On Going Validation of the 3500XL Genetic Analyzer using the AmpF*l*STR Identifiler Plus® Kit

Vanity Marie Maldonado
901-60-9424

FSC 630 Forensic Science Program
Marshall University Forensic science Program

MU Topic advisor:	Dr. Pamela Staton
Internship Agency Supervisor:	Dr. Theresa Caragine Deputy Director Department of Forensic Biology Office of Chief Medical Examiner Tel: 212-323-1250 Fax: 212-323-1590
Internship Agency:	Office of Chief Medical Examiner 421 East 26th Street New York, New York 10016 Fax: 212-323-1590
Technical Assistant:	Valerie Bostwick
Inclusive Dates of Internship:	May 28-August 23, 2012

Abstract

An internal validation of the Applied Biosystems' 3500XL Genetic Analyzer was started for the DNA Forensic Biology Laboratory of the New York City Office of Chief Medical Examiner in New York City, NY. The Applied Biosystems' AmpFℓSTR® Identifiler® Plus Amplification Kit was chosen to validate the instrument since the lab is currently validating this kit for low copy samples at twenty-nine and thirty-one cycles at half reactions. This amplification kit has shown higher sensitivity as opposed to the Applied Biosystems' AmpFℓSTR® Identifiler® Amplification Kit given that the master mix was developed to overcome inhibition.

Using exemplar samples, a sensitivity, threshold, and concordance study were performed using the Identifiler® Plus Kit on the Applied Biosystems' 3500XL Genetic Analyzer. The samples were run on an Applied Biosystems' 3130XL Genetic Analyzer as well in order to gauge the performance of the new instrument. The studies performed determined that the 3500XL Genetic Analyzer had higher sensitivity than the 3130XL Genetic Analyzer regarding peak heights and the amount of DNA needed to obtain a full profile. The threshold study showed that individual dye channel thresholds should be selected. When the same samples were ran on the 3130XL Genetic Analyzer and the 3500XL Genetic Analyzer, a concordance study verified that the correct profile was obtained for all the samples.

Introduction

Laboratories processing casework have suffered from a daunting amount of back log in cases. In order to alleviate some of the stress produced from the back log, the forensic community has been developing new processing techniques and instruments that are more effective. A new technique or instrument meant to be used for casework must be validated before it can be implemented in a laboratory's protocol for analyzing case samples. When a product is

newly released from manufacturers like Applied Biosystems or Promega, the company has already conducted a developmental validation to secure the optimal performance of the product. The product must then be subject to an internal validation for the product's use in the laboratory. An internal validation includes many of the studies done in a developmental validation, but a laboratory must conduct its own independent validation to find the optimal parameters using the protocols that will be implemented for the product for the individual laboratory.

The 3500XL Genetic Analyzer is a new series of genetic analyzer that has been released by Applied Biosystems for the use of generating DNA profiles from processed case samples. The previous model, the 3100 series has been discontinued by Applied Biosystems since June 2011, but will officially continue support until 2016.[1] Therefore, parts will be scarce and trying to get an Applied Biosystems representative to perform necessary maintenance will be hard to accomplish. Therefore, as with all new instruments or techniques used in the forensic community, the 3500XL Genetic Analyzer needs to be internally validated by individual laboratories transitioning to the new instrument before it can be used in casework.

There are many differences between the two genetic analyzers, the 3500XL Genetic Analyzer and 3130XL Genetic Analyzer being compared in this validation study. The 3500XL Genetic Analyzer has increased the number of capillaries used, allowing one to run more samples per injection while reducing the time of a run.[2] The energy being used by the instrument has also been reduced by decreasing the power supply necessary to run the instrument. The 3500XL Genetic Analyzer also makes use of consumables containing RFID (Radio Frequency Identification) tagged labels which allows for automated tracking of lot numbers and expiration dates.[4] The GS600 LIZ size standard with the option of normalization will allow the user to correct signal variations between instruments and capillaries.[3] The 3500XL instrument has a

much higher off-scale limit than the 3130XL Genetic Analyzer, making it more sensitive and will likely reduce the number of re-injections a laboratory must run to obtain a full profile. The 3130XL Genetic Analyzer can reach peaks as high as 30,000 Rfus (relative frequency units) with an optimum around 10,000 Rfus, whereas the 3500XL instrument has a height cap of about 60,000 Rfus with optimal peak heights reaching about 30,000 Rfus.

The instrument also uses a remodeled software, GeneMapper ID-X version 1.2 rather than GeneMapper ID version 3.2.1, to analyze the data which has been developed to be more user friendly. The GeneMapper ID-X version 1.2 Software possesses a layout that is very similar to that of the GeneMapper ID version 3.2.1 Software so that the transition to the ID-X version will be a smooth process. After selecting an analysis method, panels, and you hit the play button, there will be an "Analysis Requirement Check" which states if one has forgotten to assign any parameters to a sample before analysis of the run is completed.[6] The quality assessment of the allelic ladders is also done before the analysis of the injection and will exclude low quality "AL" (allelic ladder) samples.[6] Once analysis has completed, the "Analysis Summary" window will be displayed which will notify the analyst of any quality flags.[6] The program also allows one to take advantage of tools like the "Label Edit Viewer" tool for tech reviews, a "Concordance Evaluation" tool so you can see the percent match between samples that have been analyzed, and a "Report Manager" tool which will display the allele calls for the samples.[6] It has the capabilities of an expert system, but the user has to import the parameters needed for the program to analyze the system. Although it can be used as an expert system, the software is not used as such since it is against the normal forensic community practice.

Materials and Methods

The samples have to be placed through the normal DNA workflow of extracting the DNA from the oral swab, quantifying the amount of DNA present in the extract, amplifying an aliquot

of the extract, running the amplified product through the genetic analyzer, and analyzing the data from the genetic analyzer. Reference samples used for the validation studies performed had already been extracted by a member of the research group, Desarae Harmon, and needed to be quanted a minimum of three times to average the neat value to make an accurate dilution of the extract to reach the target concentrations of the sensitivity study. The reference samples used can be seen in Table 1. The samples were quanted using the Qiagen Rotor-Gene Q by trained members of the research and high sensitivity testing group. After acquiring the average neat value for the reference samples, the samples were diluted to reach three different target concentrations that would be amplified using half reactions of the AmpF*l*STR Identifiler Plus® Kit. The amplified product was then ran on the 3130XL and 3500XL genetic analyzers using different run parameters and amount of sample injected for the run. The run parameters used for the validation studies can be viewed in Table 2. First, the 3500XL had to be tested to make sure that it was functioning properly and this was done through a quality control test.

Quality Control

Before conducting the validation studies, the instrument first has to be tested to make sure that the instrument was running correctly. In order to achieve this, an injection of size standard is ran. If the electropherograms produced are clean and the size standard is called for all samples the instrument has passed. An injection of master mix, 0.3µL of LIZ and 8.7µL of Hi-Di for each sample, was plated, dispensing 9µL of the master mix into the 24 wells that make up an injection. Once the master mix was dispensed, the plate was covered with a septa, centrifuged, and put through the denaturation/cooling protocol on the thermocycler. The plate was placed on the 3500XL Genetic Analyzer and ran at I.

Sensitivity Study

For the sensitivity study, the five buccal samples chosen from laboratory personnel and the positive control of the AmpF*l*STR Identifiler Plus® Kit were quanted a minimum of three times. The average of these quants was then taken and used to make dilution sheets that would take care of the higher end of a dilution series. The averages of the Rotor-Gene values can be seen in Table 3. The samples were amped for 100pg, 250pg, and 500pg while the positive control was held at a 200pg target concentration. The 100pg and 250pg samples were prepared by first making a 1:100 dilution of the extracted stock solution. This was done by making a serial dilution, pipetting 2µL of extracted DNA into 18µL of irradiated water and then vortexing and spinning down this dilution. Then, 2µL of this newly made solution was aliquoted and dispensed into 18µL of irradiated water to complete the 1:100 dilution. A final dilution was made with a desired amount of the 1:100 diluted solution and water to achieve the target value of DNA. For the 500pg target value, a 1:20 dilution was made for each of the five extracted stock solutions, before making a final dilution with a desired amount of the 1:20 dilution and water to create the 500pg target value. The dilution calculations for each target value of DNA can be seen in Tables 4-6.

The samples were then amped using a half reaction of AmpF*l*STR Identifiler Plus® kit at twenty-nine cycles in triplicate, using the "ID+29,60min soak" protocol on the thermocycler under the user: vm. The half-reaction consists of 5µL of master mix and 2.5µL of primers. Each amp tube received 8µL of the made master mix and 5µL of the diluted samples were dispensed into the amp tube. The thermocycler protocol can be found in Table 7.

When the amp was complete, the samples were plated on the 3130XL Genetic Analyzer (Esther) at I (1kV 22 seconds). After each injection, the samples were analyzed using the

GeneMapper ID version 3.2.1 Software to make sure that the correct profiles of each sample were obtained. If the samples had drop-out they were then re-injected at IR (5kV 20 seconds).

After analyzing the samples that were ran on the 3130XL instrument the amp products were plated to be loaded onto the 3500XL Genetic Analyzer, using the same run parameters that were used on the 3130XL instrument. The 250pg samples were the first to be plated on both instruments. Master mix for the run was made by combining 0.3µL of LIZ 600 to 8.7µL of Hi-Di for each sample being loaded. For the 3500XL Genetic Analyzer set-up of the 250pg samples, 1µL of sample was dispensed into 9µL of master mix. This plate was run at I to match the 3130XL instrument parameter.

The 250pg samples were then analyzed using the GeneMapper ID-X version 1.2 Software. The "Analysis Method" was set to "ID_Plus_Analysis" and the Panels set to "Identifiler_Plus_Panels_v1_dup". The threshold for each dye channel is currently set to 50 Rfus (relative frequency units) using this method.

Once the 250pg samples were analyzed, it was decided that the samples should be injected at 1µL and 3µL on the 3500XL Genetic Analyzer. After the 100pg and 500pg samples were run on the 3130XL instrument and found to have the correct profile, the samples were run on the 3500XL Genetic Analyzer using the new volumes. The master mix to run on the 3500XL instrument was made using 0.3µL of LIZ and 8.7µL of Hi-Di for each sample, when 1µL of sample was to be loaded. When 3µL of sample was injected, the 3µL of target concentrations were dispensed into 27µL of master mix consisting of 0.9µL of LIZ and 26.1µL of Hi-Di for each sample.

For the 100pg samples, both injection volumes were run at I and IR on the 3500XL instrument to match the 3130XL Genetic Analyzer parameters.

Analytical Threshold Study

To complete the threshold study, thirty-six negative samples were amped using half reactions of AmpF*l*STR Identifiler Plus® Kit at twenty-nine cycles with the "ID+29,60minsoak" protocol on the thermocycler. This was done by adding 5µL of irradiated water to 8µL of master mix. The half reaction master mix is made using 2.5µL of primers and 5µL of master mix per sample.

Once the amp finished, the samples were then plated for the 3500XL instrument. Two injections were made, dispensing 1µL of sample into 9µL of master mix, which consisted of 0.3µL of LIZ and 8.7µL of Hi-Di per sample. The two injections were then run at both I and IR to calculate the thresholds for both parameters.

When the runs finished, the data was analyzed using the "Threshold Study" analysis method on the GeneMapper ID-X version 1.2 Software which has all the dye channels, except for orange dye channel threshold set to 1 Rfu. The orange dye channel is set at 50 Rfus. The sizing information was then exported for each injection and calculations were made to find the analytical threshold for each parameter using the negative samples. The size standard base pairs were used and peaks that fell out of a ±2 base pair range were excluded from the data used to calculate the analytical thresholds for the two run parameters on the 3500XL instrument. The IUPAC and SWGDAM method were used to do this.

IUPAC is the International Union of Pure Applied Chemistry which introduced this method to calculate analytical thresholds of a dye channel in 1995. The method uses a 1-sided significance test with the thirty-six samples at a ninety-nine percent confidence level. The equation used for this method is as follows: $AT_{M2} = \bar{Y}_{bl} + t_{1-\alpha,v}(s_{bl} / \sqrt{n})$ where AT_{M2} is the calculated analytical threshold, $\bar{Y}_{bl}$ is the average blank Rfu signal, $t_{1-\alpha,v}$ is the t-value from the student t-table given the number of samples being measured, and $s_{bl} / \sqrt{n}$ is the estimated

standard deviation of the next signal when x=0 for the negative samples used.[7] Since thirty-six negative samples were used for this calculation method, the t-value extracted from the student t-table is 2.434 at the ninety-nine percent confidence interval.

SWGDAM stands for the Scientific Working Group on DNA Analysis Methods. The group stated that an "analytical threshold may be used on two times the intensity difference between the highest peak and lowest trough within the instrumental noise data".[7] The equation using this statement is therefore, $AT = 2(Y_{max} - Y_{min})$, where AT is the calculated analytical threshold, Y_{max} is the highest peak within the instrumental noise data, and Y_{min} is the signal of the lowest trough. [7]

Both methods were used to calculate an analytical threshold. The results were compared to determine the analytical threshold for the analysis method used when analyzing data with the Genemapper ID-X version 1.2 Software.

Concordance Study

The samples used to accomplish the sensitivity study were used to evaluate the concordance of the 3500XL Genetic Analyzer to the 3100XL Genetic Analyzer using the same AmpF*l*STR Identifiler Plus® Kit. The 250pg samples were plated on the 3500XL with 1µL of sample dispensed into 9µL of master mix, consisting of 0.3µL of LIZ 600 to 8.7µL of Hi-Di. This plate was ran at I to match the 3130XL instrument run parameter. As mentioned previously in the sensitivity study, given the height of the peaks that were observed with the 250pg samples, it was decided to follow the plate set-up protocol enforced for the 3130XL Genetic Analyzer for the 100pg and 500pg samples as well. Therefore, in addition to plating 1µL of sample dispensed into 9µL of the LIZ600 and Hi-Di master mix, 3µL of sample were dispensed into 27µL of master mix consisting of 0.9µL of LIZ and 26.1µL of Hi-Di. These samples were ran at I and IR on the 3500XL Genetic Analyzer to match the parameters used on the 3130XL Genetic

Analyzer. The profiles generated for each sample on both genetic analyzers were then compared to identify if both instruments produced the same allele calls for each locus for each sample.

Results

Quality Control

The first LIZ plate injection ran on the 3500XL Genetic Analyzer was the 3500_072512_1 plate. When the results were viewed using both the 3500 Data Collection Software and the GeneMapper ID-X version 1.2 Software there was a flag for sample 8 injected with capillary 22. The spectral and spatial calibrations on file were referenced to see if there is an issue with capillary 22, but it passed and capillary 22 did not borrow from another capillary. Therefore, it must have been the injection. The plate was then re-injected as plate 3500_072612_2 to test the instrument. A bubble was detected in the instrument so the run was canceled. This injection failed. The plate was then re-injected for a third and final time as plate 3500_072612_3. The injection was completed and analyzed using the GeneMapper ID-X version 1.2 Software. Analysis of the injection, showed the calls expected for the GeneScan LIZ 600 Size Standard were found for all 24 samples. The other dye channels were completely clean, showing no contamination.

Sensitivity Study

When the 250pg and 500pg samples were ran using the I run parameter, the expected profiles for each sample were obtained when ran on both the 3130XL and 3500XL instruments. The 100pg samples had to be run at I and IR at the different volumes on the 3500XL Genetic Analyzer since they were ran at I and IR on the 3130XL instrument. When ran at I on the 3130XL Genetic Analyzer, some samples had drop-out so the plate was re-injected at IR to pull up the peaks that were present, but not called. Once this was done, full profiles were achieved at the higher run parameter. When inspecting the profiles generated by the 3500XL Genetic

Analyzer, each sample had the appropriate allele calls with no drop-out present, unlike the samples run on the 3130XL instrument.

To evaluate the sensitivity of the two instruments, the peak heights presented in the three target concentration profiles generated from both the 3130XL and 3500XL genetic analyzers were compared. The minimum, maximum peak heights, and average peak heights were compared for each run parameter on the two instruments as well as the peak height ratio and average peak height ratio. Samples that displayed drop-out were excluded from the calculations used to evaluate the sensitivity of the instruments.

The observations discovered for both instruments differed when the target concentration and run parameter were changed. When the samples were run on the 3130XL Genetic Anlyzer at I using only 1μL of sample, the 100pg samples had significant drop-out which excluded some samples from the calculated sensitivity. It was seen that the 500pg samples had a higher average peak height ratio while having a smaller peak height standard deviation and a narrower range between the minimum and maximum peak height when compared to the other target concentrations ran under these same conditions. The average peak height ratio for the 500pg samples is 12.93% higher than the average peak height ratio observed for the 100pg samples. The percent peak height ratio standard deviation and range between the maximum and minimum peak height ratio was smaller for the 500pg than those found for the 100pg, with a 7.75% and a 34.22% difference being respectively seen. The sensitivity calculations generated for the 3130XL instrument ran at I with 1μL of samples dispensed, can be found in Table 8. Calculated values for the 250pg samples were in between the values calculated for the 100pg and 500pg samples. When analyzing the peak heights and peak height ratios between sister alleles, one usually expects a 70% peak height ratio between sister alleles for the peaks to be considered

balanced. It can be as low as 50%, depending on the laboratory's protocols. While calculating the peak height ratios for the samples ran at the different target concentrations and run parameters, the number of peak height ratios that had a peak height balance of 70% or greater was also evaluated for each run on both the 3100XL and 3500XL genetic analyzers. The peak height balance calculations can be found in Table 9 for the samples ran at I on the 3130XL using 1μL of sample.

Samples ran at I on the 3500XL Genetic Analyzer using 1μL of sample, displayed similar relationships as seen as those on the 3130XL Genetic Analyzer under the same run parameter. The 100pg samples did not have any drop-out, increasing the number of loci used for its peak height calculations, therefore achieving a more accurate representation of the relationship between the target concentrations ran at this parameter. As seen in the 3130XL instrument results, the 500pg samples had a higher average peak height ratio while having a smaller peak height standard deviation and a narrower range between the minimum and maximum peak height when compared to the other target concentrations ran under these same conditions. The 500pg samples showed an average peak height ratio 14.94% higher than the average peak height ratio calculated for the 100pg samples. Just like the 3130XL Genetic Analyzer results, the percent peak height ratio standard deviation and range between the maximum and minimum peak height ratio was smaller for the 500pg than those found for the 100pg. The standard deviation encountered by the 500pg samples was 9.54% smaller than the standard deviation found for the 100pg samples and the range for the 100pg samples was 32.19% greater than the range found for the 500pg samples. The sensitivity calculations for the 3500XL instrument ran at I using 1μL of sample can be seen in Table 10 while the peak height balances can be seen in Table 11.

Samples with a target concentration of 100pg and 500pg were also ran using different conditions. Since the 100pg samples displayed drop-out when ran in the 3130XL Genetic Analyzer, the samples were ran at an IR parameter on both the 3130XL and 3500XL instruments. The 100pg and 500pg samples were also ran at I using 3μL of sample as is stated in the protocol currently practiced at this laboratory. The sensitivity results for both the IR and 3μL of sample parameters for these target concentrations can be found in Tables 12-15. When evaluating the 100pg samples under the I and IR run parameters on the 3500XL Genetic Analyzers, it can be found that there is no significant difference between the average peak height ratio and the percentage of sister alleles with a peak height ratio at or above 70%. However, there is a significant difference between the peak height values themselves at the different run parameters. The same relationships found in the different run parameters can also be viewed with the increase of sample injected into the 3500XL instrument when the 1μL and 3μL injections are compared.

Analytical Threshold Study

The thirty-six negative samples were analyzed using the GeneMapper ID-X version 1.2 Software. It was found that when run under the I parameter, the red channel possesses the highest analytical threshold for both the IUPAC and SWGDAM methods, with the threshold being about 38 and 128 Rfus, respectively, for the 3500XL Genetic Analyzer. When the data was calculated using the IR parameter, the red dye channel has the higher analytical threshold at 11 Rfus when calculated using the IUPAC method, whereas the green dye channel has the higher analytical threshold at 128 Rfus when the SWGDAM method is used. The calculations for both the IUPAC and SWGDAM methods can be seen for the 3500XL Genetic Analyzer runs at I and IR in Tables 16-19.

Concordance Study

When the profiles generated by the 3130XL and 3500XL instruments for the samples were compared to the known profiles of the reference samples used it was discovered that there was a 100% concordance. The appropriate allele calls were found for all samples, giving exact matches.

Discussion

Quality Control

Since all the required peaks were found in the LIZ standard for a complete injection on the 3500XL, the instrument has passed the quality control test and the validation studies can proceed. The instrument had to pass this test in order to ensure that the instrument itself, is working as it should. If it had not passed the test, Applied Biosystems would be consulted to fix the issue.

Sensitivity Study

The peak heights found in the 3500XL Genetic Analyzer profiles were consistently significantly higher that the peak heights found in the 31300XL Genetic Analyzer profiles for each sample. This was found for samples ran at both I and IR as well. The increase in the peak heights for the 3500XL Genetic Analyzer signifies that the instrument is much more sensitive than the 3130XL Genetic Analyzer. Since the 3500XL instrument has proven to be more sensitive when the peak heights are analyzed, it can be used when samples are deemed to have too low of a concentration and one is afraid of drop-out occurring if the samples were ran on the 3130XL instrument. The use of the 3500XL Genetic Analyzer would be more cost effective since there would be no need to try to concentrate the sample or re-inject the sample at a higher parameter as one would need to do so using the 3130XL Genetic Analyzer. An increase in the amount of sample injected also showed an increase in the peak heights observed. It is expected that as the amount of sample increases so will the peak heights seen in the profile generated. This

was clearly seen when the 1µL and 3µL injections were compared. This can also be seen in the target concentrations, with the 500pg samples having the highest peaks, the 100pg samples displaying the lowest peak heights, and the 250pg samples showing an intermediate of the other two target concentrations. Although, when the peak height ratios for the different parameters ran on both instruments were calculated, it was found that there was not a significant difference between the instruments. The peak height ratios and the peak balances were relatively close. Both instruments have a strong peak height balance for the profiles generated under different parameters.

Analytical Threshold

Both the IUPAC and SWGDAM methods produced different values for the analytical threshold for the 3500XL instrument ran at I and IR for the different dye channels. Under I parameters, both methods calculated that the red dye channel had a higher analytical threshold than the other dyes. Whereas, under IR parameters the IUPAC method calculated the red dye channel as having the highest analytical threshold, while the SWGDAM method found that the green dye channel possessed the higher analytical threshold value when compared to the other dye channels. Given the varying results, specific dye channel thresholds should be implemented when analyzing data rather than one analytical threshold for all of the dye channels.

Concordance Study

Since a 100% concordance was found when comparing the profiles generated from the 3500XL instrument to the profiles produced by the 3130XL Genetic Analyzer, the laboratory can safely transition analyzing samples with the 3500XL Genetic Analyzer from the 3130XL Genetic Analyzer without worrying that the instrument will adversely affect the profiles generated.

Conclusions and Future Work

Sensitivity Study

The sensitivity study performed was on the high end of the target concentration spectrum. It proved that the 3500XL Genetic Analyzer is more sensitive than the 3130XL Genetic Analyzer when the individual peak heights of a sample are compared to one another, but when the peak height ratios of the samples were compared between the instruments, no significant difference could be noted. The average peak height ratios for both the I and IR run parameters as well as the 1μL and 3μL injections were within a percentage of one another. Since the high end of the target concentration spectrum was tested, a set of lower dilutions should be done to evaluate and compare the performance of the two instruments when the target input reaches the low end of the spectrum. Maybe a 50 pg, 25pg, 12.5pg, and 6.25 pg could be tested to sufficiently capture the lower end of the target concentration spectrum. With this, the sensitivity study can adequately attain the limitation of both instruments as to how low a target input can be and still obtain a full profile from a sample. Given the results from the high end dilutions of the sensitivity study performed and the marketing points of Applied Biosystems regarding the 3500XL Genetic Analyzer, the newer generation will be more sensitive, having a lower target input range than the 3130XL Genetic Analyzer.

Analytical Threshold Study

The threshold study used both the IUPAC and SWGDAM guidelines for calculating an analytical threshold. The results varied between the different dye channels. Analyzing the results of the study, it would suggest that each dye channel have its own threshold rather than having an all encompassing threshold for all dyes. This threshold study only used thirty-six negatives to calculate the analytical threshold, therefore more samples could be tested to see if there are any significant changes with more data. The standard threshold most labs adhere to, that have validated the 3500XL Genetic Analyzer, is 150Rfus, but like all internal validations, this lab

must conduct a threshold study to evaluate the analytical threshold that their protocol will use for everyday casework.[1]

Concordance Study

The concordance study completed between the two generations of the genetic analyzer instrument demonstrated that the profiles obtained for all samples ran at both parameters were correct. All expected allele calls were called when using the AmpF*l*STR Identifiler Plus® Kit. In addition to this concordance study, one can be done using the 3500XL Genetic Analyzer to compare the results achieved between different amplification kits. Applied Biosystems also has an AmpF*l*STR Profiler® Kit, an AmpF*l*STR Cofiler® Kit and an AmpF*l*STR Identifiler® Kit which can be compared to the AmpF*l*STR Identifiler Plus® Kit results. Since the OCME primarily uses the AmpF*l*STR Identifiler® Kit and is in the middle of an extensive validation for AmpF*l*STR Identifiler Plus® Kit, it would benefit the lab to conduct a concordance study with these two kits on the 3500XL Genetic Analyzer. The kits can then be ran at the different parameters used in this concordance study, I and IR, or the new parameters being established for the AmpF*l*STR Identifiler Plus® Kit through the validation in progress on the 3130XL Genetic Analyzer, such as IPR, which is a 5kV and 25 second injection.

Future Work

Traditionally, an injection time and target study is conducted in the preliminary phase of a validation and since one was not done yet, it can still be ran to find the optimal parameters of the 3500XL Genetic Analyzer. This can be accomplished by using the five quanted samples used in the studies thus far. A dilution series with eight individual dilution points will be made and amplified using at least three different cycle numbers. Normal DNA levels use 28 and 31 cycles as the cycle number, but for samples needing high sensitivity conditions a cycle number of 29 or 32 is used, with the latter of each set used when a sample shows signs of degradation/inhibition.

The samples will be amplified and ran on the 3500XL Genetic Analyzer multiple times at different voltages and for various lengths of time to identify the optimal amplification load, thermal protocol, injection time, and injection voltage for normal and high sensitivity levels of the samples tested.

A precision study needs to be run using the amplified product to test the sizing of the allelic ladder and make sure that the standard deviation of the base pair sizes is lower or equal to 0.15 to ensure accurate profiles.[5] This can be done by setting up an injection of allelic ladders and running them at the optimal injection parameters discovered in the injection study for normal and high sensitivity conditions. Using the 3500XL Genetic Analyzer plate set-up software, the injection is then ran multiple times. The base pairs present within the samples are then analyzed .

To further progress the validation of the 3500XL Genetic Analyzer using the AmpF*l*STR Identifiler Plus® Kit, a reproducibility study also needs to be completed. This study is performed to verify if the profiles obtained are consistent and reliable because a set of samples will be run various times on different days and the results will be compared. A reproducibility study can be achieved by running a plate that has been set up for the 3500XL instrument three different times in three consecutive days since it is the OCME's policy that a plate set-up for capillary electrophoresis cannot be run if it is older than three days. The sample peak height, size in base pairs, and allele call consistency between different 3500XL instrument runs will be compared to one another. To test the durability of the amplified product and ascertain results that are par for the norm for real cases, the plate of amplified DNA samples can be set-up for capillary electrophoresis multiple times along a span of weeks or months. This way the reproducibility study can obtain a more realistic view of as to how long the profiles will remain full, consistent profiles. The results can then be used for a durability/sustainability test as well.

A mixture study also needs to be completed. It involves making set mixtures of the samples and testing to see at which ratio the minor component can still be detected. A simple mixture study would involve two of the five samples mentioned before. The two samples can then be amped at a low and high target value. Then, set ratio samples will be made, such as 20:1, 10:1, 5:1, 2:1, 1:1, 1:2, 1:5, 1:10, and 1:20, that will be run on the 3500XL Genetic Analyzer to see if the minor and major contributors can be differentiated from one another in the resulting profiles. The allelic drop out observed in each mixture will be noted. Male-male, male-female mixtures should be prepared using the different ratios. For a more complex mixture study, the number of contributors can be increased to see if they can be distinguished from one another and at what ratios allelic drop out can be identified.

A non-probative study can also be run using a variety of sample types that the laboratory commonly encounters during a case. This can include buccal, semen, blood, and saliva samples which this laboratory handles. The NIST panel of samples can also be used to verify the accuracy of the instrument.

A contamination study can be run using the amplification, reagent blanks, and run negatives to ensure that all the samples ran on the 3500XL are free of any corruption. The samples ran in previous studies can be used for this study. One just has to note if contamination is present in the profiles obtained from those runs. No contamination was seen in the samples ran, but an official study needs to be conducted.

A stutter study can be ran to determine the optimal marker specific stutter ratios by evaluating the stutter percentages at each marker. The samples ran in the sensitivity, reproducibility, and mixture studies can be used to calculate the percentages.

A heterozygosity study can be run. In order to do so, samples that are heterozygous at every marker presented in the AmpF*l*STR Identifiler Plus® Kit are ideal to evaluate the minimum acceptable peak height ratios for sister alleles.

These studies should be done to complete the validation of the 3500XL Genetic Analyzer. Once these studies are done and the necessary data obtained, the new instrument can be implemented into a protocol for processing everyday casework samples.

References

1. Butts, Erica L.R., Carolyn R. Hill, Margaret C. Kline, David L. Duewer, John M. Butler, Kristen L. O'Connor, Michael D. Coble, and Peter M. Vallone. "NIST Validation Studies of the 3500 Genetic Analyzer." N.p., 27 May 2011. Web. <http://www.cstl.nist.gov/strbase/pub_pres/MAAFS2011_3500validation.pdf>.
2. "3500 Series Genetic Analyzers." *3500 Series Genetic Analyzer*. Life Technologies, 2011. Web. <http://marketing.appliedbiosystems.com/mk/get/3500SERIES_MLANDING>.
3. "HID Validation: The 3500 Series Genetic Analyzers." Life Technologies, 2009. Web. <http://www3.appliedbiosystems.com/cms/groups/applied_markets_marketing/documents/generaldocuments/cms_093152.pdf>.
4. "3500 Series Genetic Analyzer: Frequently Asked Questions." *3500 Series Genetic Analyzer*. N.p., 2011. Web. <http://marketing.appliedbiosystems.com/mk/get/3500_HID_FAQS>.
5. Butler, John M., PhD. "Developmental Validation." NIST, 24 Aug. 2005. Web. <http://www.cstl.nist.gov/strbase/validation/DevelopmentalValidation.pdf>.

6. “Applied Biosystems 3500/3500xLGenetic Analyzer User Guide”. *User Guide*. N.p., 2010. Web. <http://www3.appliedbiosystems.com/cms/groups/mcb_support/documents/generaldocuments/cms_069856.pdf>
7. “Threshold Study”. *MUFSC Validation CD*. 2012.

Acknowledgements

I would like to thank Dr. Mecke and everyone at the New York City Office of Chief Medical Examiner for allowing the lab to host a Marshall University student from the Master of Forensic Science Program. I would like to acknowledge Dr. Theresa Caragine for accepting the internship proposal for this validation during the summer and for being a mentor throughout this project. I would also like to thank Dr. Caragine for her comments on this paper. Grace Axler-Diperete was a welcoming intern supervisor. Cindy Rodriquez played an instrumental role in the project, providing training in the techniques performed at this laboratory, guiding the project along from its start to the end of the internship, and for being a sounding board when issues presented themselves. Desarae Harmon also helped with training, provided examples of the paperwork used at the OCME, answered any and all questions concerning the proper protocols taken when conducting a validation, and was also a great sounding board. Without the assistance of Juliette Smith and her willingness to stay late, this project would have never started and I am eternally grateful for all her help. She is the fairy godmother of this project. I would also like to thank the rest of the high sensitivity/research group whom were all friendly and willing to help with the project. I also wish to thank the other interns in the group that directly and indirectly contributed to the progress of the validation.

I would like to gratefully acknowledgment Dr. Pamela Staton for being the Marshall University Topic Advisor of this project. Her input and understanding of the issues on this project were very helpful. I would also like to thank Valerie Bostwick for her comments and support of this project; Justin Godby for his informative and entertaining teaching techniques in the Technical Assistance Program courses that served as preparation for this validation; and Ashleigh Zeile for her insight on the project.

Tables, Charts, and Diagrams

Reference Samples	
Sample	**Sex**
PE	Female
D18	Female
D19	Male
D42	Female
D49	Male
D102	Female

Table 1. The laboratory personnel samples chosen.

Run Parameters for the 3100XL and 3500XL for the AmpF*l*STR Identifiler Plus® Kit at 29 Cycles			
Amp System/Cycle	**Specification**	**Run Module Code**	**Parameters**
ID+ 29	**Normal**	**I**	**1kV 22sec**
	High	**IR**	**5kV 20sec**

Table 2. Capillary electrophoresis run parameters for AmpF*l*STR Identifiler Plus® Kit at 29 cycles.

Rotor-Gene Quant Values for the Reference Samples Chosen for the Validation Studies											
Samples	**RG15Q071812.1500**	**RG7Q071812.1700**	**RG15Q071912.1230**	**RG7Q072312.1200**	**Donor**	**Quant Ave.**	**Neat Ave.**	**STDEV**	**MIN**	**MAX**	**RANGE**
PE_a	205.05	188.2		160.77	PE_ID+	177.04	177.04	26.66	138.45	214.89	76.44
PE_b	199.93	138.45		155.6							
PE_c	214.89	177.99		152.44							
D18_a	98.89	85.58		97.47	18	92.29	9229.22	8.73	81.56	106.76	25.2
D18_b	106.76	86.77		85.96							
D18_c	100.68	86.96		81.56							
D19_a	74.23	99.14		81.2	19	85.85	8585.00	9.31	74.23	99.14	24.91
D19_b	97.11	90.55		76.39							
D19_c	91.74	85.79		76.5							
D42_a	130.87	117.73		139.19	42	131.62	13162.33	7.54	117.73	141.38	23.65
D42_b	138	126.49		125.9							
D42_c	141.38	130.58		134.47							
D49_a	33.32	35.04		32.01	49	32.08	3208.00	2.46	26.5	35.04	8.54
D49_b	32.66	32.58		30.87							
D49_c	31.51	34.23		26.5							
D102_a		78.5	82.55	59.03	102	71.74	7174.00	9.04	59.03	82.55	23.52
D102_b		65.44	79.41	62.43							
D102_c		67.89	78.67								

Table 3. The Rotor-Gene values for the quants of the reference samples chosen for the validations studies being performed on the 3500XL. The cells highlighted in purple are the average values used to perform the dilution sheets for each target value.

100pg Dilution Table									
Starting DNA Conc. [pg/uL]	**Conc. Of Dilution (1:100)**	**Target DNA Conc. [pg total]**	**Conc. Needed for Identifiler reactions [pg/uL]**	**Desired Total Volume [uL]**	**Required DNA to Achieve Desired Volume [uL]**	**Required Water to Achieve Desired Volume[uL]**	**Tube labels for each sample per amplification set.**		
177.04	-	200	40	20.00	**4.52**	**15.48**	PE1		
9229.22	92.2922	100	20	20.00	**4.33**	**15.67**	D18a100	D18b100	D18c100
8585	85.85	100	20	20.00	**4.66**	**15.34**	D19a100	D19b100	D19c100
13162.33	131.6233	100	20	20.00	**3.04**	**16.96**	D42a100	D42b100	D42c100
3208	32.08	100	20	20.00	**12.47**	**7.53**	D49a100	D49b100	D49c100
7174	71.74	100	20	20.00	**5.58**	**14.42**	D102a100	D102b100	D102c100

Table 4. Dilution calculations for the 100pg target value. Yellow boxes represent the neat concentration while the orange boxes represent the titrated concentration.

250pg Dilution Table									
Starting DNA Conc. [pg/uL]	Conc. Of Dilution (1:100)	Target DNA Conc. [pg total]	Conc. Needed for Identifiler reactions [pg/uL]	Desired Total Volume [uL]	Required DNA to Achieve Desired Volume [uL]	Required Water to Achieve Desired Volume[uL]	Tube labels for each sample per amplification set.		
177.04	-	200	40	10.00	**2.26**	**7.74**	PE1		
9229.22	92.2922	250	50	30.00	**16.25**	**13.75**	D18a250	D18b250	D18c250
8585	85.85	250	50	30.00	**17.47**	**12.53**	D19a250	D19b250	D19c250
13162.33	131.6233	250	50	30.00	**11.40**	**18.60**	D42a250	D42b250	D42c250
3208	320.8**	250	50	30.00	**4.70**	**25.30**	D49a250	D49b250	D49c250
7174	717.4**	250	50	30.00	**2.09**	**27.91**	D102a250	D102b250	D102c250

Table 5. Dilution calculations for the 250pg target value. Yellow boxes represent the neat concentration while the orange boxes represent the titrated concentration.

500pg Dilution Table									
Starting DNA Conc. [pg/uL]	Conc. Of Dilution (1:20)	Target DNA Conc. [pg total]	Conc. Needed for Identifiler reactions [pg/uL]	Desired Total Volume [uL]	Required DNA to Achieve Desired Volume [uL]	Required Water to Achieve Desired Volume[uL]	Tube labels for each sample per amplification set.		
177.04	-	200	40	20.00	**4.52**	**15.48**	PE1		
9229.22	461.461	500	100	20.00	**4.33**	**15.67**	D18a500	D18b500	D18c500
8585	429.25	500	100	20.00	**4.66**	**15.34**	D19a500	D19b500	D19c500
13162.33	658.1165	500	100	20.00	**3.04**	**16.96**	D42a500	D42b500	D42c500
3208	160.4	500	100	20.00	**12.47**	**7.53**	D49a500	D49b500	D49c500
7174	358.7	500	100	20.00	**5.58**	**14.42**	D102a500	D102b500	D102c500

Table 6. Dilution calculations for the 500pg target value. Yellow boxes represent the neat concentration while the orange boxes represent the titrated concentration.

Thermocycler Protocol for the AmpF*l*STR Identifiler Plus® Kit at 29 Cycles				
Initial	**Number of Cycles: 29**		**Final**	**Final**
Incubation	**Denature**	**Anneal/Extend**	**Extension**	**Hold**
Hold	**Cycle**		**Hold**	**Hold**
95°C	**94°C**	**59°C**	**60°C**	**4°C**
11 min	**20sec**	**3 min**	**60 min**	∞

Table 7. Thermocycler protocol used to amp the samples used in the sensitivity study.

Sensitivity Study Results for the 3130XL ran at I with 1µL of Sample									
DNA input	Maximum Pk Ht	Minimum Pk Ht	Average Pk Ht	Maximum PHR	Minimum PHR	Average PHR	% PHR Standard Deviation	Range between Min & Max	# of loci used for PHR Calculation
100pg	1991	80	362	99.87%	16.10%	71.09%	19.17%	83.78%	180
250pg	3872	206	1006	99.45%	32.76%	76.51%	15.15%	66.69%	192
500pg	7632	465	2022	99.93%	50.37%	84.02%	11.42%	49.56%	192

Table 8. The peak heights, peak height ratios, standard deviation of the peak height ratio, and the range of the peak heights calculated for the samples ran on the 3130XL at I. The cells highlighted in pink are the average peak heights found at the different target concentrations whereas the cells highlighted in yellow display the average peak height ratios. Cells highlighted in orange note the presence of drop-out and homozygous alleles.

Peak Height Ratios for the Target Concentrations Ran at I on the 3130XL Using 1µL of Sample							
DNA input	PHRs <50%	PHRs 50-59%	PHRs 60-69%	PHRs 70-79%	PHRs 80-89%	PHRs 90-100%	TOTAL OVER 70%
100pg	15.6%	8.3%	20.6%	18.3%	18.9%	18.3%	55.6%
250pg	5.7%	6.8%	15.1%	25.5%	24.5%	22.4%	72.4%
500pg	0.0%	3.1%	9.9%	15.4%	27.1%	40.6%	83.1%

Table 9. Peak height balances found between sister alleles for the target concentrations ran at I on the 3130XL using 1µL of sample. The cells highlighted in yellow show the percentage of the results that display a peak height ratio of 70% or more.

Sensitivity Study Results for the 3500XL ran at I with 1µL of Sample									
DNA input	Maximum Pk Ht	Minimum Pk Ht	Average Pk Ht	Maximum PHR	Minimum PHR	Average PHR	% PHR Standard Deviation	Range between Min & Max	# of loci used for PHR Calculation
100pg	7634	124	1476	99.26%	14.39%	68.97%	20.95%	84.88%	191
250pg	12126	943	3575	99.81%	33.35%	76.12%	15.16%	66.46%	192
500pg	34541	2104	8985	99.91%	47.22%	83.91%	11.41%	52.69%	192

Table 10. The peak heights, peak height ratios, standard deviation of the peak height ratio, and the range of the peak heights calculated for the samples ran on the 3500XL at I. The cells highlighted in pink are the average peak heights found at the different target concentrations whereas the cells highlighted in yellow display the average peak height ratios. Cells highlighted in orange note the presence of homozygous alleles.

Peak Height Ratios for the Target Concentrations Ran at I on the 3500XL Using 1µL of Sample							
DNA input	PHRs <50%	PHRs 50-59%	PHRs 60-69%	PHRs 70-79%	PHRs 80-89%	PHRs 90-100%	TOTAL OVER 70%
100pg	19.4%	7.3%	19.4%	18.3%	17.3%	18.3%	53.9%
250pg	5.7%	8.9%	15.6%	24.5%	23.4%	21.9%	69.8%
500pg	0.5%	2.1%	8.9%	17.1%	26.6%	40.6%	84.3%

Table 11. Peak height balances found between sister alleles for the target concentrations ran at I on the 3500XL using 1µL of sample. The cells highlighted in yellow show the percentage of the results that display a peak height ratio of 70% or more.

Sensitivity Study Results for 100pg ran on the 3500XL at I and IR for the 1µL and 3µL Injected									
DNA input	Maximum Pk Ht	Minimum Pk Ht	Average Pk Ht	Maximum PHR	Minimum PHR	Average PHR	% PHR Standard Deviation	Range between Min & Max	# of loci used for PHR Calculation
100pg (1µL)	7634	124	1476	99.26%	14.39%	68.97%	20.95%	84.88%	191
100pg (1µL)	9505	137	1698	99.64%	14.12%	68.78%	20.93%	85.52%	191
100pg (3µL)	16535	219	3101	99.78%	14.54%	68.90%	21.09%	85.24%	191
100pg (3µL)	19922	8891	3745	99.91%	14.74%	68.51%	20.83%	85.16%	191

Table 12. The sensitivity results calculated for 100pg samples ran at I and IR on the 3500XL using 1µL and 3µL injections. The rows highlighted in a purple-grey are the injections ran at I while the rows highlighted in aqua are the injections rat at IR. The column highlighted in pink notes the average peak heights found for the different run parameters. The column highlighted in yellow is the average peak height ratios for the run parameters.

Peak Height Ratios for 100pg ran on the 3500XL at I and IR for the 1µL and 3µL Injected							
DNA input	PHRs <50%	PHRs 50-59%	PHRs 60-69%	PHRs 70-79%	PHRs 80-89%	PHRs 90-100%	TOTAL OVER 70%
100pg (1µL)	19.4%	7.3%	19.4%	18.3%	17.3%	18.3%	53.9%
100pg (1µL)	19.9%	8.4%	17.3%	19.9%	15.7%	18.8%	54.5%
100pg (3µL)	20.4%	7.3%	19.4%	17.8%	17.3%	17.8%	52.9%
100pg (3µL)	20.4%	7.3%	17.3%	23.0%	14.1%	17.8%	55.0%

Table 13. Peak height balances found between sister alleles for the target concentrations ran at I and IR on the 3500XL using 1µL and 3µL of sample. The rows highlighted in a purple-grey are the injections ran at I while the rows highlighted in aqua are the injections rat at IR. The column highlighted in yellow shows the percentage of the results that display a peak height ratio of 70% or more.

Sensitivity Study Results for 1µL and 3µL Injected on the 3500XL at I									
DNA input	Maximum Pk Ht	Minimum Pk Ht	Average Pk Ht	Maximum PHR	Minimum PHR	Average PHR	% PHR Standard Deviation	Range between Min & Max	# of loci used for PHR Calculation
100pg (1µl)	7634	124	1476	99.26%	14.39%	68.97%	20.95%	84.88%	191
100pg (3µl)	16535	219	3101	99.78%	14.54%	68.90%	21.09%	85.24%	191
500pg (1µl)	34541	2104	8985	99.91%	47.22%	83.91%	11.41%	52.69%	192
500pg (3µl)	46587	25599	13288	99.83%	49.88%	83.98%	11.52%	49.95%	192

Table 14. The sensitivity results for the 100pg and 500pg samples ran on the 3500XL at I using 1 µL and 3µL injections. The rows highlighted in a darker purple-grey are the 1µL injections whereas the rows highlighted in a lighter grey are the 3µL injections. The column highlighted in pink notes the average peak heights found for the different injections. The column highlighted in yellow is the average peak height ratios for the injections.

Peak Height Ratios for 1µL and 3µL Injected on the 3500XL at I							
DNA input	PHRs <50%	PHRs 50-59%	PHRs 60-69%	PHRs 70-79%	PHRs 80-89%	PHRs 90-100%	TOTAL OVER 70%
100pg (1µl)	19.4%	7.3%	19.4%	18.3%	17.3%	18.3%	53.9%
100pg (3µl)	20.4%	7.3%	19.4%	17.8%	17.3%	17.8%	52.9%
500pg (1µl)	0.5%	2.1%	8.9%	17.1%	26.6%	40.6%	84.3%
500pg (3µl)	0.0%	3.1%	9.4%	16.7%	26.0%	40.6%	83.3%

Table 15. Peak height balances found between sister alleles for the target concentrations ran at I on the 3500XL using 1µL and 3µL of sample. The rows highlighted in a darker purple-grey are the 1µL injections whereas the rows highlighted in a lighter grey are the 3µL injections. The column highlighted in yellow shows the percentage of the results that display a peak height ratio of 70% or more.

IUPAC Method for 3500XL Samples Ran at I						
						n=36
DYE	AVERAGE	STDEV	MIN	MAX	AT	SQRT(36)
BLUE	6.23	2.66	1	25	7.31	6
GREEN	11.87	4.10	3	38	13.53	6
YELLOW	22.67	6.74	5	47	25.40	6
RED	33.67	9.64	10	74	37.58	6

Table 16. The IUPAC calculations for the analytical threshold of the samples ran at I on the 3500XL for the different dye channels. The highest calculated analytical threshold has been circled in a bright red hue.

SWGDAM Method for 3500XL Samples Ran at I						
DYE	AVERAGE	STDEV	MIN	MAX	Lowest Trough	AT
BLUE	6.23	2.66	1	25	1	48
GREEN	11.87	4.10	3	38	3	70
YELLOW	22.67	6.74	5	47	5	84
RED	33.67	9.64	10	74	10	128

Table 17. The SWGDAM calculations for the analytical threshold of the samples ran at I on the 3500XL for the different dye channels. The highest calculated analytical threshold has been circled in a bright red hue.

IUPAC Method for 3500XL Samples Ran at IR						
						n=36
DYE	AVERAGE	STDEV	MIN	MAX	AT	SQRT(36)
BLUE	3.48	2.38	1	24	4.44	6
GREEN	4.00	4.60	1	65	5.87	6
YELLOW	5.95	2.90	1	40	7.13	6
RED	9.18	3.75	2	34	10.70	6

Table 18. The IUPAC calculations for the analytical threshold of the samples ran at IR on the 3500XL for the different dye channels. The highest calculated analytical threshold has been circled in a bright red hue.

SWGDAM Method for 3500XL Samples Ran at IR						
DYE	AVERAGE	STDEV	MIN	MAX	Lowest Trough	AT
BLUE	3.48	2.38	1	24	1	46
GREEN	4.00	4.60	1	65	1	128
YELLOW	5.95	2.90	1	40	1	78
RED	9.18	3.75	2	34	2	64

Table 19. The SWGDAM calculations for the analytical threshold of the samples ran at IR on the 3500XL for the different dye channels. The highest calculated analytical threshold has been circled in a dark green hue.