Internal Validation of the Applied Biosystems® GlobalFiler™ Express PCR Amplification Kit

---

Shanna K. Saunders, B. S., Graduate Student, Marshall University Forensic Science Center, 1401 Forensic Science Drive, Huntington, WV 25701

Agency Supervisor and Reviewer: Kyra Groeblinghoff, DNA Technical Leader
St. Louis County Police Crime Laboratory, 111 South Meramec, Clayton, MO 63105

Technical Assistant and Reviewer: Season Seferyn, M.S.F.S., Parentage DNA Analyst
Marshall University Forensic Science Center, 1401 Forensic Science Drive, Huntington, WV 25701

MU Topic Advisor and Reviewer: Dr. Pamela Staton, Ph.D. Graduate Program Coordinator
Marshall University Forensic Science Center, 1401 Forensic Science Drive, Huntington, WV 25701

Abstract

The conventional process for obtaining DNA profiles from reference standards includes extraction and quantitation steps which add cost and time to the workflow. Advancements in buffers and amplification kits allow these steps to be modified or removed which can decrease costs and increase time savings allowing these resources to be redirected to casework processing. The Applied Biosystems® GlobalFiler™ Express DNA Amplification Kit utilizes these advancements and is a novel kit for the St. Louis County Police Crime Laboratory. Based on the most recently published FBI Quality Assurance Standards (Sept. 2011), any novel PCR amplification chemistry must be internally validated in order to ensure the reliability of data for use in a crime laboratory setting (10). The purpose of this study was to conduct an internal validation of the Applied Biosystems® GlobalFiler™ Express DNA Amplification Kit on the Applied Biosystems® 3500 Genetic Analyzer for use at the St. Louis County Police Crime Laboratory. During this validation, sensitivity, threshold, contamination, reproducibility, concordance, and mixture studies were completed to verify that the kit would produce reliable and reproducible results. Critical parameters were also tested and selected in order to optimize the protocol for the laboratory. Ideally the protocol would optimize the first pass rate. The first pass rate refers to obtaining a full profile the first time samples are taken through the DNA workflow. Overall, this validation demonstrated that the Applied Biosystems® GlobalFiler™ Express DNA Amplification Kit would produce consistent and reliable results. The National DNA Index System (NDIS) is currently evaluating the kit. Pending approval by NDIS, the Applied Biosystems® GlobalFiler™ Express Kit will be put into use by the St. Louis County Police Crime Laboratory. Future validations will take place to make the Applied Biosystems® GlobalFiler™ DNA Amplification Kit an option as well.

Introduction

The quality of a DNA Amplification Kit directly affects the recovery of DNA typing results which makes it critical for these kits to be evaluated before being implemented. The St. Louis County Police Crime Laboratory currently uses the Applied Biosystems® AmpFLSTR® Identifiler® PCR Amplification Kit (Life Technologies™, Foster City, CA) which includes 15 loci and Amelogenin. The FBI's Combined DNA Index System (CODIS) Core Loci Working Group recommendations are changing from 13 core CODIS loci to 20 required and 3 recommended loci, meaning a new kit will be needed (8). The GlobalFiler™ PCR Amplification Kits (Life Technologies™, Foster City, CA) include a 6-dye assay that targets 21 autosomal STR loci, one Y STR locus, one Y insertion/deletion locus (Yindel), and the sex-determining marker, Amelogenin, for a total of 24 loci (4). These kits satisfy the new recommendations being set out by the FBI as well as having many advantages over other kits. The St. Louis County Police Crime (SLCPD) Laboratory has chosen this kit for a number of reasons.

The 6-dye chemistry allows for additional spacing between loci to increase resolution as well as allowing more mini short tandem repeats (STRs) to be included in the kit. Ten mini STR loci are included and 97% of the alleles fall under the 400 base pair mark meaning better results for degraded samples (14). The previously used Penta loci were not listed as recommended CODIS core loci and are not included in the GlobalFiler™ Kits. Other amplification kits on the market include these loci leaving less room for recommended loci. A Y marker (Yindel) is also included in the kit which not only verifies the sex of the person who gave the sample but helps to detect Y deletions that give a false homozygote (i.e. female) at the Amelogenin locus.

The laboratory chose to not only upgrade to the 24 loci, 6-dye amplification kit, but they also decided to implement the GlobalFiler™ Express Kit, a direct amplification kit, for known samples. Because it is a direct amplification kit, the GlobalFiler™ Express Kit will decrease the amount of time and money the lab is putting into single-source samples. Direct amplification kits allow for the omission of the extraction and quantification steps and focuses directly on the amplification step. Instead of an extraction step, the GlobalFiler™ Express Kit uses a proprietary buffer called the PrepNGo Buffer that lyses the cells in the same tube that the sample is pulled from to set up the amplification step. Cell lysis by the PrepNGo Buffer takes approximately twenty minutes.

Not only do direct amplification kits eliminate steps in the workflow but they also have shorter amplification times to decrease overall laboratory time. This particular kit uses a thermal profile that is complete in approximately 23 to 24 minutes and a capillary electrophoresis (CE) run time of approximately forty five minutes per injection (1). A non-direct kit normally has an amplification time of approximately three hours and a run time of approximately 50 minutes per injection plus an hour extraction and 2 hour quantitation. Therefore, the time difference between direct and non-direct kits is approximately five hours.

The Applied Biosystems® GlobalFiler™ Express Kit is also cost effective. Previously, the St. Louis County Crime Laboratory used the AmpFLSTR® Identifiler® Amplification kit which required an extraction and quantitation step. Under these conditions, the total cost of one sample is approximately $185 from extraction to analysis. With the Express kit, an extraction and quantitation step are not performed and the total cost of a sample is approximately $105 from sample preparation to analysis. These calculations do not include analyst salary. The PrepNGo™ Buffer used with the Express kit is $500 per bottle but since the validation was done

with a significantly smaller volume than recommended, the extra cost did not attribute much to the overall price per sample.

With the previously used AmpFLSTR® Identifiler® Amplification Kit, the cost of an amplification kit was $3,379.81 whereas a GlobalFiler Express Kit is $3,880.00 (both have 200 reactions). Even though the GlobalFiler Express Kit is slightly more expensive, the cost savings in extraction and quantitation supplies, particularly the quantitation controls, makes up for the additional cost of the direct amplification kit. The lab is transitioning to exclusive use of the Applied Biosystems® 3500 Genetic Analyzer with the AmpFLSTR® Identifiler® Amplification Kit until the new kits and loci are validated with NDIS approval. Therefore, additional costs associated with the 3500 Genetic Analyzer reagents are not considered.

Single-source samples are usually of good quality and do not need to be reamplified or rerun to produce a full profile. With the abundance of these types of samples and the lack of need to quantitate them, per Standard 9.4 of the FBI's Quality Assurance Standards (6), the decrease in time and cost for the GlobalFiler™ Express Amplification Kit is optimal for these types of samples.

Because of the benefits of the Applied Biosystems® GlobalFiler™ Express Kit, an internal validation was performed at the St. Louis County Police Crime Lab to allow the GlobalFiler™ Express Kit to be implemented into the work flow for single-source samples. Sample preparation, injection time, analytical threshold, and stochastic threshold were determined to optimize the protocol for analysis of FTA cards and buccal swabs. Other studies including sensitivity, precision, concordance, reproducibility, contamination, and internal stutter were completed to ensure that the kit would produce reliable profiles on a variety of common

sample types. An additional population study was conducted to determine if the local population allele frequencies varied significantly from the national population.

NDIS is currently reviewing internal validation data from validation laboratories to determine if the desired performance can be obtained from the GlobalFiler™ Express Amplification Kit.

Methods and Materials

Samples were punched or cut and then immersed in varying volumes of PrepNGo™ Buffer and incubated at various temperatures for at least twenty minutes. For an Extraction Negative Control (ENC), 200 µL of PrepNGo™ Buffer was used.

Samples were amplified with 6 µL of Master Mix, 6 µL of the Primer Set from the GlobalFiler™ Express Amplification Kit, and 3 µL of the sample. For the Amplification Negative Control (ANC), 3 µL of PrepNGo™ solution was added to the Master Mix and Primer Set. For the Amplification Positive Control (APC), a variation in the ratio of 007 Control to PrepNGo™ Buffer was analyzed. The GeneAmp® PCR System 9700 was set in Max ramping mode and used to amplify the samples. The thermal profile consisted of an initial hold at 95°C for one minute, 27 cycles of denaturation at 94°C for 3 seconds and anneal/extend at 60°C for 30 seconds, and then a final extension at 60°C for 8 minutes followed by the final hold at 4°C.

The Applied Biosystems® 3500 Genetic Analyzer (Life Technologies™, Foster City, CA) was used to separate and detect the amplified product. 1 µL of the amplified samples and controls (including allelic ladders) was combined with 9.5 µL of Hi-Di Formamide and 0.5 µL of GeneScan® 600 LIZ size standard. The GlobalFiler™ Allelic Ladder was included once every three injections. The samples and ladders were injected at 1.2 kV for varying injection times and

separated in Performance Optimized Polymer-4 (POP-4) at 13 kV for varying amounts of time depending on injection time (from 1550 seconds for 5 second injection to 1210 for 20 second injection). All data was analyzed using GeneMapper® *ID-X* version 1.4.

Cycle Number

Prior to any validation studies, a 27 cycle thermal profile was decided on since the suggested range is 25 to 27 cycles and the cycle number can always be reduced if necessary (14).

Sensitivity Studies / Non-Probative Sample Type Studies

Volume of PrepNGo™ Buffer/Number of Punches

Since the Applied Biosystems® GlobalFiler™ Express Kit uses PrepNGo™ Buffer to lyse cells and pull the DNA into solution instead of a classic extraction step, the procedure for use of the PrepNGo™ Buffer needed to be evaluated. With treated FTA paper, Applied Biosystems'® preparation guidelines suggest putting one 1.2mm punch of each sample directly into a plate well and adding 10µL of the amplification master mix and primer set (3). They do not suggest adding the punches to PrepNGo™ Buffer since the cells have already been lysed by the FTA paper. Punching samples directly into wells may introduce unnecessary contamination since punches seem to 'jump' between wells due to static on the 96-well plate. In addition, the St. Louis County Police Crime Laboratory performs the sampling step during the Biology screening section prior to DNA. Therefore, it is infeasible to punch a sample into a plate as only one case is worked on at a time in the Biology section and cases are batched together for increased throughput during DNA analysis. The St. Louis County Police Laboratory decided they would have the Biologists take punches in 2 mL tubes and the DNA Analysts would add PrepNGo™ Buffer. The DNA Analysts would then incubate and pull from this solution for

amplification. This allows the prepared sample to also be conserved in case re-amplification is required.

Since the previous ratio of one punch to 10µL of Master Mix and Primer Set solution would leave a very small volume of sample left in the 2 mL sample tube, the laboratory decided to have the Biologists take two punches of each sample into each 2 mL sample tube. Because the concentration of DNA that is put into the amplification process needs to be optimized but no quantitation step is incorporated in an Express Kit workflow, a variation of PrepNGo™ Buffer was added to punches from two samples. This was done to determine what volume provided the optimum DNA concentration for amplification and to establish the sensitivity of the kit. Aliquots of 20µL, 50µL, 100µL, and 200µL of PrepNGo™ Buffer were added to the two punches from two samples and allowed to stand at room temperature for at least twenty minutes before being amplified. After determining the ideal volume of buffer to add to each sample, two samples were prepared with one punch and two punches to verify that the number of punches in the chosen volume were optimal.

For swab substrates, Applied Biosystems® suggests placing the entire swab head into 400µL of PrepNGo™ Buffer (2). For conservation purposes, the lab wanted to minimize the amount of sample consumed so they decided on a ½ cutting of the swab in 200µL of PrepNGo™ Buffer. Two different non-probative sample types (q-tip oral swabs and Whatman® comb-style buccal swabs) that the laboratory normally receives were prepared as described above and run to confirm that this preparation would still give optimal results.

The amount of time that the samples spent in PrepNGo™ Buffer before amplification was also evaluated. Samples were amplified after being extracted in PrepNGo™ buffer for 20

minutes, 24 hours, and 48 hours. The peak heights and number of full profiles were compared to determine if samples would produce better results if they were to sit in PrepNGo™ Buffer for longer than the recommended twenty minutes (1).

Micro Punch Sampling / Cross Contamination Study

A single 1.2 mm Harris Micro Punch™ Tip was used to sample all FTA blood cards. A single punch on a 'blank' or new FTA card was proposed to clean the punching apparatus in between each sample. To ensure that the tool was not a source of contamination, two 'blank' punches were taken after each sample so that the second 'blank' punch would show if contamination was carried to the next sample during casework. The punches were put in the decided optimal amount of PrepNGo™ Buffer (20μL), and taken through the laboratory work flow. This was done for three samples giving a total of 6 'blank' punches.

Incubation Temperature

Four FTA samples and twelve swab samples were subjected to three different incubation temperatures (room temperature, 37°C, and 56°C) for twenty minutes. The samples were amplified and separated on the Applied Biosystems® 3500 Genetic Analyzer at the varying injection times. The overall quality of the resulting profiles was compared for the various incubation times.

Injection Time Study

Four FTA card samples and two swab samples that had been previously amplified for other studies were injected at 5 second, 10 second, and 15 second injections. All other parameters of the injection and electrophoresis were maintained. Later, a comparison of 10

second, 15 second, and 20 second injections was done to optimize the protocol for low yield samples. An examination of the number of full profiles and the average peak heights was completed to compare the injection times. A comparison between injection times with allelic ladders was also performed since the quality and quantity of the allelic ladder would be the same for each injection. This was to determine an optimal injection time and an acceptable injection time range based on sample type.

Analytical Threshold Study

Eight amplification negative samples were run in duplicate at three different injection times and then analyzed at a threshold of 1 Relative Fluorescence Unit (RFU). The peak data was separated by dye color and the average peak heights, standard deviation of the peak heights, and maximum peak height were calculated for all five dyes. The minimum peak height was set to 1 RFU since that was the analytical threshold. Two equations were used to determine the analytical threshold.

The first equation used was suggested by Scientific Working Group for DNA Analysis Methods (SWGDAM) and is as follows (12):

Equation 1:

$$Analytical\ Threshold\ = 2(Maximum\ peak\ height - Minimum\ peak\ height)$$

The second equation determines the Limit of Detection which is another name for the Analytical Threshold and is as follows (7):

Equation 2:

$$Analytical\ Threshold\ = Average\ peak\ height + (3 \times Standard\ Deviation\ peak\ height)$$

Both equations were used to calculate the analytical threshold for each dye at each injection time and the larger threshold was chosen for conservative purposes.

Stochastic Threshold

Approximately 85 samples run in a previous study were used again to determine the stochastic threshold. The average peak height ratio (PHR) and standard deviation of peak height ratio per dye was calculated for these samples. Since there is no quantitation data for a sensitivity study with a direct amplification kit, a single equation was used to calculate the Stochastic Threshold for each dye. The equation used was recommended in a previous validation study and is as follows (13):

Equation 3:

$$Stochastic\ Threshold = Analytical\ Threshold\left(\frac{1}{Average\ PHR - 3 \times Standard\ Deviation\ PHR}\right)$$

As a secondary check, the highest false homozygote peak height for each dye was determined. The peak height of this peak should be under the stochastic threshold calculated by Equation 3. When it was not below the stochastic threshold, the highest surviving false homozygote peak height per dye was rounded up to the nearest multiple of ten and the stochastic threshold was adjusted as necessary.

Contamination Study

A 96 well plate with allelic ladders and samples was set up in a checkerboard orientation with run negatives (Formamide with LIZ 600) in between to check for contamination between injections. The allelic ladders were placed starting in well A1 and descended diagonally across

the plate so that an allelic ladder was injected into each capillary. The run negatives were then analyzed to ensure there was no contamination between injections. In addition, all lysed and amplified sets had ENCs and ANCs to ensure all lysing and amplification reagents are free of contamination.

Precision Study (within injection)

Twenty-four allelic ladders were run during three different injections at all three injection times (5, 10, and 15 seconds). Each injection and injection time was evaluated separately. Each locus was then individually evaluated. The standard deviation in base pair size per allele and the difference in maximum base pair size from minimum base pair size were calculated. The results were examined to assure that the standard deviation of base pair size was less than 0.15 base pairs and the difference in maximum and minimum base pair size was less than ± 0.5 base pairs. Ideally, the variation within the injections will fall below these limits indicating minimal variation in sizing precision for all alleles within the injection.

Reproducibility Study (between injection / injection days)

The ladders used for the Precision Study were set up on three different days and analyzed for reproducibility between days. The allele calls and sizing precision between days were compared to make certain that the same peaks were present and were sized within the ± 0.5 base pair window with a standard deviation less than 0.15 base pairs.

Concordance Study

One hundred Caucasian and one hundred African American FTA blood card samples as well as 30 internal swab (q-tip oral swabs and Whatman® comb-style buccal swabs) samples

were previously amplified and run with the AmpFLSTR® Identifiler® PCR Amplification Kit. These same samples were amplified and run with the GlobalFiler™ Express PCR Amplification Kit. Both sets of data were analyzed using the GeneMapper® ID-X Software and then compared to determine if the two kits produced concordant profiles.

Population Study

With the incorporation of six new autosomal loci, local frequencies of these loci need to be compared to national databases to verify that the local population does not vary significantly from the national population. One hundred Caucasian and one hundred African American local genotypes were compared to 361 Caucasian and 342 African American national genotypes provided by the National Institute of Standards and Technology (NIST) to determine if the local frequency of alleles is comparable to the national frequencies. The genotypes were entered into the Promega® Powerstats V12 excel program to generate the frequency of each allele at the new loci and a p-value for each new locus was calculated. The p-value expresses how concordant the local allele frequencies are with those national frequencies provided by NIST (9). Chi squared and global chi-squared distributions were also evaluated. The power of inclusion and heterozygosity per locus were evaluated based on statistics published by Promega® (5).

Internal Stutter Study

Eighty samples were analyzed with the stutter filters set to zero for all loci to establish appropriate stutter ratios for the St. Louis County Police Crime Laboratory with the Applied Biosystems® GlobalFiler™ Express Kit on the Applied Biosystems® 3500 Genetic Analyzer. When the data was analyzed in GeneMapper® ID-X, all other artifacts besides minus eight, minus four, minus two (for SE33 and D1S1656), and plus four stutter were removed so that only

these artifacts could be analyzed. The analyzed data was then exported to excel and separated by locus. Each stutter artifact was labeled accordingly. The ratio for each category of stutter was then calculated and the average and standard deviation for the stutter ratio within each locus was calculated. The following equation was used to determine the internal stutter cutoff.

Equation 4:

$$\begin{aligned} &Internal\ Stutter\ Ratio \\ &\quad = Calculated\ Average\ Stutter\ Ratio + (3\ \times Standard\ Deviation\ of the\ Stutter\ Ratio) \end{aligned}$$

These ratios were then compared to the stutter ratios recommended by Life Technologies®. Life Technologies® only had minus four stutter ratios for all loci with one plus four stutter ratio for D22S1045 and minus two stutter ratios for SE33 and D1S1656. With the noted exceptions, Life Technologies® did not provide stutter ratios for minus eight, minus two, and plus four stutter ratios so these could not be compared with the St. Louis County Police Crime Laboratory results.

Mixture Study

Three samples, two female and one male, were prepared separately with PrepNGo™ Buffer and then combined in different ratios during the amplification set up. These three samples were chosen because in previous studies they showed similar peak heights suggesting similar concentrations. The ratios chosen were 1:2, 1:1, 2:1, and 1:1:1. The total amount of template input is 3 µL so the ratios reflect the volumes of each sample added during the amplification set up. The samples were amplified and then run on the 3500 Genetic Analyzer at 10 and 20 second injection times. The profiles were analyzed with GeneMapper® ID-X and then evaluated to ensure that the Applied Biosystems® GlobalFiler™ Express Amplification Kit could detect

mixtures. The expected mixture proportions were compared to the RFU ratios to determine if the kit would amplify each contributor according to the amount of input DNA added to the sample.

Pass Rate

Previously run data was compiled and analyzed by the number of full profiles that were obtained from the total number of samples that were run when preparing them at the optimal parameters. A comparison of the first pass rate to the final pass rate was used to show how the parameters of the protocol were optimized to obtain the best profiles for analysis while decreasing the amount of time and resources used. Samples that were previously used to optimize the sample preparation were not included in this study. Any samples that did not produce full profiles after optimizing the protocol were taken through the normal laboratory workflow. The samples were extracted using the EZ1® DNA Investigator Kit on the Qiagen® EZ1® (Qiagen®, Hilden, Germany) and quantitated using Plexor® HY (Promega®, Madison, Wisconsin). They were then amplified with the Identifiler® Amplification Kit and separated on the Applied Biosystems® 3500 Genetic Analyzer to simulate the results that would be produced once the GlobalFiler™ Amplification Kit is released. These samples were run with 5, 10, and 15 second injection times.

Results

Sensitivity Studies / Non-Probative Sample Type Studies

Volume of PrepNGo™ Buffer & Number of Punches

The current protocol employed by the laboratory designates a fifteen second injection as a standard injection time. A range of injection times were used to allow for flexibility with low, medium, and high quantity samples.

Two FTA card samples in 20, 50, 100, and 200 µL of PrepNGo™ Buffer were injected at each injection time in duplicate and were analyzed for the number of complete profiles, number of incomplete loci, and the amount of total allelic dropout (Table 1). With a ten second injection time, full profiles in all FTA card samples were only seen with 20 µL of PrepNGo™ Buffer. FTA samples extracted in 20 and 50 µL of PrepNGo™ Buffer yielded complete profiles with a 15 second injection time. The increased injection time improved allele recovery. The range of dropout in the incomplete profiles ranged from 3 alleles at five second injection in 20 µL PrepNGo™ Buffer to 37 alleles at five second injection in 200µL of PrepNGo™ Buffer. Increasing the volume of PrepNGo™ Buffer increased the rate of dropout for the FTA samples.

Table 1 - Volume of PrepNGo™ Buffer Study: Number of complete and incomplete profiles for different PrepNGo™ Buffer volumes with varying injection times. Two sample punches were used in this comparison.

Buffer Volume (µL)	Injection Time (s)	# Complete Profiles (out of 4)	Average # of Incomplete Loci	Average # of Missing Alleles (Dropout)
20	5	2	2	3.5
	10	4	-	-
	15	4	-	-
50	5	0	9.5	16
	10	2	4.5	6.5
	15	4	-	-
100	5	2	20	36
	10	2	14	23.5
	15	2	7	9
200	5	0	20.25	37
	10	0	14	21.5
	15	0	9	12.25

With two punches of the FTA card in 20 µL of PrepNGo™ Buffer, full profiles were seen for every injection time. One punch of the FTA card in 20 µL of PrepNGo™ Buffer yielded complete profiles with the ten and fifteen second injection times. One punch of the FTA card in 20 µL of PrepNGo™ Buffer did not produce full profiles in all samples when run with a five second injection time. These parameters resulted in 2 to 3 dropout alleles per sample. Full profiles at all injection times were seen with the two swab samples that were run.

Cross Contamination Study

Contamination was only seen in the first cleaning 'blank' punch samples. Two of the six fifteen second injection samples were the only samples to have contamination. All other first punch 'blanks' were true blanks. All second punch 'blanks' were true blanks indicating one blank punch between samples is sufficient to clean the tool (i.e. no DNA was transferred to the second blank punch which would be the next sample during casework).

Incubation Temperature

All four FTA samples that were incubated at 37°C produced full profiles (Table 2). Two of the four samples incubated at 56°C produced full profiles with the two incomplete profiles only missing a total of five loci. Room temperature samples only produced one complete profile out of four and had up to fourteen missing loci in one sample. Average peak height ratios (PHR) are comparable between the room temperature, 37°C, and 56°C incubations and ranged from 78% to 89%. The highest peak heights (RFU) in each sample were seen in the 37°C samples. Sample peak heights from the 37°C incubation were between 1.359 and 8.855 times higher than the peak heights from the same sample incubated at room temperature. Sample peak heights

from the 56°C incubation were between 1.353 and 4.428 times higher the peak heights from the same sample incubated at room temperature.

Table 2 – Incubation Temperature: Comparison of Peak Heights between incubation times for four different FTA samples.

Sample	Temp (°C)	Missing Loci	Average PHR	Ratio of Average PH to RT	Ratio of Average PH to 56
A101	RT	0	0.8419	-	-
	37	0	0.8175	3.175	1.136
	56	0	0.8639	2.822	-
A102	RT	9	0.8672	-	-
	37	0	0.8172	5.237	2.967
	56	4	0.7854	1.825	-
A103	RT	14	0.8400	-	-
	37	0	0.8952	**8.855**	**2.997**
	56	1	0.8475	**4.428**	-
A104	RT	1	0.8247	-	-
	37	0	0.8271	1.359	1.027
	56	0	0.8249	1.353	-

Injection Time Study

Although 3 injection times were validated (10, 15, and 20 seconds), the laboratory decided to include a 10 and 20 second injection time within their protocol to account for both high and low yield samples. Even though a ten second injection yielded lower peak heights, 97.6% (83 of 85) of FTA blood card samples were full profiles. When swab samples were analyzed, a longer injection time was needed. A 20 second injection time yield peak heights that are between 1.55 and 2.75 times larger than peak heights from a 10 second injection (Table 3 and 4). Because of the increase in peak height yielded by the 20 second injection time, more profiles were recovered for swab samples. Therefore, a combination of the two was decided upon to suffice for high (FTA blood card) and low (swab) yield samples.

Table 3 – Injection Time Study: The average and maximum peak heights of allelic ladders between injection times.

	Injection Times (s)	Average PH (RFU)	Max PH (RFU)
Ladders	20	9178.916667	11505
	15	7929.6875	9871
	10	3606.958333	4439

Table 4 – Injection Time Study: Comparison of injection time ratios including the minimum and maximum ratio between injection times.

Ratio	Average Ratio	Max. Ratio	Min. Ratio
20 sec : 15 sec	1.53866257	1.87218893	1.195071
20 sec : 10 sec	2.08646688	2.74823768	1.553612
15 sec : 10 sec	1.53866257	1.87218893	1.195071

Analytical Threshold Study

The data from eight amplification negative samples was used with Equation 1 and Equation 2. Equation 1 suggested by Scientific Working Group for DNA Analysis Methods (SWGDAM) Mixture Interpretation Guidelines Section 1.1 resulted in the largest analytical threshold per dye color in each injection time (Table 5) (12). After rounding peak heights up, the maximum threshold calculated by Equation 1 was 80 RFU in yellow with a 10 second injection where the maximum threshold calculated by Equation 2 was 30 RFU. The maximum calculation (Equation 1) per dye color resulted in the analytical threshold for Blue being 50 RFU, Green being 70 RFU, Purple being 60 RFU, Red being 70 RFU, and Yellow being 80 RFU. After analyzing some data with the calculated analytical threshold and calculating the suggested stochastic threshold, it was decided to have the same analytical threshold of 80 RFU for all dyes. There was no correlating increase in the calculated analytical threshold with an increase in injection time.

Table 5 – Analytical Threshold Study: The data used for and the comparison between the two equations used to estimate the Analytical Threshold per dye for a ten second injection.

	10 seconds injection					
Dye	Average PH (RFU)	St. Dev. PH (RFU)	Max PH (RFU)	Min PH (RFU)	Equation 1: 2*(Max-Min)	Equation 2: Avg. + 3 Std. Dev.
Blue	7.431	2.483	18	1	34	14.879
Green	12.01	3.635	27	1	52	22.914
Purple	9.593	2.737	25	1	48	17.805
Red	11.14	3.781	29	1	56	22.484
Yellow	5.386	1.992	37	1	72	11.362

Stochastic Threshold

After deriving the Analytical Threshold, PHR averages, and standard deviation per dye color, the results were used to estimate the stochastic threshold (Equation 3) (Table 6). The results were then rounded up to the nearest multiple of ten suggesting the stochastic threshold for Blue should be 140 RFU, Green should be 150 RFU, Purple should be 140 RFU, Red should be 170 RFU, and Yellow should be 130 RFU.

Table 6 - Stochastic Threshold Study: The calculations used in Equation 3 to estimate the Stochastic Threshold per dye.

	Average PHR	Std. Dev. PHR	Analytical Threshold (RFU)	Stochastic Threshold (RFU)	Max. False Homozygote PH (RFU)
Blue	0.8778	0.0884	80	130.60	N/A
Green	0.8594	0.0985	80	141.86	262
Purple	0.8792	0.0939	80	133.92	270
Red	0.8577	0.1234	80	164.11	221
Yellow	0.8990	0.0843	80	123.82	179

To determine if the calculated stochastic threshold would be appropriate, samples with low level data (known dropout) were examined for false homozygotes above the stochastic threshold from Equation 3. All samples were analyzed at the analytical threshold and only loci with the sister allele falling below the analytical threshold (true dropout) were examined. The

highest peak height per dye of the false homozygotes was 262 RFU for Green, 270 RFU for Purple, 221 RFU for Red, and 179 RFU for Yellow. There were no false homozygotes called for Blue. Because all of these false homozygote peak heights are above the calculated stochastic threshold, a new, conservative Stochastic Threshold was estimated to be 300 RFU for every dye.

Contamination Study

All of the run negatives, consisting of only formamide and LIZ 600, placed between the ladders and samples did not contain any called peaks except peaks that could be identified as pull-up from the ILS LIZ 600. Figure 1 shows an example of the peaks called. Because these allele calls could be attributed to pull-up, all run negatives were true negatives. It should be noted that the St. Louis County Police Crime Laboratory does not use run negatives as their amplification negative control is used as their capillary electrophoresis negative control. All ENCs and ANCs were free of amplified DNA at all injection times utilized.

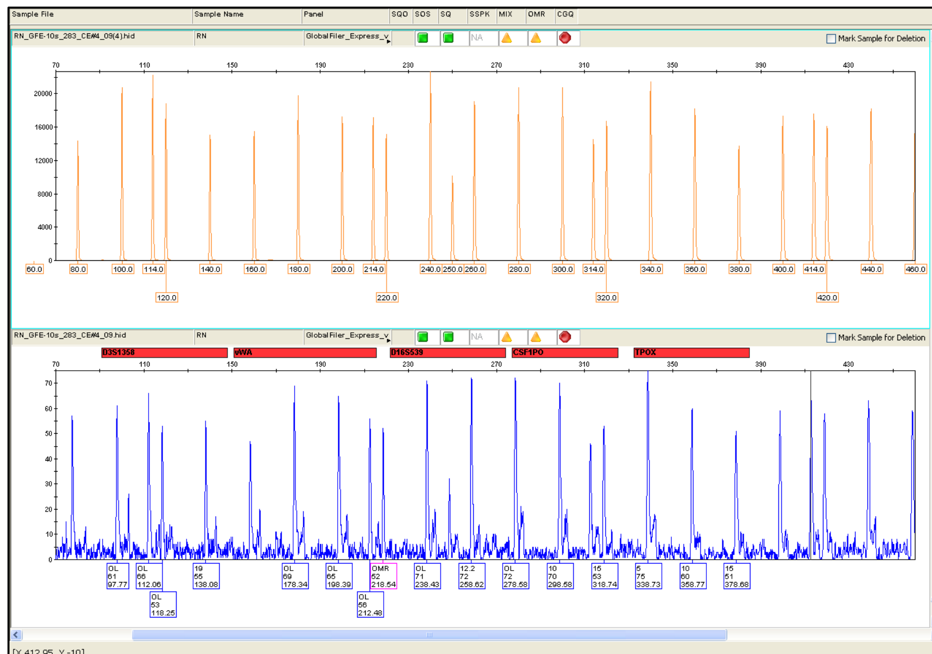

Figure 1 – Contamination Study: The electropherogram of a run negative with allele calls in the blue dye channel resulting from pull-up from the ILS (orange dye channel).

Precision Study

All 24 allelic ladders were called as expected. The standard deviation and the difference between the maximum and minimum base pair size were calculated for each allelic ladder. All standard deviation were under 0.15 base pairs and all differences between maximum and minimum base pairs size were under 0.5 base pairs. The maximum standard deviation seen for a ten second injection was 0.0637 from the 34.2 allele in the SE33 loci and the maximum difference between base pair size within a bin was 0.20 from the 32.2 allele in the SE33 loci.

Reproducibility Study

All the ladders compared within this study were sized within the ± 0.5 base pair window with a standard deviation less than 0.15 base pairs and produced the same profiles between days. These results demonstrate that the Applied Biosystems® GlobalFiler™ Express Kit gives reproducible data for analysis.

Concordance Study

Out of the 200 FTA blood card samples that were previously run with AmpFLSTR® Identifiler® Kit, all of the samples were concordant when run with Applied Biosystems® GlobalFiler™ Express Amplification Kit. Of the 30 internal swab (q-tip oral swabs and Whatman® comb-style buccal swabs) samples, 26 full profiles were obtained which were concordant with the AmpFLSTR® Identifiler® Kit. For the four samples from which incomplete profile were obtained, the loci that were recovered were concordant with the AmpFLSTR® Identifiler® Kit. Four off-ladder allele calls from the AmpFLSTR® Identifiler® Kit analysis were given actual allele calls with the Applied Biosystems® GlobalFiler™ Express Amplification Kit due to the extra alleles added to its ladder.

Population Study

All six new loci had similar frequencies to those provided by NIST. The p-values for each individual locus for the Caucasian and African American samples were above 0.05 (5%) meaning there is not a significant difference between the local and national allele frequencies (Table 7). Further testing included the global chi-squared distribution test which tested all the new loci at once. The null hypothesis used was that the two sample groups (local and national) were taken from the same population. The resulting global chi-squared distribution p-value was 0.612 for Caucasians and 0.831 for African Americans confirming that the local allele frequencies are not significantly different than the national allele frequencies (9). A comparison of the power of inclusion and heterozygosity showed local Caucasian and African American genotypes were similar to national genotypes (Table 8 and 9) (5). SE33 had the highest PIC and the highest percentage of heterozygosity indicating it to be a very discriminating locus that will be valuable in mixture interpretation. The St. Louis County Police Crime Laboratory will use the allele frequencies provided by NIST and incorporated by the FBI Popstats program for future casework.

Table 7 – Population Study: The internal sample p-values for the new GlobalFiler™ loci for the Caucasian and African American populations. A p-value less than 5% (0.05) indicates a significant variation from the national population.

p - values					
D1S1656		**D2S441**		**D10S1248**	
Caucasian	Black	Caucasian	Black	Caucasian	Black
0.204	0.636	0.943	0.201	0.947	0.792
D12S391		**D22S1045**		**SE33**	
Caucasian	Black	Caucasian	Black	Caucasian	Black
0.377	0.318	0.543	0.899	0.238	0.747

Table 8 – Population Study: The power of inclusion (PIC) and heterozygosity statistics for the new GlobalFiler™ loci between local and national Caucasian populations.

		Caucasian					
		D1S1656	D2S441	D10S1248	D12S391	D22S1045	SE33
Local	PIC	0.890	0.710	0.720	0.880	0.670	0.940
	Heterozygotes	85.00%	77.00%	75.00%	87.00%	78.00%	96.00%
National	PIC	0.891	0.735	0.719	0.881	0.677	0.946
	Heterozygotes	92.52%	78.67%	76.45%	89.75%	75.35%	95.01%

Table 9 – Population Study: The power of inclusion (PIC) and heterozygosity statistics for the new GlobalFiler™ loci between local and national African American populations.

		African American					
		D1S1656	D2S441	D10S1248	D12S391	D22S1045	SE33
Local	PIC	0.850	0.740	0.750	0.840	0.800	0.920
	Heterozygotes	88.00%	78.00%	76.00%	87.00%	84.00%	95.00%
National	PIC	0.852	0.723	0.775	0.844	0.801	0.926
	Heterozygotes	87.13%	78.95%	82.16%	86.26%	80.41%	92.69%

Internal Stutter Ratios

The average, standard deviation, and average plus three times the standard deviation of the stutter ratios were calculated per locus and are included in Table 10. Internal percentages were calculated for minus eight, minus four, minus two, and plus four stutter. The Life Technologies™ stutter values were then compared to the internal values (Table 10). The internal minus four stutter values were consistent with the Life Technologies™ recommendation. A higher D22S1045 minus four stutter percentage was seen internally but upon further evaluation, the increased average and standard deviation may be influenced by plus four stutter from the sister allele. The plus four and minus two stutter values were comparable to the Life

Technologies® recommendations when present. D22S1045 plus four stutter was significantly higher than the Life Technologies™ stutter ratios but this may be due to the fact that only 2 plus four stutter artifacts seen at this locus (9.39% and 14.41% stutter ratio). During the internal stutter analysis, DYS391 did not produce any minus four stutter and D1S1656 did not produce any minus two stutter and are represented by 0.00. Yindel and Amelogenin do not produce stutter artifacts and are also represented by 0.00. Since the stutter values are mostly comparable, the Life Technologies™ stutter ratios that are already within the panel will be used in GeneMapper analysis. Interpretation Guidelines will indicate if a peak above the GeneMapper filters will be called as stutter or if it is an indication of a mixture or contamination .The minus eight or plus four stutter values obtained from this validation will be used as a guide for analysts to determine if a peak in those positions are stutter products or possible contamination.

Table 10 – Internal Stutter Ratios: The comparison between the internal and Life Technologies recommended stutter percentages at every loci.

	Internal			Life Technologies
Marker	**Stutter %**	**Std. Dev.**	**Equation 4: Stutter % + 3Std. Dev.**	**Stutter %**
D3S1358	8.19	1.78	13.53	12.45
vWA	6.99	1.94	12.82	12.33
D16S539	5.58	1.63	10.48	10.17
CSF1PO	5.21	1.25	8.97	11.4
TPOX	2.76	1.03	5.84	5.43
Yindel	0.00	-	-	0.00
AMEL	0.00	-	-	0.00
D8S1179	6.03	2.36	13.11	10.22
D21S11	6.84	1.38	10.99	11.42
D18S51	7.56	2.04	13.68	13.47
DYS391	0.00	-	-	8.54
D2S441	5.10	1.25	8.85	8.75
D19S433	7.03	1.50	11.54	10.58
TH01	2.11	0.60	3.91	5.24
FGA	7.28	1.43	11.57	11.96

D22S1045	9.06	2.87	**17.67**	8.27
(Plus 4)	11.90	3.54	**22.53** (2 cases)	17.3
D5S818	6.32	1.30	10.22	10.84
D13S317	4.98	1.20	8.58	9.98
D7S820	4.74	1.65	9.71	10.21
SE33	8.82	2.22	15.47	14.42
(Minus 2)	2.99	0.31	3.92	4.97
D10S1248	7.81	1.46	12.21	12.5
D1S1656	7.60	1.95	13.44	13.08
(Minus 2)	0.00	-	-	1.79
D12S391	8.80	2.61	16.62	15.08
D2S1338	8.17	1.63	13.07	12.45

Mixture Study

Mixture ratios of 1:2, 1:1, 2:1, and 1:1:1 were made to determine if the GlobalFiler™ Express Amplification Kit could detect mixture samples. All four ratios showed full profiles of each component with proper peak height ratios between components. Varying injection times between 10 and 20 second injections did not affect the profiles or the component ratios.

Figure 2 – Mixture Study: The blue channel electropherogram from a 1:1:1 mixture for a ten (top) and twenty (bottom) second injection time (showing no change in allele calls or mixture ratio proportions between injection times).

Pass Rate

The first run that was considered included only FTA card samples. To prepare these samples, two punches were taken and cells were lysed in 20 μL of PrepNGo™ Buffer for 20 minutes at room temperature. One 'blank' cleaning punch was taken between each sample. The samples were then amplified with 27 cycles and injected for 10 seconds. 83 of the 85 samples run produced full profiles giving a first pass rate of 97.65%. When the samples were subject to 10 and 20 second injections as well as incubated at 37°C for 20 minutes, all samples produced full profiles.

Buccal swabs proved to be more challenging than FTA card samples. The first pass rate for buccal swabs was 61.90% with only 13 of 21 samples producing full profiles. To prepare

these samples, half of a swab was submerged in 200 µL of PrepNGo™ Buffer for 20 minutes. Half of the samples were incubated at 37°C and half were incubated at 56°C. The samples were then amplified with 27 cycles and injected for 10 and 20 seconds. Between incubation times, 37°C produced 3 of 6 full profiles (50.00%) and 56°C produced 4 of 6 full profiles (66.67%). Between injection times, a 10 second injection produced 7 of 12 full profiles (58.33%) and a 20 second injection produced 6 of 9 full profiles (66.67%). Even after modification to incubation time, injection time, and the time in the PrepNGo™ buffer, the best standing pass rate is 66.67% (16 of 24). This last run was prepared with half of a swab was submerged in 200 µL of PrepNGo™ Buffer for 48 hours. Samples were incubated at 37°C then amplified with 27 cycles and injected for 10 and 20 seconds.

Since no significant increase in pass rate was seen with modification to the protocol, the remaining 4 samples of the total 31 buccal samples were subject to an EZ1 extraction, quantitation with Plexor® HY and amplified with the Identifiler® Amplification Kit. The samples were separated with 5, 10, and 15 second injections and analyzed using GeneMapper® ID-X. All gave full and concordant profiles with the previous Identifiler® Amplification Kit results.

Discussion and Conclusions

The Applied Biosystems® GlobalFiler™ Express Amplification Kit was found to produce reliable, reproducible, and robust results when used on the Applied Biosystems® 3500 Genetic Analyzer. The validation studies that were performed helped to define the parameters that produced these results. For FTA card samples it was determined that two 1.2 mm punches will be submerged in 20 µL of PrepNGo™ Buffer solution and incubated for 20 minutes in a

37°C oven before amplification. In between samples, a ‘blank’ punch of a clean FTA card will be taken and discarded to prevent cross contamination between samplings. For swab samples, the optimal procedure is ½ of the swab will be submerged in 200 µL of PrepNGo™ Buffer solution and incubated for 20 minutes in a 37°C oven before amplification. Either type of sample can be left in the PrepNGo™ Buffer solution for at least 48 hours without deleterious effects. If the swab samples did not produce full profiles with the optimized protocol, they would be taken through the normal laboratory workflow with the Applied Biosystems® Identifiler® Amplification Kit.

A 27 cycle thermal profile was decided on before the validation studies were started. The internal validation suggests that a 10 second injection at 1.2 kV should be used for FTA card samples and a 20 second injection time at 1.2 kV should be used for buccal samples. The St. Louis County Police Crime Laboratory has decided to implement a combination of the 10 second and 20 second injection time to ensure that low yield and high yield samples will be able to produce optimal results without rerunning the samples. An analysis threshold per dye was found and an overall threshold was established to make implementation into analysis easier. An analysis threshold of 80 RFU for all dyes will be used in the final protocol. The established stochastic threshold of 300 RFU for all dyes allowed profiles without false homozygotes to be called.

The validation studies that were performed also demonstrated the precision within an injection and the reproducibility and absence of contamination between injections. Reproducibility was also shown to exist between three separate amplification and CE runs on three different days. Contamination was not seen within any ENCs or ANCs indicating the reagents used were free of contamination. Per the SWGDAM Quality Assurance Standards

Clarification Document, if any reagents or water are not added to the associated samples, a combined ENC and ANC can be used for direct amplification kits. When this kit is implemented into the laboratory protocol, a Combined Negative Control (CNC) may be used instead of the separate ENC and ANC (11).

Samples previously analyzed with AmpFLSTR® Identifiler® Amplification Kit on the Applied Biosystems® 3500 Genetic Analyzer were concordant when analyzed with the GlobalFiler™ Express Amplification Kit on the Applied Biosystems® 3500 Genetic Analyzer. Internal stutter ratios were determined and compared to Life Technologies'™ recommended stutter ratios and the Life Technologies™ stutter ratios will be used for future analysis. The local allele frequencies were also comparable to national allele frequencies. NIST national population allele frequencies will be used in casework statistical calculations.

Overall the Applied Biosystems® GlobalFiler™ Express Amplification Kit was shown to produce optimal results from common types of reference samples while being cost and time effective. The validation studies performed gave internal recommendations for sampling, incubation time, analytical threshold, stochastic threshold, and stutter ratios. Through these studies it was determined that the Applied Biosystems® GlobalFiler™ Express Amplification Kit used on the Applied Biosystems® 3500 Genetic Analyzer could be trusted to produce reliable genotypes from single-source samples for the St. Louis County Police Crime Laboratory. The National DNA Index System (NDIS) is currently evaluating the kit. Pending approval by NDIS, this kit is recommended for use by the St. Louis County Police Crime Laboratory on future reference samples to increase throughput and cost efficiency. Future validations will be done to allow the Applied Biosystems® GlobalFiler™ DNA Amplification Kit to be implemented with casework unknown samples.

Acknowledgements

This validation benefitted from the input of Kyra Groeblinghoff, Season Seferyn, and Dr. Pamela Staton, all of who provided assistance and critiques of the presentation of these studies. Thanks should also be given to Jennifer Hayden for giving instruction throughout this project as well as to Chris Thatch and Allison Gapinski who provided support in the production of this paper. The author would like to thank the analysts and employees of the St. Louis County Police Crime Laboratory who gave valuable input to this project and made the experience within their lab worthwhile. This project was supported by Award No. 2009-IJ-CX-K11 awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. The opinions, findings, and conclusions or recommendations expressed in this publication/program/exhibition are those of the author(s) and do not necessarily reflect the views of the Department of Justice.

References

1. Applied Biosystems. GlobalFiler™ Express PCR Amplification Kit – PCR Amplification and CE Quick Reference. 2012.

2. Applied Biosystems. GlobalFiler™ Express PCR Amplification Kit – PCR Setup Swab Substrate Quick Reference. 2012.

3. Applied Biosystems. GlobalFiler™ Express PCR Amplification Kit – PCR Setup Treated Paper Substrate Quick Reference. 2012.

4. Applied Biosystems. GlobalFiler™ Express PCR Amplification Kit User Guide. Rev 10/2012.

5. Butler, John M., Carolyn R. Hill, and Michael D. Coble. Variability of New STR Loci and Kits in US Population Groups. Promega. 2012.< http://www.cstl.nist.gov/strbase/NISTpop.htm>.

6. DNA Advisory Board. Quality Assurance Standards for Forensic DNA Testing Laboratories. Forensic Science Communications, July 2000, Volume 2, Number 3.

7. Grgicak, Catherine M., Analytical Thresholds: Determination of Minimum Distinguishable Signals. ISHI. San Antonio. 2010. Lecture.

8. Hares, D.R. (2012a) Expanding the CODIS Core Loci in the United States. FSI: Genetics 6(4), e52-54.

9. Lucy, David. *Introduction to Statistics for Forensic Scientists*. Chichester, West Sussex, England: Wiley, 2005. Print.

10. "Quality Assurance Standards for Forensic DNA Testing Laboratories Effective 9-1-2011." *FBI*. FBI, 14 Apr. 2011. Web. 30 July 2013

11. Scientific Working Group on DNA Analysis Methods – QAS Clarification Document. May 16, 2013.

12. SWGDAM Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories. January 14, 2010.

13. Thatch, Chris, Danielle Imes, Valerie Bostwick, Pamela J. Staton. Internal Validation of PowerPlex® 16 HS with the Applied Biosystems® 3500xL Genetic Analyzer. 2012.

14. Wang, Dennis. Development of a “Global” STR Multiplex for Human Identification Analysis. ISHI. Nashville. 2012. Lecture.