973(15)30026-0 [pii] 10.1016/j.fsigen.2015.06.005. PubMed PMID: 26197946.
26. Parson W, Ballard D, Budowle B, Butler JM, Gettings KB, Gill P, et al. Massively parallel sequencing of forensic STRs: Considerations of the DNA commission of the International Society for Forensic Genetics (ISFG) on minimal nomenclature requirements. *Forensic Sci Int Genet.* 2016;22:54-63. Epub 2016/02/05. doi: S1872-4973(16)30009-6 [pii] 10.1016/j.fsigen.2016.01.009. PubMed PMID: 26844919.
27. Silvia AL, Shugarts N, Smith J. A preliminary assessment of the ForenSeq FGx System: next generation sequencing of an STR and SNP multiplex. *Int J Legal Med.* 2016. Epub 2016/10/28. doi: 10.1007/s00414-016-1457-6 10.1007/s00414-016-1457-6 [pii]. PubMed PMID: 27785563.
28. Elena S, Alessandro A, Ignazio C, Sharon W, Luigi R, Andrea B. Revealing the challenges of low template DNA analysis with the prototype Ion AmpliSeq Identity panel v2.3 on the PGM Sequencer. *Forensic Sci Int Genet.* 2016;22:25-36. Epub 2016/02/02. doi: S1872-4973(15)30045-4 [pii] 10.1016/j.fsigen.2015.07.011. PubMed PMID: 26828903.
29. Fordyce SL, Mogensen HS, Borsting C, Lagace RE, Chang CW, Rajagopalan N, et al. Second-generation sequencing of forensic STRs using the Ion Torrent HID

- STR 10-plex and the Ion PGM. *Forensic Sci Int Genet.* 2015;14:132-40. Epub 2014/12/03. doi: S1872-4973(14)00202-6 [pii]  
10.1016/j.fsigen.2014.09.020. PubMed PMID: 25450784.
30. Guo F, Zhou Y, Liu F, Yu J, Song H, Shen H, et al. Evaluation of the Early Access STR Kit v1 on the Ion Torrent PGM platform. *Forensic Sci Int Genet.* 2016;23:111-20. Epub 2016/04/16. doi: S1872-4973(16)30060-6 [pii]  
10.1016/j.fsigen.2016.04.004. PubMed PMID: 27082757.
31. Zeng X, King J, Hermanson S, Patel J, Storts DR, Budowle B. An evaluation of the PowerSeq Auto System: A multiplex short tandem repeat marker kit compatible with massively parallel sequencing. *Forensic Sci Int Genet.* 2015;19:172-9. Epub 2015/08/05. doi: S1872-4973(15)30049-1 [pii]  
10.1016/j.fsigen.2015.07.015. PubMed PMID: 26240968.
32. Marshall PL, King JL, Budowle B. Utility of amplification enhancers in low copy number DNA analysis. *Int J Legal Med.* 2015;129(1):43-52. Epub 2014/05/23. doi: 10.1007/s00414-014-1021-1. PubMed PMID: 24848516.
33. Aponte RA, Gettings KB, Duewer DL, Coble MD, Vallone PM. Sequence-based analysis of stutter at STR loci: Characterization and utility. *Forensic Science International: Genetics Supplement Series.* 2015;Series 5:e456-e8.
34. Just RS, Moreno LI, Smerick JB, Irwin JA. Performance and concordance of the ForenSeq system for autosomal and Y chromosome short tandem repeat sequencing of reference-type specimens. *Forensic Sci Int Genet.* 2017;28:1-9. Epub 2017/01/28. doi: S1872-4973(17)30001-7 [pii]  
10.1016/j.fsigen.2017.01.001. PubMed PMID: 28126691.
35. Bregu J, Conklin D, Coronado E, Terrill M, Cotton RW, Grgicak CM. Analytical thresholds and sensitivity: establishing RFU thresholds for forensic DNA analysis. *J Forensic Sci.* 2013;58(1):120-9. Epub 2012/11/08. doi: 10.1111/1556-4029.12008. PubMed PMID: 23130820.
36. Mulero JJ, Chang CW, Lagace RE, Wang DY, Bas JL, McMahon TP, et al. Development and validation of the AmpFISTR MiniFiler PCR Amplification Kit: a MiniSTR multiplex for the analysis of degraded and/or PCR inhibited DNA. *J Forensic Sci.* 2008;53(4):838-52. Epub 2008/06/11. doi: JFO760 [pii]  
10.1111/j.1556-4029.2008.00760.x. PubMed PMID: 18540972.

37. Borsting C, Morling N. Next generation sequencing and its applications in forensic genetics. *Forensic Sci Int Genet.* 2015;18:78-89. Epub 2015/02/24. doi: S1872-4973(15)00031-9 [pii]  
10.1016/j.fsigen.2015.02.002. PubMed PMID: 25704953.
38. Gettings KB, Kiesler KM, Faith SA, Montano E, Baker CH, Young BA, et al. Sequence variation of 22 autosomal STR loci detected by next generation sequencing. *Forensic Sci Int Genet.* 2015;21:15-21. Epub 2015/12/25. doi: S1872-4973(15)30091-0 [pii]  
10.1016/j.fsigen.2015.11.005. PubMed PMID: 26701720.

## 6. Disseminations of Research Findings

This project resulted in one accepted manuscript for publication and two additional publications are planned. Outcomes of this work were presented on two forensic meetings (oral presentations) as well as on several in-house educational seminar for approximate 150 criminalists at the OCME.

### Publication

- Almalki N, Chow HY, Sharma V, Hart K, Siegel D, Wurmbach E. (2016) "Systematic assessment of the performance of Illumina's MiSeq FGx™ Forensic Genomic System." *Electrophoresis* Dec 12; doi: 10.1002/elps.201600511.

### Publications planned:

- Qualitative and quantitative assessment of Illumina's forensic STR and SNP kits on MiSeq FGx™
- Assessment of mixed and degraded DNA utilizing Illumina's MiSeq FGx™ Forensic Genomic System

### Conferences:

- Green Mountain DNA Conference 2016: "Evaluation of next generation sequencing platforms for forensic casework" (oral presentation)
- Next Generation Dx Summit 2017: "Evaluation of Illumina's MiSeq FGx Forensic Genomic System" (oral presentation)



## **7. Tables and Figures**

### **7.1 Tables**

7.1.1 Table 1: Overview of experimental runs

7.1.2 Table 2: Effects of Experimental Conditions on Read Numbers

7.1.3 Table 3: Sequence variants

7.1.4 Table 4: Mixture Study: two-person mixtures

7.1.5 Table 5: Mixture Study: more than two-person mixtures

7.1.6 Table 6: Illumina and PP Fusion for one sample that was degraded by boiling

7.1.7 Table 7: Illumina and PP Fusion for degraded blood and teeth samples

7.1.8 Table 8: Costs and Throughput

### **7.2 Figures**

7.2.1 Figure 1: Sensitivity testing

7.2.2 Figure 2: Flow Cell Capacity

7.2.3 Figure 3: Allele Coverage Ratios (ACRs)

7.2.4 Figure 4: Number of Reads - Balance

**7.1.1 Table 1: Overview of experimental runs**

Run #	Comments	# of samples	Experiment: samples: DNA input	Total DNA [pg]	Primer Mix	ERN <sup>1</sup>
1	Concordance: M4-M6, F4-F6 Mixed DNA	32	Concordance : 3M and 3F: 1000pg, mixed DNA samples	N/A	Kit A (Lot#20140818) not expired	ND
2	Confirmation of sensitivity	32	Sensitivity: 3M and 3F: ~1000pg, ~500pg, ~200pg, ~100pg, ~50pg	12,100	Kit A (Lot#20140818) expired for one month	ND
3	Concordance: M7, M8 and F7	32	Concordance: 2M and 1F: ~500pg	N/A	Kit A (Lot#20140818) expired for four months	ND
4	Concordance: M9	32	Concordance: 1M: 1000pg	N/A	Kit B (Lot#20140815) expired for nine months	ND
5	Concordance and sensitivity: M1-M3, F1-F3	32	Concordance and sensitivity: 3M and 3F: 800pg, 400pg, 200pg, 100pg, 50pg	10,300	Kit A (Lot#3206C085) not expired	681
6	Repeat of Expt. 5	32	Concordance and sensitivity: 3M and 3F: 800pg, 400pg, 200pg, 100pg, 50pg	10,300	Kit A (Lot#3206C085) not expired	506
7	More targets	32	Same samples as in Expt. 5 and 6	10,300	Kit B (Lot#3206C086) not expired	471
8	Flow cell capacity	32	Sensitivity: 1M and 1F: 100pg, 50pg (26 samples:1000pg)	27,300	Kit A (Lot#3206C085) not expired	ND
9	Repeat of Expt. 8	32	Sensitivity: 1M and 1F: 100pg, 50pg (26 samples:1000pg)	27,300	Kit A (Lot#3206C085) not expired	ND
10	Standard run	32	Benchmark run: 1000pg	31,000	Kit A (Lot#3206C085) not expired	1589
11	More targets	32	Same samples as in Expt. 10	31,000	Kit B (Lot#3206C086) not expired	890
12	More samples	96	Each at 1000pg	95,000	Kit A (Lot#3206C085) not expired	774
13	Sensitivity	32	Same samples as in Expt. 10: 500pg	27,300	Kit A (Lot#3206C085) not expired	459
14	Sensitivity	16	Sensitivity test: 14 samples: 100pg	2,400	Kit A (Lot#3206C085) not expired	357
15	Repeatability	32	Repeat of Expt. 10: (11 weeks later)	31,000	Kit A (Lot#20140818) not expired	757
16	DNA degradation	32	Each at 1000pg	31,000	Kit A (Lot#2567C122) not expired	ND

<sup>1</sup>average Experiment Read Number (for all samples and loci)

ND: not determined

**7.1.2 Table 2: Effects of Experimental Conditions on Read Numbers**

<b>Test</b>	<b>Comparison of Experimental Runs<sup>1</sup></b>	<b>Fold-Change of Average Read Numbers of Correct Alleles</b>	<b>Results</b>
1. Experimental repeat	5 / 6	1.3	Good experimental replication.
2. Kit stability (testing 11 weeks apart)	10 / 15	2.1	Aged kit appeared to decline in activity.
3. Varying the numbers of samples – 32 vs. 96	10 / 12	2.1	Increasing the number of samples reduced read numbers.
4. Varying the numbers of targets <sup>3</sup> - Primer Mix A vs. Primer Mix B	10 / 11	1.8	Increasing the number of targets reduced read numbers.
	5 / 7	1.4	
	6 / 7	1.1	
5. Varying DNA input <sup>2</sup> between runs - 1ng vs. 500pg	10 / 13	3.5	Reducing DNA input reduced read numbers.
6. Varying DNA input <sup>2</sup> within a single run - 800, 400, 200, 100, and 50pg DNA. Read number comparisons only between 800 and 50pg.	5	7.4	Smaller amounts of DNA within the same run had lower read numbers.
	6	4.5	
	7	6.8	

<sup>1</sup>The ERNs of the respective experimental runs were used to calculate the fold-change.

<sup>2</sup>DNA input refers to the amount of sample DNA used at the library preparation stage.

<sup>3</sup>Primer Mix A has 152 loci; Primer Mix B has 228 loci

7.1.3 Table 3: Sequence variants

	Locus	Allele	Sequence
aSTRs	D3S1358	15	TCTA(TCTG)2(TCTA)12
			TCTA(TCTG)1(TCTA)13
		16	TCTA(TCTG)2(TCTA)13
			TCTA(TCTG)3(TCTA)12
		17	TCTA(TCTG)2(TCTA)14
			TCTA(TCTG)3(TCTA)13
	D5S818	10	(AGAT)11
			(AGAT)10 AGAG
		11	(AGAT)12
			(AGAT)11 AGAG
		12	(AGAT)13
			(AGAT)12 (AGAG)
		13	(AGAT)14
			(AGAT)13 AGAG
	D8S1179	13	(TCTA)1 (TCTG)1(TCTA)11
			(TCTA)13
		14	(TCTA)1(TCTG)1(TCTA)12
			(TCTA)2(TCTG)1(TCTA)11
			(TCTA)14
		15	(TCTA)1 (TCTG)1(TCTA)13
	(TCTA)2 (TCTG)1(TCTA)12		
	WWA	14	(TCTA)1(TCTG)1(TCTA)1(TCTG)4(TCTA)3(TCCA)1(TCTA)3(TCCA)2 (TCTA)1(TCTG)3(TCTA)10(TCCA)1(TCTA)1
	FGA	22	(TTTC)3 (TTTT) (TTCT) (CTTT)14 CTCC (TTCC)2 (TTTC)2 (TTTT)2 (TTCT) (CTTT)14 CTCC (TTCC)2
	D13S317	11	(TATC)12(AATC)1 ATC (TATC)2 (TTTC) (TGTC)2
			(TATC)11(AATC)2 ATC (TATC)2 (TTTC) (TGTC)2
		12	(TATC)13(AATC)1 ATC (TATC)2 (TTTC) (TGTC)2
			(TATC)12(AATC)2 ATC (TATC)2 (TTTC) (TGTC)2
		13	(TATC)14(AATC)1 ATC (TATC)2 (TTTC) (TGTC)2
			(TATC)13(AATC)2 ATC (TATC)2 (TTTC) (TGTC)2
		14	(TATC)14(AATC)2 ATC (TATC)2 (TTTC) (TGTC)2
			(TATC)15(AATC)1 ATC (TATC)2 (TTTC) (TGTC)2
	D21S11	29	(TCTA)4(TCTG)6(TCTA)3 TA(TCTA)3 TCA(TCTA)2 TCCA TA(TCTA)11
			(TCTA)5(TCTG)6(TCTA)3 TA(TCTA)3 TCA(TCTA)2 TCCA TA(TCTA)10
			(TCTA)6(TCTG)5(TCTA)3 TA(TCTA)3 TCA(TCTA)2 TCCA TA(TCTA)10
		30	(TCTA)2 TATA (TCTA)3 (TCTG)5 (TCTA)3 TA(TCTA)3 TCA(TCTA)2 TCCA TA(TCTA)11
			(TCTA)6(TCTG)5(TCTA)3 TA(TCTA)3 TCA(TCTA)2 TCCA TA(TCTA)11
			(TCTA)5(TCTG)6(TCTA)3 TA(TCTA)3 TCA(TCTA)2 TCCA TA(TCTA)11
			(TCTA)5(TCTG)5(TCTA)3 TA(TCTA)3 TCA(TCTA)2 TCCA TA(TCTA)12
			(TCTA)4(TCTG)6(TCTA)3 TA(TCTA)3 TCA(TCTA)2 TCCA TA(TCTA)12
	D12S391	18	(AGAT)11(AGAC)6(AGAT)1
			(AGAT)10(AGAC)7(AGAT)1
		19	(AGAT)12(AGAC)6(AGAT)1
			(AGAT)11(AGAC)7(AGAT)1
		20	(AGAT)12(AGAC)7(AGAT)1
			(AGAT)11(AGAC)9
(AGAT)13(AGAC)6 (AGAT)1			
21		(AGAT)12(AGAC)9	
	(AGAT)12(AGAC)8 (AGAT)1		
	(AGAT)4 (AGGT)1 (AGAT)9(AGAC)6 (AGAT)1		
D2S1338	19	(TGCC)7(TTCC)12	
		(TGCC)6 (TTCC)13	
	20	(TGCC)7(TTCC)13	
		(TGCC)7 (TTCC)10 GTCC (TTCC)2	
21	(TGCC)7(TTCC)14		

2015-DN-BX-K005  
Validation of Illumina's MiSeq FGx Platform for Transition to Casework

		23	(TGCC)7 (TTCC)11 GTCC (TTCC)2
			(TGCC)7(TTCC)13 GTCC (TTCC)2
			(TGCC)6 (TTCC)14 GTCC (TTCC)2
		24	(TGCC)7(TTCC)14 GTCC (TTCC)2
			(TGCC)6(TTCC)15 GTCC (TTCC)2
	D2S441	10	(TCTA)10
	(TCTA)8 TCTG (TCTA)1		
	D9S1122	11	(TAGA)11
			(TAGA)1 TCGA (TAGA)9
		12	(TAGA)12
			(TAGA)1 TCGA (TAGA)10
		14	(TAGA)14
	(TAGA)1 TCGA (TAGA)12		
D4S2408	9	(ATCT)9	
(ATCT)1 (GTCT)1 (ATCT)7			
Y, X-STRs	DYF387S1	37	(AAAG)3(GTAG)1(GAAG)4(AAAG)2(GAAG)1(AAAG)2(GAAG)10(AAAG)14
			(AAAG)3(GTAG)1(GAAG)4(AAAG)2(GAAG)1(AAAG)2(GAAG)9(AAAG)15
	DXS10135	22	(AAGA)3 GAAAG(GAAA)19 AGAGAATAGAAAAGAAGA...
			(AAGA)3 GAAAG(GAAA)15 (GGAA)3 (GAAA)1AGAGAATAGAAAAGAAGA...
		26	(AAGA)3GAAAG(GAAA)21(GGAA)(GAAA)1AGAGAATAGAAAAGAAGA...
			(AAGA)3 GAAAG(GAAA)23 AGAGAATAGAAAAGAAGA...

**7.1.4 Table 4: Mixture Study: two-person mixtures**

AMEL		No. of Contributors	No. of contributors and Sex	Ratio	Intended DNA mixture	
Ratio (Y/X)	Sex				No. of contributors and Sex	Ratio
1.9	M	2	2M	1:1	2M	1:1
-	F	2	2F	1:1	2F	1:1
0.47	MF	2	1M+1F	1:1	1M+1F	1:1
1.6	M	2	2M	3:1 or 4:1	2M	4:1
-	F	2	2F	4:1 or 5:1	2F	4:1
0.52	MF	2	1M+1F (M>F)	3:1	1M+1F	4:1
0.18	MF	2	1M+1F (F>M)	4:1 or 6:1	1F+1M	4:1

**7.1.5 Table 5: Mixture Study: more than two-person mixtures**

AMEL		No. of Contributors	No. of contributor and Sex	Ratio	Intended DNA mixture	
Ratio (Y/X)	Sex				No. of contributors and Sex	Ratio
1.0	M	3	2M+1F (M>F)	1:1:1	2M+1F	1:1:1
0.22	MF	3	2F+1M (F>M)	1:1:1	2F+1M	1:1:1
0.46	MF	4	2M+2F	1:1:1:1	2M+2F	1:1:1:1
0.42	MF	6	3M+3F	3M (1:1:1) 3F (?)	3M+3F	1:1:1:1:1:1

**7.1.6 Table 6: Illumina and PP Fusion for one sample that was degraded by boiling**

Boiling Time (mins)	DI <sup>a</sup>	Illumina drop-outs		PowerPlex Fusion drop-outs				
		Loci	No. of total drop-outs per sample	Yellow	Black	Green	Red	No. of total drop-outs per sample
0	0.8	D22S1045	1	-	-	-	-	-
5	0.9	-	0	-	-	-	-	-
10	1.3	-	0	-	-	-	-	-
15	1.7	-	0	-	-	Penta D	D22S1045	2
20	2.5	-	0	TPOX	D13S317, Penta E	Penta D	D22S1045	5
30	4.1	Penta E	1	D7S820, D5S818, TPOX	D13S317, Penta E	CSF1PO Penta D	D22S1045	8
40	8.7	-	0	D7S820, D5S818, TPOX	D13S317, Penta E	D2S1338 CSF1PO Penta D	FGA, D22S1045	10
50	17.6	D22S1045, Penta E	2	D7S820, D5S818, TPOX	D2S441, D10S1248, D13S317, Penta E	D2S1338, CSF1PO, Penta D	D19S433, FGA, D22S1045	13
60	42.5	D22S1045, D12S391, Penta E	3	D21S11, D7S820, D5S818, TPOX	D1S1656, D2S441, D10S1248, D13S317, Penta E	D2S1338, CSF1PO, Penta D	D19S433, FGA, D22S1045	15

<sup>a</sup>DI: Degradation Index

**Illumina:** D22S1045 (193-229bp), D12S391 (237-281), Penta E (362-467bp)

**PP Fusion: Yellow:** D21S11 (198-266.5), D7S820 (267-316), D5S818 (316.5-379), TPOX (387-443.5)

**Black:** D1S1656 (151-207), D2S441 (207.5-247.5), D10S1248 (248-295), D13S317 (295.2-350), Penta E (354.5- 474.9)

**Green:** D2S1338 (218-299), CSF1PO (309-363), Penta D (370-461)

**Red:** D19S433 (191.5-256.5), FGA (257-415), D22S1045 (420-472)



**7.1.7 Table 7: Illumina and PP Fusion for degraded blood and teeth samples**

Boiling Time (mins)	DI <sup>a</sup>	Illumina drop-outs		PowerPlex Fusion drop-outs				
		Loci	No. of total drop-outs per sample	Yellow	Black	Green	Red	No. of total drop-outs per sample
Blood card 1	2.5	-	0	-	D13S317 Penta E	Penta D	D22S1045	4
Blood card 2	5.3	-	0	-	Penta E	CSF1PO	D22S1045	3
Tooth sample 1 <sup>b</sup>	2.5	D22S1045, D12S391, Penta D Penta E	4	D7S820, D5S818, TPOX	D10S1248, D13S317, Penta E	CSF1PO, Penta D	D22S1045	9
Tooth sample 2	4.6	-	0	-	D13S317	CSF1PO	-	2

<sup>a</sup>DI: Degradation Index

<sup>b</sup>: sample concentration was 100pg

**Illumina:** D22S1045 (193-229bp), D12S391 (237-281), PentaD (209-293), PentaE (362-467bp)

**PP Fusion: Yellow:** D21S11 (198-266.5), D7S820 (267-316), D5S818 (316.5-379), TPOX (387-443.5)

**Black:** D1S1656 (151-207), D2S441 (207.5-247.5), D10S1248 (248-295), D13S317 (295.2-350), PentaE (354.5-474.9)

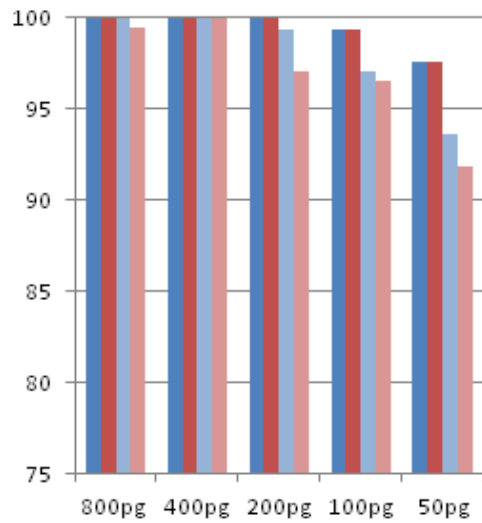
**Green:** D2S1338 (218-299), CSF1PO (309-363), PentaD (370-461)

**Red:** D19S433 (191.5-256.5), FGA (257-415), D22S1045 (420-472)

**7.1.8 Table 8: Costs and Throughput**

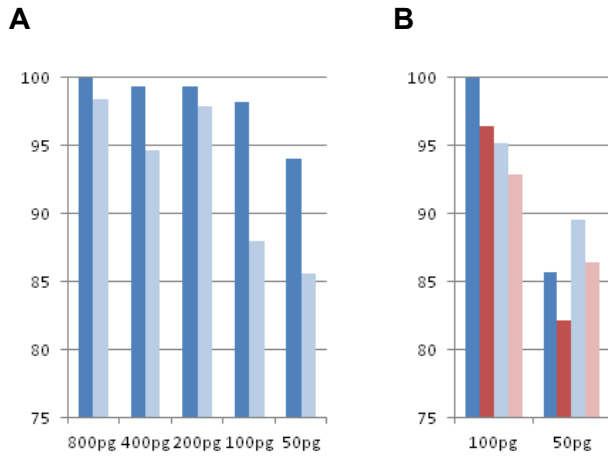
<b>Illumina's MiSeq FGx Platform (Kit A: 152 Loci and Kit B: 228 Loci)</b>		
<b>Kit Specification</b>	<b>Kit Price</b>	<b>Cost per sample</b>
ForenSeq DNA Signature Prep Kit	\$ 15,500 for 384 reactions	\$ 40.36
MiSeq FGx Reagents Kits	\$ 1,400 for 32 reactions	\$ 43.75
<b>Total</b>		<b><u>\$ 84.11</u></b>
<b>Preparation Time:</b>		
Step	Specification	Time (Hour)
1	PCR 1 Prep	1
2	PCR 1	3.5
3	PCR 2 Prep	1
4	PCR 2	1.5
5	Purification	1
6	Normalization	1
7	Denaturation and Pooling	1
8	Instrumental Run	30
<b>Total</b>		<b><u>41 hours per 32 samples</u></b>
<b>PowerPlex Fusion (24 loci)</b>		
<b>Kit Specification</b>	<b>Kit Price</b>	<b>Cost per sample</b>
Co-Amplification and Fluorescent Detection of 24 Loci	\$ 4,651 for 200 reactions	\$ 23.25
POP – Polymer for Genetic Analyzer	\$ 588 up to 250 samples	\$ 2.35
HiDi Formamide	\$ 120 up to 300 reactions	\$ 0.40
<b>Total</b>		<b><u>\$ 26.00</u></b>
<b>Preparation Time:</b>		
Step	Specification	Time (Hour)
1	PCR prep	1
2	PCR	1.5
3	CE prep	a
4	Capillary Electrophoresis	1
<b>Total</b>		<b><u>4.5 hours per 16 samples</u></b>

### 7.2.1 Figure 1



**Sensitivity testing:** The average outcome of Illumina's ForenSeq Primer Mix A for six samples (3 male and 3 female) is shown in percent (y-axis) for given DNA input (x-axis). Expt. 5 is shown in blue and Expt. 6 in red (same samples, freshly diluted). The darker shades represent the outcome for aSTRs and the lighter shades for the X- and Y-STRs.

### 7.2.2 Figure 2



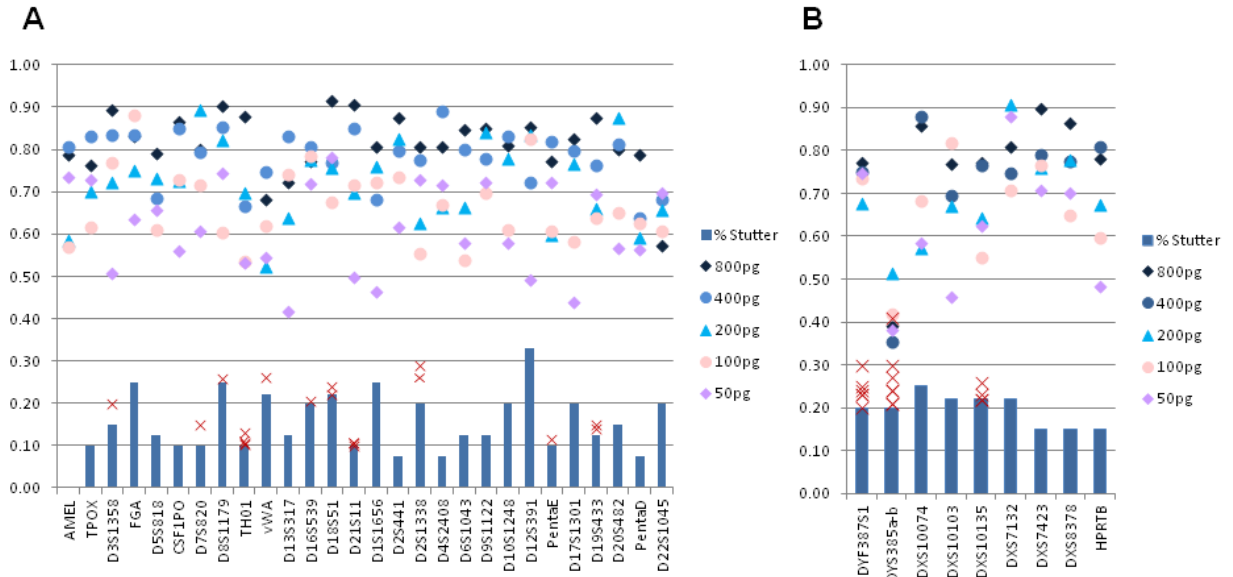
#### Flow Cell Capacity:

**A)** Illumina's ForenSeq™ DNA Signature Prep Kit B (Expt. 7): the average outcome of the six samples (same as in Expt. 5 and 6) is shown in percent.

**B)** Four samples (2 male and 2 female) were run with additional 26 samples at approximately 1,000pg DNA input plus controls (Expt. 8), average outcome is shown in blue. The repeated run (Expt. 9) is shown in red (same samples, freshly diluted).

X-axis: DNA input per sample; Y-axis: Correct outcome in percent.

7.2.3 Figure 3

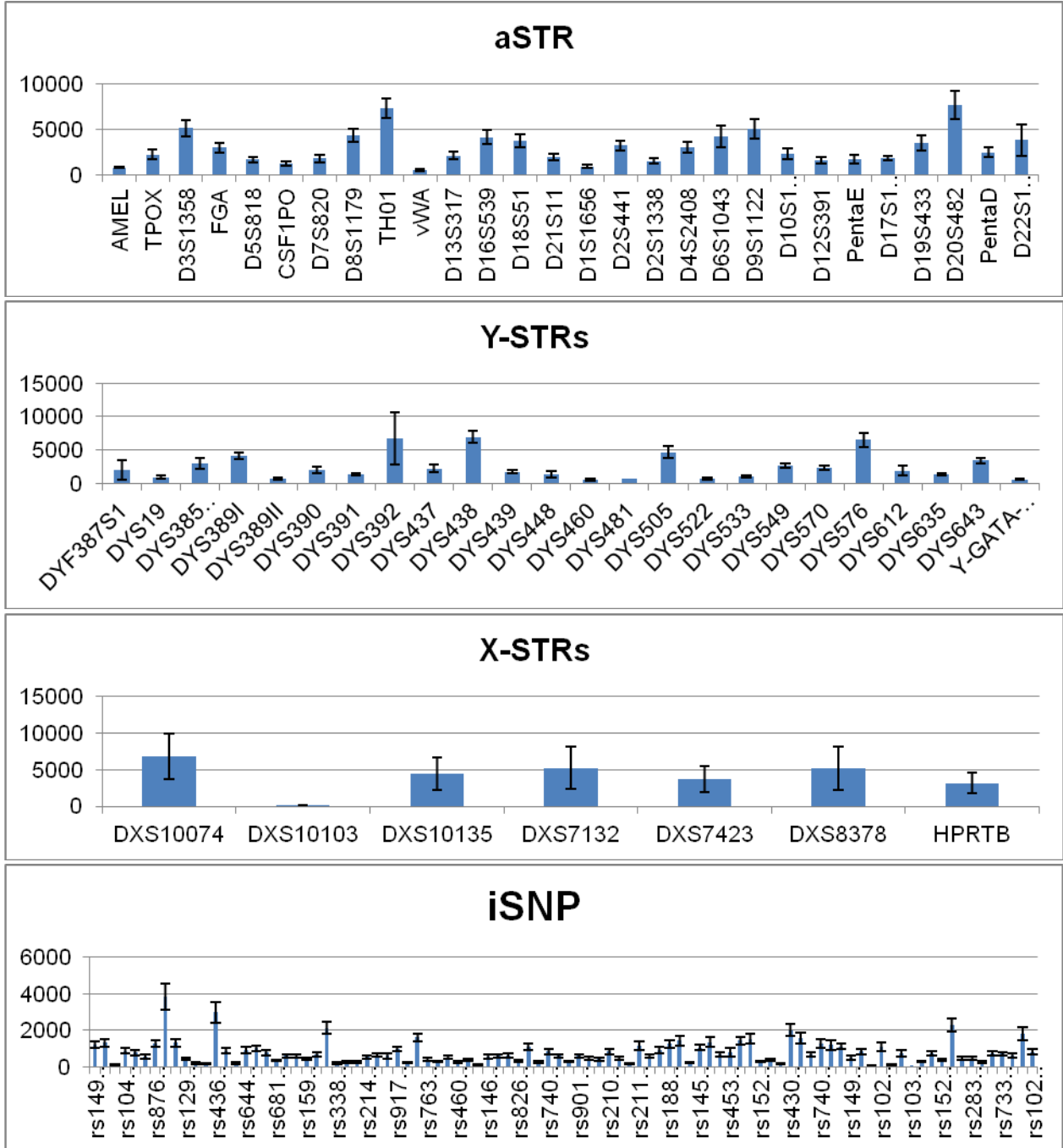


The Allele Coverage Ratios (ACRs) are shown for Expt. 5, as average from the six samples for each locus and DNA input [pg]. The bars for each locus indicate the stutter filter (for percent, multiply value on Y-axis by 100). The crosses over some bars show typed stutter for Expt. 5. The STR loci are shown on the X-axis.

A) ACRs and stutter for the aSTRs

B) ACR and stutter for the Y- and X-STRs (DYF387S1 and DYS385a-b show two products per male).

7.2.4 Figure 4



**Number of Reads – Balance:** Shown are the average read numbers from Expt. 10 with their standard deviations (SD) for aSTRs, Y-STRs, X-STRs and iSNPs.