2021

PowerPlex® Fusion 6C System versus PowerPlex® Fusion 5C: A Comparison of Performance Metrics

Caitlin M. McCaughan
Virginia Commonwealth University

PowerPlex® Fusion 6C System versus PowerPlex® Fusion 5C: A Comparison of Performance Metrics

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Forensic Science at Virginia Commonwealth University.

by
Caitlin M. McCaughan, B.S.

Susan A. Greenspoon, Ph.D., Research Mentor
Virginia Department of Forensic Science-Central Laboratory

Virginia Department of Forensic Science
Virginia Commonwealth University
Summer 2020-Spring 2021
Richmond, VA

February 16, 2021

Acknowledgments

I would like to take this opportunity to recognize the many people who assisted in the completion of this project. First, I would like to thank Dr. Susan Greenspoon for her time and advice, both of which were invaluable to this research and my education. I would like to thank my research advisor, Dr. Catherine Connon, for her knowledge and guidance throughout this project. Additionally, I would like to thank my committee member, Cathryn Shannon, for her support and help, both in the lab and with all the subsequent writings. I would also like to thank the Virginia Department of Forensic Science for providing the space, materials, and instrumentation to complete this project and the Forensic Biology and Research section, especially Felicia Brown and Bob Scanlon, for welcoming me and always helping me out. Lastly, I would like to thank the Promega Corporation for providing their Fusion 6C amplification kit for use in this research.

Table of Contents

Abstract

Following the FBI mandated expansion of the CODIS core loci from 13 to 20, several manufacturers developed short tandem repeat (STR) typing kits in response to the new criteria. One such manufacturer was the Promega Corporation, which released the PowerPlex® Fusion 5C megaplex STR typing kit (Fusion 5C) in 2012. Currently, the Virginia Department of Forensic Science (VDFS) utilizes this amplification kit for both casework and database applications. In 2015, Promega released the PowerPlex® Fusion 6C STR typing kit (Fusion 6C), which contains three additional loci beyond those included in PowerPlex® Fusion 5C, and a sixth color channel. The power of discrimination increases with the inclusion of these additional loci. However, this alone does not justify the replacement of Fusion 5C for Fusion 6C, as multiple costly and time-consuming validation studies would need to be performed before any switch. Given all of the processes to be validated and optimized in the VDFS laboratory prior to implementation, to be beneficial, it would need to be demonstrated that Fusion 6C showed several additional advantages in performance beyond an increase in the already large discriminatory power of Fusion 5C. In this study, both STR amplification kits were assessed for performance metrics in several aspects, including baseline noise, sensitivity, pull-up, allele ambiguity, mixture analysis, and degraded sample analysis. This research found that Fusion 6C demonstrated a lower baseline noise level, less ambiguous pull-up (versus a true minor peak) in three- and four-person mixtures, and statistically significant higher allele counts in four-person mixtures compared to Fusion 5C. Fusion 6C also utilizes a shorter polymerase chain reaction (PCR) cycling procedure, taking approximately 60 minutes to Fusion 5C's 90 minutes. Disadvantages of Fusion 6C include lower sensitivity and more bin overlap. Fusion 5C and 6C are comparable in two-person mixture pull-up and allele counts, three-person mixture allele counts, and likelihood ratios from probabilistic modeling of three- and four-person mixtures. Based on the results, Fusion 6C possesses several critical advantages in addition to the higher possible power of discrimination, and it will be recommended that the Virginia Department of Forensic Science consider implementing this amplification kit for both caseworking and database applications.

Keywords: Forensic science, DNA amplification, PowerPlex® Fusion, short tandem repeats

Introduction

Promega Corporation's (Madison, WI) PowerPlex® Fusion megaplex short tandem repeat (STR) profiling kit was commercially released in 2012. This STR typing kit contains 22 autosomal STR loci, one Y-chromosome STR locus, and amelogenin as a sex indicating locus. The development of this kit, which included an expansion in the number of STR loci from previous kit formulations, was in response to several different factors. Following a recommendation from the Combined DNA Index System (CODIS) Core Loci Working Group in 2011, the Federal Bureau of Investigation (FBI) required that all forensic laboratories participating in the National DNA Index System (NDIS) increase the number of core loci from 13 to 20 by 2017 (1, 2, 3). This new requirement was prompted by the fact that required core loci can differ between databases, as seen with the CODIS of the United States and the European Standard Set (ESS) (2). The expanded number of loci would also provide better resolution of kinship in closely related communities (4). These STR loci fulfill the requirement for upload to both CODIS and other international DNA databases, facilitating global comparisons. In addition, the updated 20 core loci increase discriminatory power, decrease the probability of an adventitious match, and provide greater suppression of polymerase chain reaction (PCR) inhibitors, which is useful for direct amplification (1). Moreover, additional loci and the subsequent increase of power of discrimination are especially conducive for missing person cases and identification of closely related individuals (4). The increase in the CODIS core loci required for upload has led to the development of updated STR typing kits to include all of these loci, as well as additional loci. Initially, two such kits were developed containing the expanded CODIS core loci, the Promega Corporation's PowerPlex® Fusion 5C with 24 loci and Applied Biosystems' (Foster City, CA) GlobalFiler™ PCR Amplification Kit with 24 loci (1, 4).

The PowerPlex® Fusion 5C System (Fusion 5C) consists of the expanded core 20 CODIS loci: CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D1S1656, D2S441, D10S1248, D12S391, D22S1045, D2S1338 and D19S443; as well as Penta E and Penta D (highly discriminatory pentanucleotide STR loci), amelogenin, and DYS391 (4). Amelogenin is a sex indicating marker. DYS391 is a Y-chromosome STR marker that detects male DNA and can provide an independent confirmation of sex when null Y-chromosome allele deletions at the amelogenin locus occur (1, 5). The system uses five dyes to simultaneously amplify and fluorescently detect all 24 of these loci. Previous studies have shown Fusion 5C as a sensitive, specific, and robust STR typing system (6). DNA from as little as 100-500 picograms (pg) have yielded full profiles (4, 6). The reproducibility and precision of this system has been well documented, as well as concordance between other similar STR kits developed by different manufacturers (6). The Fusion 5C kit is an accurate STR typing kit for forensic casework samples, and this has been shown in many validation studies (4, 5, 6). Fusion 5C, a five-dye multiplex, was specifically developed to detect amplified sample fragments using the Applied Biosystems™ 3130/3130xl Genetic Analyzers; however, it can also be analyzed by a newer, higher throughput capillary instrument, the 3500 series (7). The Applied Biosystems' 3500/3500xL Genetic Analyzer, which has the capability to perform fragment analysis with up to six dyes and inject 24 samples simultaneously, can separate and detect Fusion 5C fragments (8). The advantages of the 3500/3500xL instrument, with greater automation and sample throughput, implored many forensic laboratories to purchase this system.

Since the initial release of the Fusion 5C system, Promega developed a six-color STR typing system, compatible with the 3500 series of CE systems. PowerPlex® Fusion 6C (Fusion 6C) was developed with 6-dye chemistry, enabling the inclusion of more loci within a single

reaction, and increasing the power of discrimination (9). The Fusion 6C System allows the amplification and fluorescent detection of 27 loci simultaneously. These expanded loci include the core 20 CODIS loci; three Y-STRs (DYS391, DYS576, and DYS570); and four additional loci (including amelogenin, Penta D, Penta E, and SE33) (9). The inclusion of amelogenin, DYS391, DYS576, and DYS570 allow for sex determination and male component detection, as well as increased discriminatory power in mixtures with male contributors (9, 10). In addition to the Fusion 5C loci, Fusion 6C includes SE33, DYS576, and DYS570. SE33 is a locus that can be highly discriminatory, due to high heterozygosity and variability (1, 11). However, there are documented problems with using SE33 for human identity, including heterozygote peak height imbalance, separation issues, a wide allele range, and a high mutation rate (11). Both DYS576 and DYS570 are rapidly mutating Y-STRs and are included in Fusion 6C to improve mixture analysis and aid in differentiation of close male relatives (1, 10).

A developmental validation performed by Ensenberger et al. demonstrated Fusion 6C has high specificity for human DNA, high sensitivity with minimal DNA input, and concordance across laboratories (12). Sensitivity studies analyzing Fusion 6C have been abundant, with consistent results showing a minimum of 250 pg of input DNA for the detection of 100% of expected alleles above analytical threshold (10, 12, 13). Ensenberger et al. notes that in the presence of inhibitors, Fusion 6C performs at a similar level to Fusion 5C (12). Just like Fusion 5C, Fusion 6C provides a robust, accurate, and sensitive STR typing kit for casework and DNA databasing (10, 13). Advantages of this kit include uniform heterozygote peak heights, balanced intra-color signal, and improved preservation of mixture ratios when compared with similar kits by other manufacturers (1). Fusion 6C also has a shorter cycling time for amplification, speeding up the workflow. While both kits employ rapid PCR cycling procedures, Fusion 5C takes

approximately 90 minutes, while Fusion 6C takes around 60 minutes (1, 7, 9). Fusion 5C utilizes a traditional three-step PCR process with an additional temperature control parameter, whereas Fusion 6C utilizes a two-step PCR cycle without an extension step (7, 9).

Forensic laboratories must perform an extensive validation of new technologies prior to implementation. The time, cost, and effort is multiplied if a laboratory utilizes a probabilistic modeling system for mixture interpretation, which must be validated for each DNA profiling kit used. In addition, many forensic laboratories house a databanking facility which will typically employ the identical STR profiling kit as the caseworking entity in the laboratory. Not only is this costly in relation to the purchase of new reagents and occasional new instrumentation, but it is also costly in the requirement of manpower to perform validation and training of personnel. A small increase in discriminatory power may not be sufficient to motivate a laboratory to invest the resources necessary for changing from one DNA profiling kit to another. Thus, the goal of this comparative study is to determine if the use of the six-dye technology in Fusion 6C, as well as different dye chemistries and cycling parameters, provide for such an enhancement in performance when compared to Fusion 5C to justify validation and implementation of this kit and discontinuation of Fusion 5C for casework and convicted offender database applications.

As technology improves and forensic laboratories detect and amplify lower concentrations of DNA, stochastic effects are more likely to present in the form of artifacts and baseline noise (14). Tay et al. determined that the Fusion 5C typing kit presented a higher baseline noise level as compared with the Applied Biosystems VeriFiler™ Plus kit and the QIAGEN Investigator® 24plex QS kits (Hilden, Germany) (14). This increased baseline noise was most prevalent in the yellow dye channel (TMT-ET) (14). The presence of more background noise in the Fusion 5C system can lead to difficulty in determining whether a peak is noise, some

other artifact, or a true allele. Not only can this impact human interpretation of the STR typing results, which requires the measurement and application of an analytical threshold, it may also impact probabilistic modeling systems that do not apply an analytical threshold, such as TrueAllele® Casework (Cybergenetics, Pittsburg, PA), which measures locus specific baseline noise de novo for each electropherogram (15). Excessive baseline noise may obscure true peaks, therefore eliminating potentially informative alleles. To filter out baseline noise, an analytical threshold is often implemented in forensic laboratories, based upon limit of detection values for each dye channel. The limit of detection is the minimum threshold of signal that can be distinguished from noise. Forensic laboratories err on the side of conservativism to minimize the effects of baseline noise by setting minimum peak height thresholds. This threshold must be set to a value where low-level true signal is not obscured, however, since this is a setting based upon validation samples and not performed *de novo* for each electropherogram, often true signal is filtered out with noise. Forensic laboratories differ in the approaches for setting an analytical threshold. One common method for measuring the analytical threshold value is to average the baseline noise peak heights that are unrelated to any PCR products for each color channel and add three standards deviations (16). Another method is to determine the maximum peak height from baseline noise at each dye channel and multiply by two. Both procedures have shown concordance at the Virginia Department of Forensic Science (VDFS), the laboratory system participating in this study, in determining the threshold value. Additionally, some laboratories may choose one analytical threshold value for all color channels based on the data, and others set different limits for each dye. Incorporating different amplification kits will result in different limit of detection values. A direct comparison of baseline noise between Fusion 5C and Fusion 6C must be performed in order to determine if the kits will have different analytical thresholds.

The lower the limit of detection, the smaller the analytical threshold can be, which refers to the lowest value (RFU) that DNA can be reliably distinguished from noise, strengthening the assumption that low-level true signal is not being obscured or filtered out.

In addition to the higher baseline noise seen in Fusion 5C, Tay et al. also found that the Fusion 5C System produced an average of 11.5 artifacts when the recommended amount of 500 pg input DNA was used (14). While data directly comparing the prevalence of artifacts between STR typing kits is limited, it is important to note differences may exist between kits and impact STR typing interpretation. A common artifact seen after capillary electrophoresis is pull-up. Pull-up is caused by overlap of the emission spectra between different dye colors (17). The fluorescent dyes used are designed to emit light at different wavelengths; however, there is overlap when the emission spectra are close in wavelength. For example, the blue dye channel usually emits light between 450 and 495 nm, and the green dye channel emits approximately between 495-570 nm. This spectral overlap can cause peaks from one dye to bleed through into the channel of another. This is especially problematic when allele fragments causing the pull-up effect have a high fluorescent signal and align with bins for loci in the other dye channels, sometimes leading to uncertainty as to whether what is observed is pull-up, stutter, or an actual allele.

Fusion 5C displays overlap between bins of differing color channels. For example, physical bins located in the blue channel can overlap with bins present within loci in the green, yellow, and red channels. Bin overlap can result in ambiguous alleles when pull-up occurs during capillary electrophoresis, and this results in increased time spent in data analysis. Since Fusion 5C tends to produce artifacts, including pull-up, and has overlapping bins, it is worthwhile to compare and assess the extent of bin overlap in Fusion 6C. If Fusion 6C demonstrates less

overlap of bins, then it may possess a reduced likelihood of pull-up being falsely labeled by the software or manually as an allele or stutter.

Other challenges faced during data interpretation include mixtures, degraded samples, and low-template samples. The ambiguity of mixture profiles can make data analysis difficult. In their respective developmental validations, Fusion 5C produced 88% minor contributor alleles at a 1:9 ratio, while Fusion 6C produced 96% unique minor contributor alleles at a 1:9 ratio (4, 12). While this difference may not seem large, it has a potential impact on the success rate of developing a meaningful profile from a low-template contributor. A meaningful profile is one where sufficient information is developed to draw a conclusion. Part of the difference can be explained by the three Y-STRs present in Fusion 6C, which aid in mixture deconvolution, as well as SE33, which can be highly discriminatory. Assessing the differences in mixture interpretation between the two kits, with and without including the additional loci in Fusion 6C, will allow a direct comparison to determine which kit can be more successfully utilized to deconvolute mixtures.

The VDFS uses probabilistic modeling to assess statistical power in complex mixtures (18, 19). Probabilistic modeling (also known as probabilistic genotyping, PG) using the TrueAllele® Casework System models baseline noise *de novo* for each mixture to render a statistical electropherogram (15). It uses Markov chain Monte-Carlo (MCMC) searching to sample the statistical space for which an explanation (genotypes of the contributors to the mixture and their relative contributor proportions) is rendered. (15). It can then predict likelihood ratios for many combinations of parameters that explain the DNA, and ultimately, a likelihood ratio is employed to determine an individual's association to the deconvoluted mixture profile (15). Sampling baseline noise *de novo* without an analytical threshold allows interpretation for

increasingly complex and low-template mixtures and can provide information on the comparative sensitivity and accuracy of Fusion 5C and Fusion 6C, as well as any effects stemming from baseline noise.

In addition, mixtures can have exacerbated effects due to pull-up. Interpreting mixtures, including determining the number of contributors and assessing artifacts, is already complicated. The more contributors present, the more challenging it is to distinguish between true alleles, stutter, and pull-up. Comparing Fusion 5C and Fusion 6C for pull-up in increasingly complex mixtures can be extrapolated to assess which kit demonstrates less ambiguous artifacts that can be resolved upon human inspections, and therefore is easier to analyze.

Degradation occurs when the DNA molecule is no longer intact, for reasons such as environmental conditions, UV exposure, heat, and time (20). Amplicons smaller in length are more likely to be intact and successfully typed since during the degradation process, the DNA molecules become increasingly more fragmented. Fusion 5C has nine loci less than 220 base pairs, and 6C has ten (7, 9). The additional locus less than 220 base pairs, since this amplicon is smaller in length, may provide Fusion 6C a small advantage for successful typing of severely degraded samples. Quintin et al. performed a recent study analyzing the Fusion 5C kit's performance with degraded DNA samples and found that Fusion 5C consistently gave accurate typing results of degraded DNA (20).

As with degraded DNA samples, low-template samples are also challenging for analysts to interpret. There is variation in defining what constitutes low-template, but it is typically considered to be at approximately 100 pg or less of template DNA (21, 22). A variety of factors, including more sensitive instrumentation and the advancement of statistical interpretation software, have expanded the capabilities of forensic laboratories to interpret lower template

samples (14). Newer STR kit development has heavily focused on increasing sensitivity of amplification, and full profiles have been demonstrated from 100 pg of input DNA, making the definition of low-template DNA ambiguous (14, 22). Similar to degraded DNA samples, low-template samples display a loss of allelic and locus information; however, this pattern is not as predictable as seen in degraded samples. Butler writes that specific loci may be more likely to exhibit allelic dropout, especially when larger in size, but that each STR typing kit will vary in the loci more prone to dropout due to the design of the primers and other differences in conditions (22). Thus, each STR typing kit must be assessed for which loci tend to lose allelic information first with restricted template quantities. Stochastic effects can greatly impact the proclivity of allelic information first lost, leading to a high degree of variability and a loss of predictability. (22). Comparing the performance of Fusion 5C and Fusion 6C with low-template DNA input will determine the use of which kit provides more success in obtaining complete profiles.

Fusion 6C contains two rapidly mutating Y-STR loci, DYS570 and DYS576. The mutation rate for these two loci is 1 in 101.7 and 1 in 74.375, respectively (23). This is far more often than the average mutation rate of around 1 in 1000 for more commonly utilized loci in forensic DNA typing (24). Based on this, it is possible that the inclusion of the rapidly mutating Y-STRs in Fusion 6C will improve mixture deconvolution and discriminate between male relatives more efficiently than Fusion 5C. Assessing male relatives using the Y-STRs contained in Fusion 6C will provide an indication if the rapidly mutating Y-STRs result in differences between patrilineally related people.

Advantages of the Fusion 6C System are well studied, and include strong heterozygous peak height balance, intracolor signal balance, preservation of mixture ratios, high discriminatory

power, and robustness (1, 2, 13). None of the disadvantages documented with the Fusion 5C STR kit, such as high baseline noise and artifacts, have been reported in the literature regarding Fusion 6C. There is little data directly comparing Fusion 5C and Fusion 6C, and this study aims to provide clarity regarding the advantages and disadvantages to Fusion 6C when compared with Fusion 5C. Based on the information obtained from this study, a recommendation will be provided to the Virginia Department of Forensic Science on whether a full validation of Fusion 6C and its replacement of Fusion 5C for casework and databasing operations is merited.

2. Materials and Methods

2.1. DNA Collection, Extraction, and Quantification

All samples used in this project were collected by the research section at the Virginia Department of Forensic Science to mimic casework and databank samples. Sample types included reference buccal swabs, FTA blood cards, mock forensic-type samples from proficiency tests, mixtures, and various tissue from cadavers and blood samples exposed to various conditions to produce degraded DNA. All samples used for testing, except for mixture preparation, were extracted using the Promega DNA IQ™ system (Promega, Madison, WI) following the VDFS optimized protocol for manual extraction (25). For two-person mixtures, the Promega DNA IQ™ system was used on the Biomek® NX^{P} Automation Workstation (Beckman Coulter, Brea, CA; n=48) following VDFS protocols (25). For three- and four-person mixtures, six buccal swabs were extracted using an organic extraction method in order to obtain the highest possible DNA yield. The organic extraction procedure used followed protocols from the VDFS (25). Extracted DNA was stored in a -20˚C freezer.

Extracted DNA was quantified via qPCR using Promega's PowerQuant® System with Applied Biosystems' (Foster City, CA) QuantStudio™ 5 instrument following manufacturer recommended methods (26). Standards were included in each run with known concentrations of 50 ng/µL, 2 ng/µL, 0.08 ng/µL, 0.0032 ng/µL, and no template, where two of each were included to give a total of 10 DNA standards per qPCR reaction. In addition to concentration of autosomal DNA, Y-chromosome DNA concentration, autosomal/Y-chromosome DNA ratio, and degradation index were also generated.

2.2. PCR Amplification

DNA samples were amplified using both the PowerPlex® Fusion 5C and PowerPlex® Fusion 6C amplification kits. Half-volume reactions (12.5 µL) utilizing 0.5 ng of DNA (in 5 µL) and 28 cycles were performed for single-source and two-person mixtures with Fusion 5C and 6C, the former of which was previously validated by VDFS and the latter of which was optimized in this study (27). Table 1 contains PCR parameters used for both Fusion 5C and Fusion 6C in this study. Both the GeneAmp™ PCR System 9700 (Thermo-Fisher Scientific, Waltham, MA) and Applied Biosystems' Veriti were used for STR amplification, with caution taken to ensure the same instrument was used for all comparative studies.

Half-volume reactions were also processed as described above for three- and four-person mixtures, as well as degraded samples, for both of the Fusion kits; 0.5 ng of DNA was amplified using Fusion 5C and 0.625 ng was amplified using Fusion 6C. The difference in input for these samples followed a DNA target optimization of Fusion 6C performed in this study for half-volume reactions, where 0.625 ng of total DNA allowed more consistent peak heights between full-volume Fusion 6C reactions following manufacturer recommendations, and half-volume Fusion 6C reactions.

Full-volume reactions (25 µL) utilizing 1 ng of DNA (in 15 µL) and 29 cycles were also performed for single-source samples and two-person mixtures with Fusion 6C, as recommended by the manufacturer (9).

2.2.1 Half Reaction Optimization of Fusion 6C

Fusion 6C was optimized for cycle number using half-volume reaction conditions. Master mix, amplification grade water, and the primer mixture were added in amounts of 2.5 µL to the reaction cocktail. DNA was added at a concentration of 1 ng/ 7.5 µL. PCR cycle settings of 27 and 28 cycles were compared with the full-volume setting of 29 cycles. The average peak heights across color channels (RFU) were compared to determine which conditions most mimicked full-volume reaction results. In addition, allele counts were used to compare the numbers of observed alleles between the three conditions. Both average peak heights and observed allele counts were used to determine which cycle number led to half-volume reaction profiles that were most similar to the full-volume reaction (1 ng/ 15 µL).

2.2.2. Target DNA Input

To optimize the input quantity of DNA for Fusion 6C, the target amount of DNA was found by determining which of three half reaction inputs, 1 ng, 0.75 ng, and 0.5 ng, performed most similarly to a 1 ng full-volume reaction (n=3). This was measured by comparing the average peak heights (RFU) per each color channel across samples. The peak heights were normalized for homozygosity, by dividing a homozygous peak height by two to account for the inflated value. As with the cycle optimization, the goal of the comparison was to find which half-volume target DNA input was most similar to the full reaction (1 ng) without overblowing the signal and giving a higher average peak height than the full reaction.

2.3. STR Separation and Profile Analysis

Amplified DNA was detected and analyzed using the Applied Biosystems 3500xL Genetic Analyzer, with a 36-cm capillary array, POP-4 polymer, and a 1.2 kV injection voltage, for 12 and 24 second (s) injection times in both Fusion 5C and Fusion 6C amplified samples. Capillary electrophoresis parameters followed VDFS outlined conditions for Fusion 5C, with an input of 1 µL of amplified DNA or ladder per well, plus 9.5 µL of formamide and 0.5 µL of WEN-Internal Lane Standard (ILS) (28). Wells without DNA included 1 µL of formamide, for a total of 11 µL per well. For Fusion 6C, manufacturer recommended protocols were used, which included the parameters mentioned above (29). Samples were analyzed with GeneMapper® ID-X version 1.4 (Applied Biosystems) following VDFS procedures for analysis of capillary electrophoresis results for Fusion 5C, and utilizing manufacturer recommendations for Fusion 6C (18, 29). Data analysis was carried out using VDFS analysis protocols for Fusion 5C, using an analytical threshold of 75 relative fluorescence units (RFU) for each color channel with the exception of the orange dye channel, which had an analytical threshold of 50 RFU. The analytical thresholds for Fusion 6C for both full-volume and half-volume reactions were determined in this study for each color channel, ranging from 40-60 RFU and 65-102 RFU, respectively (see Table 2, Table 3).

2.3.1. Limit of Detection

The limit of detection (LOD) was determined for full-volume and half-volume reactions in Fusion 6C. LOD was measured by determining the maximum baseline noise in RFU per sample and per color channel, specifically noise that was not affected by alleles, pull-up, stutter, or other PCR artifacts. The limit of detection was found by identifying the maximum noise values across all samples for each of the five color channels (blue, green, yellow, red, and

purple), and then multiplying that maximum value by two for each color channel. A total of 56 measurements were used for full-volume determination, and 17 were used for half-volume reaction LOD determination.

2.4 Comparative Analyses

Many additional studies were performed directly comparing Fusion 5C to Fusion 6C in performance and critical aspects. These sub-studies are broken down in further detail here.

2.4.1. Number of Alleles Detected

Using the paired T-test ($\alpha = 0.05$) function in Microsoft Excel, Fusion 5C and Fusion 6C were analyzed for significant differences in autosomal allele counts when using the same sample amplified under each kit's optimized half-volume reaction conditions; using a hypothesized mean difference of zero when comparing loci present in both Fusion 5C and Fusion 6C. When SE33 was included in the calculations, the hypothesized mean difference was adjusted to two, to account for the two possible alleles present in Fusion 6C that were not present in Fusion 5C. Both two-tailed T-tests and one-tailed T-tests were performed to examine if there was a significant difference between kits, and to assess if Fusion 6C provided a higher average allele count than Fusion 5C.

2.4.2. Sensitivity

To assess sensitivity, two-fold serial dilutions were performed on the same DNA sample from a reference buccal swab. Final DNA quantities were 600 pg, 300 pg, 150 pg, 75 pg, 37.5 pg, 18.75 pg, 9.375 pg, 4.69 pg, and 2.34 pg. Samples were amplified with both Fusion 5C and with Fusion 6C using half-volume reactions (0.5 ng in 5 µL). Sensitivity was assessed by

determining which amplification kit yielded a higher percentage of alleles detected for each quantity, which was calculated by the following formula:

$$\frac{Number\ of\ Alleles\ Detected}{Number\ of\ Expected\ Alleles} \times 100\%$$

2.4.3. Bin Overlap

Bin overlap between color channels was determined for both Fusion 5C and Fusion 6C. In GeneMapper® ID-X, physical bins represented by the gray bars were counted if they exhibited overlap with a bin in any another color channel by an amount greater than 0.1 base pairs. Overlapping bins were counted for each color channel, except for the orange, since it contains the internal lane standard (ILS).

2.4.4. Mixture Studies

Two-person mixture samples from proficiency test samples (n=6) and previously created male: female mixtures (n=48) were used to analyze both pull-up and detected allele counts for Fusion 5C and Fusion 6C. The two-person mixtures were processed singularly using full-volume reactions for Fusion 6C (25 µL), and the validated half-volume reaction for Fusion 5C (12.5 µL). The proficiency test samples were processed following the procedure outlined earlier. The male: female mixtures had been previously extracted by another analyst, using the Biomek® NX^P Automation Workstation, and the workflow was resumed from the quantitation step onwards.

Three-person mixtures were created using ratios of the single source DNA samples at 60:30:10, 70:15:15, and 80:10:10 (n=9) for a total input DNA amount of 0.5 ng for Fusion 5C and 0.625 ng for Fusion 6C.

Four-person mixtures were created using ratios of the single source DNA samples at 70:10:10:10, 50:20:15:15, and 60:15:15:10 (n=12) for a total input DNA amount of 0.5 ng for Fusion 5C and 0.625 ng for Fusion 6C.

2.4.4.1. Pull-up Analysis

For each set of mixtures (two-, three- and four-person) total pull-up was measured by counting pull-up events noted during human data review that were located in allele calling bins. This was done using GeneMapper® ID-X for samples amplified with both Fusion 5C and Fusion 6C. In addition, any pull-up that led to a false allele call was identified after analysis by comparing mixture profiles to each contributor's reference profile. Alleles falsely identified were examined to determine if pull-up played a factor in the incorrect assignment as a real peak. This was done by reexamining the profile to determine if a parent allele was in the pull-up position for the false allele, classified by a range of plus or minus 1.0 base pairs, in another color channel.

In three- and four-person mixture samples, amplification using 0.5 ng in 5 µL for Fusion 5C and 0.625 ng in 5 µL resulted in an incomparable difference in peak heights; the Fusion 5C peak heights were higher on average than the Fusion 6C peak heights. To adjust for this, and continue using the same samples, the remainder of the prepared mixtures (7.5 µL) was included in a re-amplification using Fusion 6C for a total of 0.9375 ng of total DNA input.

2.4.4.2. Autosomal Allele Counts of Mixtures

Two-, three-, and four-person mixtures were used to assess any differences in allele counts between samples amplified with Fusion 5C and Fusion 6C. Autosomal allele counts were determined by counting the number of detected alleles, excluding the SE33 locus to allow for direct comparison between kits. Significant difference in average allele count was calculated

using a paired T-test (α =0.05) with a hypothesized mean difference of zero, as detailed in section 2.4.1.

2.4.4.3. Probabilistic Modeling

Two three-person and two four-person mixtures that indicated the presence of dropout were selected for analysis using TrueAllele® Casework. Both Fusion 5C and Fusion 6C amplified products of the chosen mixture samples were included for the analysis. Procedures followed those outlined in the VDFS manual (19). Likelihood ratios were calculated in the software. The SE33 locus was not included for the statistical assessment since the goal was to determine if baseline noise and other differences between the loci in common would result in greater contributor specific allelic information for one STR kit versus the other using the probabilistic modeling system.

2.4.5. Degraded Samples

FTA blood card samples (n=11) previously treated under a variety of conditions were used to perform a degradation study. These conditions are as following: one month at room temperature under sunlight (n=1), one month at 37˚C (n=1), one month at 56˚C (n=1), one month at 80˚C (n=2), three months at room temperature in a damp environment (n=2) three months at 56˚C (n=1), and three months at 80˚C (n=3). In addition to these conditions, samples were kept at room temperature for almost 20 years post treatment. The sample set also consisted of five tissue samples from a brain, heart, rib, muscle, and liver from the same individual, stored for an unknown length of time at -20˚C. The tissue samples were known to be degraded prior to the storage over time. The tissues had been previously typed and were identified as having differing degrees of DNA degradation depending on the tissue source. The degradation index was

determined using the PowerQuant® System, and autosomal allele counts were calculated to assess the tolerance of both Fusion 5C and Fusion 6C to degraded samples. Significance tests were conducted, using a paired one-tailed T-test to determine if Fusion 6C was providing additional allelic information over Fusion 5C in degraded samples (α=0.05).

2.5. Rapidly Mutating Y-STRs

DNA was extracted from a father and his two sons and amplified using only Fusion 6C. Alleles from DYS570 and DYS576 were compared between the three samples, specifically to assess if any random mutations arose between the two generations.

3. Results and Discussion

3.1. Half Reaction Optimization of Fusion 6C

Currently, VDFS uses half-volume reactions for Fusion 5C. To allow for direct comparison to Fusion 5C, half-volume conditions first had to be optimized for Fusion 6C. This would limit any variations associated with reducing the volume of the amplification reaction. Autosomal allele counts for samples amplified with a full-volume reaction, following the manufacturer recommended cycling protocols, were compared with those from half-volume reaction samples. Conditions tested included amplification using 27-and 28-cycle processes. The 29-cycle setting was not tested using the half-reaction since it had been reported that reduced volume reactions increase sensitivity, and thus, are not allowed for sample upload to NDIS (22, 30). When the samples were compared, those amplified with 28 cycles were most consistent with the autosomal allele counts obtained from the full-volume reaction (Figure 1). Peak heights were also assessed to determine the optimal cycle number most similar to the full-volume reaction and manufacturer conditions. The peak heights were most similar between 29-cycle full-volume and

28-cycle half-volume (Figure 2). Based on this, the half-volume reaction protocol for Fusion 6C was set to include a 28-cycle process (Table 1).

3.2. Target DNA Input

The manufacturer recommends 1 ng/ 15 µL for DNA input in a Fusion 6C reaction (9). Since a half-volume reaction was necessary for this study, a target optimization was performed to determine which amount of DNA input was most effective at a 5 µL maximum volume. Average peak heights across dye channels were compared between half-volume reactions with DNA inputs of 1 ng, 0.75 ng, and 0.5 ng, with the goal of performing most similarly to the full-volume reaction with 1 ng of DNA. At a 12s injection time, the input DNA that resulted in peak heights most similar to a full-volume reaction was between 0.5 ng and 0.75 ng (Figure 3). Standard deviations were calculated to show the variation observed. Homozygous alleles were normalized to account for some of the variation seen in the average peak heights, likely due to amplification efficiency at the homozygous loci analyzed. At a 24s injection time, the same trend was observed, where the optimal DNA target amount was between 0.5 ng and 0.75 ng (Figure 4). Therefore, the optimal input of DNA for a Fusion 6C half-volume reaction that most closely mimics a full-volume reaction was determined to be 0.625 ng for this study.

Using the values obtained for PCR cycling and DNA input, half-volume reactions amplified with Fusion 6C were able to be directly compared to Fusion 5C amplified samples processed with the VDFS standard workflow.

3.3. Limit of Detection

Establishing an analytical threshold limits variability in calling a true peak from a false one when analyzing capillary electrophoresis data. It also minimizes the impact of any potential

baseline noise. The limit of detection determination assists in setting the value of the analytical threshold. The goal is to ensure no true peaks are being left out, as well as no noise being confused as a real allele peak. While the accuracy of the distinction between real low-level alleles and noise cannot be 100%, a high level of confidence can be provided to the process by requiring >99% confidence interval. Fusion 5C, having already been validated at VDFS, has a limit of detection at 75 RFU for each of the four-color channels, excluding orange. This one setting was adopted for ease of application since both the blue and green channels showed significantly lower baseline noise. The calculated values were 37 (blue) and 58 (green) for 12s and 24s data combined; as reported in the VDFS Validation of the PowerPlex® Fusion System for Use with the ABI Genetic Analyzer: Limit of Detection summary.

Using full-volume and half-volume reactions, the limit of detection for Fusion 6C was established. This was for the purpose of analyzing data from samples amplified with Fusion 6C, as well as determining the effects of baseline noise on both kits. Since the limit of detection for Fusion 5C had already been established in a previous validation study, those values were used for comparison with Fusion 6C.

Since Fusion 6C has an additional color channel, all five of the dyes were analyzed for a limit of detection. The orange dye channel (ILS) is the same as used in Fusion 5C, and this was not reanalyzed as no analytical threshold is required for the ILS. For full-volume reactions, the limit of detection was measured to be 62, 58, 62, 62, and 60 for blue, green, yellow, red, and purple, respectively, at a 12s injection (Table 2). At a 24s injection, the limits of detection were 54, 62, 102, 64, and 40 in the same color order (Table 2). The yellow channel showed an outlier in one of the samples at the 24 second injection time. Since there was no other explanation for this value other than high baseline noise, this was left in the evaluation. Therefore, the yellow

channel was assigned an analytical threshold of 102 RFU (Table 3). Based on the remaining values, analytical thresholds were assigned for full-volume 6C reactions at 65 RFU for the blue, green, red, and purple channel (Table 3). These values were rounded up from the limit of detection values, to simplify the settings and provide a conservative estimate. It is unclear if there is less baseline noise for the Fusion 6C system than the Fusion 5C using full-volume conditions. However, in order to properly compare the two, the limit of detection for each color channel in Fusion 6C had to be determined using the optimized half-volume reaction, to allow for direct comparison and conclusions. At a 12s injection time, the limit of detection values for Fusion 6C half-volume reactions were measured to be 36, 52, 30, 48, and 38, for blue, green, yellow, red, and purple, respectively (Table 2). At a 24s injection time, the values were recorded at 36, 58, 28, 52, and 32 (Table 2). The analytical threshold values for Fusion 6C half-volume reactions, based on the limit of detection data, were set at 40 RFU for the blue, yellow, and purple channels, and 60 RFU for the green and red channels (Table 3). These values allow for direct comparison to Fusion 5C, and all five color channels have an analytical threshold slightly less than the one value adopted for Fusion 5C (75 RFU). There may be slightly less baseline noise present in samples amplified with Fusion 6C. This may be a slight advantage to this amplification kit, as less noise is beneficial both for data analysis and the accurate calling of allele peaks, as well as for the efficiency of probabilistic modeling systems like True Allele® Casework, that measure baseline noise *de novo*.

3.4. Sensitivity

Previous studies have determined that both Fusion 5C and Fusion 6C are sensitive, and as little as 250 pg can yield 100% of expected alleles using both amplification kits (6, 10, 12 ,13). However, the two kits have not been directly compared for sensitivity, particularly using half-

volume reactions. The same sample was diluted from a range of 600 pg to 2.34 pg. The same diluted sample was amplified with both Fusion 5C and Fusion 6C and the results were compared to determine which was more sensitive. The percentage of alleles detected was assessed for each, by dividing the number of observed alleles by the number of expected alleles and multiplying by 100%. Both Fusion 5C and Fusion 6C showed full profiles down to 75 pg (Figure 5). Fusion 6C samples lost allelic information more rapidly than Fusion 5C as DNA quantity continued to decrease. As seen in Figure 5, at least some allelic information was retained for both Fusion 5C and Fusion 6C for the half-volume reactions until the minimum quantity tested, 2.34 pg, but Fusion 5C displayed a higher allele count than Fusion 6C from 37.5 pg and below. It is also important to note that Fusion 6C contains one additional autosomal allele, SE33, and this was included for sensitivity purposes. Percent profile was used for the comparison to account for this. As such, Fusion 5C appears to be more sensitive than Fusion 6C, especially in a half-volume capacity. A less sensitive amplification kit can lower the effects of instrument noise and contaminants. However, the major drawback is potential loss of allelic information, especially in samples compromised in quality and/or template, as often seen in forensic casework.

3.5. Bin Overlap

A previous issue identified with Fusion 5C is the pull-up of signal into bins of other color channels. This can result in a false allele call with traditional genotyping software if the pull-up effect is large enough to cross the analytical threshold. While the pull-up requires human examination, it can complicate the data analysis and lead to either falsely calling the pull-up a true allele, or more commonly, calling the true allele pull-up since it can be so difficult to distinguish between the two alternatives when analyzing complex mixtures containing low-level, along with high-level, contributors. With the addition of an extra dye channel in Fusion 6C, it

was possible that the bin overlap could be decreased since the loci could be more spaced out across five dye channels. Less bin overlap would reduce the potential number of false allele calls from pull-up into other color channels. This effect was studied by counting overlapping bins in both amplification kits. Fusion 5C demonstrated significantly less overlap of bins than Fusion 6C (Table 4). This is likely due to the addition of three loci to Fusion 6C that are not present in Fusion 5C. Even though the additional color channel provides more space for the STR loci, the additions crowd the multiplex even more. This, plus the overlap from the purple channel, is likely why Fusion 6C has significantly more bin overlap. Pull-up was further analyzed to assess if factors other than bin overlap resulted in differing amounts of pull-up between STR amplification kits.

3.6. Pull-up Analysis

Pull-up was compared between Fusion 5C and 6C in two-, three-, and four-person mixtures. Results were analyzed first by GeneMapper® ID-X, and then by human analysis for artifact determination. For two-person mixtures, more pull-up into bins was noted in Fusion 5C (Table 5). Both 5C and 6C had two false allele calls due to pull-up after both analyses and comparison to reference profiles (Table 5). Table 5 demonstrates total pull-up observed in two-person mixture samples, as well as false allele calls due to pull-up, with the allele to the left of the pull-up designating the parent allele and the tan fill indicating stutter was a contributing factor. Since these two allele calls were missed after human review, it can be assumed that these were ambiguous. It is worth stating that Fusion 5C samples showed more pull-up than Fusion 6C for two-person mixture samples, as any artifacts interfere with downstream analysis. However, the focus of this study is to assess ambiguous pull-up between Fusion 5C and Fusion 6C, to

determine if either amplification kit is generating less. In these two-person mixture samples, the same number of false alleles calls due to pull-up were observed between both amplification kits.

For three- and four-person mixtures, Fusion 5C had 25 pull-up events into allele bins, compared to zero in Fusion 6C. However, there was a peak height disparity despite using the optimal quantity of DNA for both kits. Samples were amplified with a new target of 0.9375 ng for Fusion 6C, with the goal of raising the peak heights to determine if the lopsided pull-up was a factor of the smaller peak heights observed in 6C amplified samples. The DNA template increase resulted in the observation of more comparable peak heights, and the study was performed again. Fusion 6C samples resulted in twice as much pull-up than the same samples amplified with Fusion 5C at 0.5 ng input DNA (Figure 6). However, much of the pull-up seen in these Fusion 6C amplified samples was easily detectable by human review, and after analysis, there were only two allele calls found to be false due to pull-up (Figure 6). This is one-third the amount seen in Fusion 5C, which showed six mistaken alleles. The incorrectly assigned pull-up peaks were missed due to their size, which was consistent with the sizes of other alleles from lower-level contributors to the mixture samples. This finding suggests that in Fusion 6C the allele calls are less ambiguous than samples amplified with Fusion 5C. More certainty in distinguishing artifacts from a true allele is beneficial to a forensic laboratory. This can reduce analyst time spent in data analysis, and ultimately reduce turnaround time for evidence. Especially in three- and four-person mixtures, where the potential for ambiguity during interpretation increases, Fusion 6C can be a real advantage.

3.7. Autosomal Allele Counts of Mixtures

The number of autosomal alleles present in two-person mixture samples amplified with both Fusion 5C and 6C were determined and compared. Since SE33 is only in Fusion 6C, it was

not included to allow for direct comparison, but it is important to note that this locus increases the power of discrimination capabilities of Fusion 6C. For two-person mixtures, there was no significant difference (α=0.05) found between kits at 12s and 24s, with p-values of 0.587 and 0.505, respectively (Table 6). A one-tailed T-test, ($H_A = \mu_{1(5C)} < \mu_{2(6C)}$), demonstrated that Fusion 6C autosomal allele counts were not significantly higher than Fusion 5C, with p-values of 0.706 and 0.253, respectively (Table 6). Two-person mixture samples injected at 24 seconds had a higher average allele count in Fusion 6C than Fusion 5C. While not found to be statistically significant, even a single allele may increase the power of discrimination, which is advantageous.

Three- and four-person mixtures were also amplified with both Fusion 5C and Fusion 6C to determine any difference in autosomal allele count. For three-person mixtures, there was no statistical support that allele counts were higher in Fusion 6C with a p-value of 0.935 for a one-tailed T-test. There was no significant difference between Fusion 5C and Fusion 6C allele counts with a p-value of 0.129 for a two-tailed test (Table 7). Fusion 5C showed a greater number of samples with a higher percent profile than Fusion 6C (Table 8, Figure 7). Again, SE33 was not included.

For four-person mixtures, a one-tailed T-test provided statistical support that Fusion 6C had a greater average allele count than Fusion 5C, with a p-value of 0.026 (Table 7). This indicates that even without SE33, more allelic information was obtained in four-person mixtures when amplifying with Fusion 6C. The four-person mixtures amplified with Fusion 6C also showed more complete percent profiles than those amplified with Fusion 5C (Table 9, Figure 8). Based on this study, Fusion 6C provides more allelic information than Fusion 5C as the complexity of the mixture increases. Again, this study did not include SE33, which unquestionably provides additional allelic information and discrimination. This study also did

not include the two rapidly mutating Y-STRs also included in Fusion 6C, DYS570 and DYS576. These Y-STRs can provide additional help in mixture deconvolution, such as determining the number of male contributors (2).

3.8. Probabilistic Modeling

For more complicated mixtures, the VDFS utilizes probabilistic genotyping to determine likelihood ratios of possible contributors. Two samples from each of the three- and four-person mixtures were chosen, specifically ones that exhibited significant drop-out in the profile with peaks below the analytical threshold, for further analysis using the True Allele® Casework System (TA). The samples were amplified with both Fusion 5C and Fusion 6C to allow for a direct comparison of the likelihood ratios produced by TA. The resulting likelihood ratios for the three-person mixtures did not differ by more than two ban (logarithmic units), which is the threshold set by Cybergenetics to verify that the mixture deconvolution process was reproducible (28). This indicates that no significant differences were found between the discriminatory power of three-person mixtures amplified with Fusion 5C versus Fusion 6C. For the four-person mixture samples analyzed by TA, the same conclusion was reached. No significant differences in match scores were found. For these complex mixtures, it appears Fusion 5C and Fusion 6C are comparable without the inclusion of SE33. It is important to note that SE33 was not included for the analysis with True Allele. This would undoubtedly provide a difference in the observed likelihood ratio.

3.9. Degraded Samples

Fusion 6C contains one additional locus than Fusion 5C under 220 base pairs in product length. Amplicons with lengths greater than 250 bp are less likely to successfully type in

degraded samples (31). When DNA degradation is a factor, the smaller loci tend to amplify more consistently. The additional locus under 220 base pairs in Fusion 6C is likely to provide extra allelic information in samples that have been degraded. To study this, samples of varying degrees of degradation were amplified using both Fusion 5C and Fusion 6C. The degree of degradation was measured using the PowerQuant® qPCR system, which provides a degradation index derived from dividing the autosomal target concentration by the degradation target concentration. A value of 2.0 or greater indicates a progressively greater level of deterioration. Autosomal alleles were counted and compared between kits. As expected, as the degradation index increased in these samples, the autosomal allele count decreased in both Fusion 5C and Fusion 6C. SE33 was excluded at first to allow a direct comparison between Fusion 5C and Fusion 6C. Fusion 6C showed four samples with higher allele counts than Fusion 5C (Figure 9). This is out of 16 total samples, displaying that Fusion 6C is providing more information in about 25% of samples than Fusion 5C (Figure 9). Fusion 5C showed three samples with higher allele counts that Fusion 6C (Figure 9). The samples show the same autosomal allele counts in 9 of the 16 samples (Figure 9). However, no statistically significant support was found that Fusion 6C provided more allelic information with a p-value of 0.319, using an α value of 0.05 (Table 10). SE33 was then included in the comparison to determine if there was any advantage in typing degraded samples in Fusion 6C including all of its autosomal loci. When SE33 was included, Fusion 6C demonstrated higher allele counts than Fusion 5C in 10 of the 16 samples, or 62.5% (Figure 10). Fusion 5C demonstrated three samples with higher allele counts than Fusion 6C (Figure 10). SE33 was then included in the statistical calculation, accounted for by increasing the hypothesized mean difference to 2, and again no statistically significant support was found that 6C provided more

allelic information, with a p-value of 0.107 (Table 10). Even though more alleles were seen in samples amplified with Fusion 6C, the difference in allele counts was not statistically significant.

Next, each individual locus that 5C and 6C share were analyzed for significant differences in allele counts for degraded samples using one-tailed paired T-tests ($H_A = \mu_{1(5C)} < \mu_{2(6C)}$). This yielded two autosomal loci with significant differences in means. The first was D10S1248, with an average autosomal allele count of 1.313 for Fusion 5C and 1.625 for Fusion 6C. The p-value was 0.0480, significant with 0.05 as the α value (Table 11). D10S1248 is in the same location and color channel in both kits, and future work will be needed to specifically assess why this locus performed more optimally in Fusion 6C. The second locus with a significant difference in means was FGA, with an average autosomal allele count of 1.438 for Fusion 5C and 1.750 for Fusion 6C, and a p-value of 0.0270 (Table 11). FGA is located in the purple channel of Fusion 6C and has a smaller product size than its counterpart in Fusion 5C, which explains the success in amplifying in severely degraded samples.

Lastly, DYS391, a Y-STR present in both 5C and 6C, showed a significant increase in average alleles in Fusion 6C with a p-value of 0.0002 (Table 11). The average allele count in Fusion 5C was 0.3130, and the average allele count in Fusion 6C was 0.875. This indicates that more than double the amount of allelic information is gained when using Fusion 6C to amplify DYS391 in the same degraded samples. This Y-STR is in a new location in Fusion 6C, located in the purple channel and with a smaller product size. Again, since the product size of DYS391 is much smaller than in Fusion 5C, it is intuitive that it would amplify more successfully in degraded samples.

In this study, it has been shown that Fusion 6C is more successful at amplifying three loci in degraded samples than Fusion 5C. An improved performance with the three individual loci

discussed is interesting, especially since it did not result in a significant difference overall. This could be the result of a small sample size (n=16), and future research comparing the ability of Fusion 6C to amplify degraded samples may provide more clarity. This could be significant, as often forensic casework samples are severely compromised and degraded. SE33, DYS570, and DYS576 were not included in the direct comparison of individual loci since Fusion 5C does not have them. These can further discriminate the DNA profile.

3.10. Rapidly Mutating Y-STRs

The two rapidly mutating Y-STRs, DYS570 and DYS576, included in Fusion 6C were examined for the ability to distinguish between male related individuals. Samples from a father and his two sons were processed using Fusion 6C. Genotypes at these two loci were determined and compared. The results showed that all three of the profiles were consistent at the Y-STR loci. Based on this, there is not enough information to determine how effective these two loci are at distinguishing related male individuals, however more extensive studies have demonstrated the utility of rapidly mutating loci in discriminating between closely and distantly related male individuals (32). This is a very small sample size, and more work should be done in the future to determine the utility of rapidly mutating Y-STRs included in multiplex kits.

Conclusions

The research objective was to investigate performance metrics of Fusion 5C and Fusion 6C, initiating a direct comparison, that could then be used to determine whether or not the Virginia Department of Forensic Science should consider the validation and implementation of Fusion 6C. It is not worthwhile for the laboratory to make the change to Fusion 6C if the only notable advantage is the additional autosomal locus and two Y-STRs. In order to devote the time

and resources necessary to implement a new amplification kit, there had to be significant advantages to using Fusion 6C.

Based on the results of this study, Fusion 6C demonstrated possibly less baseline noise, and as a result, a slightly lower analytical threshold setting. This can greatly aid in peak determination, and the lower threshold may allow more potential alleles to be called. Fusion 6C utilizes a short cycling protocol, which speeds up the analysis workflow, and is an advantage in time savings. Another benefit of using Fusion 6C for amplification when compared to Fusion 5C is the less ambiguous pull-up observed. This also saves time in analysis, lessening the challenge of allele determination. Lastly, this study found that Fusion 6C provided significantly more information in four-person mixtures than Fusion 5C. Since forensic samples can often be challenging mixtures, any additional allelic information is valuable.

Fusion 5C also provides some benefits over Fusion 6C. A major advantage of Fusion 5C is the increased sensitivity observed in this study. Increasing the sensitivity of amplification kits has been a goal of the manufacturers since the increase in CODIS core loci. The values of sensitivity should not be understated. However, there is a slight disadvantage to over sensitivity expressed in the form of instrument noise and artifacts, which was initially one of the complaints reported regarding Fusion 5C. So, while the sensitivity of Fusion 5C over Fusion 6C is an advantage, it may also be the cause of the increased noise and artifacts. Another benefit of Fusion 5C is less bin overlap observed between color channels. This can result in less chance that pull-up will fall into an allele calling bin of another locus. This is likely the cause of the higher number of pull-up caught during human analysis of three- and four-person mixtures samples using Fusion 6C. Again though, the pull-up was ultimately determined to be more ambiguous in Fusion 5C, which can complicate analysis.

Overall, supported by previous research, both Fusion 5C and Fusion 6C are robust, sensitive, and advantageous typing kits for STR amplification. This study has shown that Fusion 6C possesses additional advantages other than the increased power of discrimination from SE33, DYS570, and DYS576. Based on the performance metrics and critical aspects determined in this study, a recommendation to consider the switch to Fusion 6C from Fusion 5C will be submitted to the VDFS.

This comparative study provides a preliminary investigation into the feasibility of implementing Fusion 6C as the STR amplification kit at the Virginia Department of Forensic Science, with broad applications for other forensic laboratories. As STR amplification kits continue to improve and update, other laboratories may find a similar dilemma with new amplification kits and upgrades to other technologies. In this case, Fusion 6C demonstrated several critical aspects that can provide additional benefits to using it for STR amplification. Making any change in the DNA workflow requires a full validation effort, and this is costly. By performing this pre-validation study, it provides these other laboratories both with the specific results from this research, as well as a model for future considerations.

This research identifies an additional area of study to be further addressed. As seen in the results, Fusion 5C was more sensitive, but resulted in slightly higher baseline noise and more ambiguous pull-up. Fusion 6C, on the other hand, was less sensitive but demonstrated decreased artifacts. As technology continues to improve, both in amplification kits as well as instrumentation, DNA analysis will have the capabilities to be even more sensitive. Future research should target this and determine where the line between sensitivity and decreased artifacts is most optimal.

References

1. Lin S, Li C, Ip SCY. A selection guide for the new generation 6-dye DNA profiling systems. Forensic Sci Int Genet 2017; 30:34-42. https://doi.org/10.1016/j.fsigen.2017.05.010

2. Tan JYY, Tan YP, Ng S, Tay AS, Phua YH, Tan WJ, et al. A preliminary evaluation study of new generation multiplex STR kits comprising of the CODIS core loci and the European Standard Set loci. J Forensic and Legal Med 2017; 52:16-23. https://doi.org/10.1016/j.jflm.2017.07.017

3. Hares DR. Selection and implementation of expanded CODIS core loci in the United States. Forensic Sci Int Genet 2015; 17:33-34. https://doi.org/10.1016/j.fsigen.2015.03.006

4. Oostdik K, Lenz K, Nye J, Schelling K, Yet D, Bruski S, et al. Developmental validation of the PowerPlex® Fusion System for analysis of casework and reference samples: A 24-locus multiplex for new database standards. Forensic Sci Int Genet 2014; 12: 69-76. https://doi.org/10.1016/j.fsigen.2014.04.013

5. Oostdik K, Ensenberger M, Sprecher C, Bourdeau-Heller J, Krenke B, Storts D. Bridging Databases for Today and Tomorrow: The PowerPlex® Fusion System. Promega Corporation. 2012. https://www.promega.com/resources/profiles-in-dna/2012/bridging-databases-for-today-and-tomorrow-the-powerplex-fusion-system/

6. Pfoser K, Owen S. Evaluation of the PowerPlex® Fusion System for Use on the ABI PRISM® 310 Genetic Analyzer. Promega Corporation. Updated 2013. https://www.promega.com/resources/profiles-in-dna/2013/evaluation-of-the-powerplex-fusion-system-for-use-on-the-abi-prism-310-genetic-analyzer/

7. Promega Corporation. PowerPlex® Fusion System for Use on the Applied Biosystems™ Genetic Analyzers Technical Manual. Madison, WI: Promega Corporation, Revised 04/2017. https://www.promega.com/~/media/files/resources/protocols/technical%20manuals/101/powerplex%20fusion%20system%20protocol.pdf

8. Thermo Fisher Scientific, Inc. Applied Biosystems ™ 3500/3500xL Genetic Analyzer User Guide. Waltham, MA:ThermoFisher Scientific, 2018. https://assets.thermofisher.com/TFS-Assets/LSG/manuals/100031809_3500_3500xL_Software_v3_1_UG.pdf

9. Promega Corporation. PowerPlex® Fusion 6C System Technical Manual. Madison, WI: Promega Corporation, Revised 01/2015. https://www.promega.co.jp/lit/pdf/PowerPlexFusion6c.pdf

10. Cisana C, Cerri N, Bosetti A, Verzeletti A, Cortellini V. PowerPlex® Fusion 6C System: evaluation study for analysis of casework and database samples. Croat Med J 2017; 58:26–33. 10.3325/cmj.2017.58.26

11. Borsuk LA, Gettings KB, Steffen CR, Kisler KM, Vallone PM. Sequence-based U.S. population data for the SE33 locus. Electrophoresis 2018; 39(21):2694-2701. https://doi.org/10.1002/elps.201800091

12. Ensenberger MG, Lenz KA, Matthies LK, Hadinoto GM, Schienman JE, Przech AJ, et al. Developmental validation of the PowerPlex® Fusion 6C System. Forensic Sci Int Genet 2016; 21: 134-144. https://doi.org/10.1016/j.fsigen.2015.12.011

13. Boavida A, Bogas V, Sampaio L, Gouveia N, Porto MJ, Corte-Real F. PowerPlex® Fusion 6C system: internal validation study. Forensic Sci Research 2018; 3:2: 130-137. 10.1080/20961790.2018.1430471

14. Tay JW, Murakami JA, Cooper PL, Rye MS. Sensitivity and baseline noise of three new generation forensic autosomal STR kits: PowerPlex® Fusion, VeriFilerTM Plus and Investigator® 24plex QS. Forensic Sci Int: Reports 2019; 1: 100049. https://doi.org/10.1016/j.fsir.2019.100049

15. Greenspoon SA, Schiermeier-Wood L, Jenkins BC. Establishing the Limits of TrueAllele® Casework: A Validation Study. J Forensic Sci 2015; 60: 1263-1276. doi:10.1111/1556-4029.12810

16. Bregu J, Conklin D, Coronado E, Terrill M, Cotton RW, Grgicak CM. Analytical Thresholds and Sensitivity: Establishing RFU Thresholds for Forensic DNA Analysis. J Forensic Sci 2013; 58 (1): 120-29. https://doi-org.proxy.library.vcu.edu/10.1111/1556-4029.12008

17. Wang ZF, Dai SP, Lian JY, Chen HF, Ye WH, Cao HL. Allele Size Miscalling due to the Pull-Up Effect Influencing Size Standard Calibration in Capillary Electrophoresis: A Case Study Using HEX Fluorescence Dye in Microsatellites. Genotyping 2018; 3: 31-46. 10.5772/intechopen.73028 https://www.researchgate.net/publication/325879587_Allele_Size_Miscalling_due_to_the_PullUp_Effect_Influencing_Size_Standard_Calibration_in_Capillary_Electrophoresis_A_Case_Study_Using_HEX_Fluorescent_Dye_in_Microsatellites

18. Forensic Biology Procedures Manual 210-D2016. Interpretation of PowerPlex® Fusion CE Data. Virginia Department of Forensic Science. Revised December 23, 2019. https://www.dfs.virginia.gov/wp-content/uploads/2020/01/210-D2016-FB-PM-Interpretation-of-Fusion-Data.pdf

19. Forensic Biology Procedures Manual 210-D2020. TrueAllele® Casework System. Virginia Department of Forensic Science. Revised September 26, 2017. https://www.dfs.virginia.gov/wp-content/uploads/2017/10/210-D2020-FB-PM-TrueAllele-Casework-System.pdf

20. Quintin DM, Scheinman JE, Adamowicz DSP. Assessment of PowerPlex® Fusion 5C's Ability to Type Degraded DNA. Science & Justice 2020; 60(5): 423-431. https://doi.org/10.1016/j.scijus.2020.05.004

21. Evans J, Hadi S. Low Template DNA: Tad, Touch and Traces. J Foren Path 2018; 3:112. https://www.longdom.org/open-access/low-template-dna-tad-touch-and-traces.pdf

22. Butler JM. Advanced Topics in Forensic DNA Typing: Methodology. Waltham, MA: Elsevier Inc., 2012.

23. Willuweit S, Roewer L. YHRD: Y-Chromosome STR Haplotype Reference Database. January 7, 2021. https://yhrd.org/

24. Carracedo A, Lareu MV. Deoxyribonucleic Acid | Significance. Oxford, UK: Elsiever Inc., Encyclopedia of Forensic Sciences 2000; pp 541-545. https://doi.org/10.1006/rwfs.2000.0473

25. Forensic Biology Procedures Manual 210-D2004. Extraction of DNA. Virginia Department of Forensic Science. June 30, 2020. https://www.dfs.virginia.gov/wp-content/uploads/2020/07/210-D2004-FB-PM-Extraction-of-DNA.pdf

26. Promega Corporation. PowerQuant® System Technical Manual. Madison, WI: Promega Corporation, Revised 01/2020. https://www.promega.com/resources/protocols/technical-manuals/101/powerquant-system-protocol/

27. Forensic Biology Procedures Manual 210-D2007. PowerPlex® Fusion Amplification and Long Term Storage. Virginia Department of Forensic Science. June 30, 2020. https://www.dfs.virginia.gov/wp-content/uploads/2020/07/210-D2007-FB-PM-PP-FUSION-Amp-and-Storage.pdf

28. Forensic Biology Procedures Manual 210-D2010. CE For PowerPlex Fusion. Virginia Department of Forensic Science. December 27, 2017. https://www.dfs.virginia.gov/wp-content/uploads/2018/01/210-D2010-FB-PM-CE-Fusion-4050-QR2.pdf

29. Promega Corporation. PowerPlex Fusion 6C for Use on the Applied Biosystems® Genetic Analyzers Technical Manual. Madison, WI: Promega Corporation, Revised 12/2018. Accessed June 29, 2020. https://www.promega.com/~/media/files/resources/protocols/technical%20manuals/101/powerplex%20fusion%206c%20system%20protocol.pdf

30. National DNA Index System (NDIS) Operational Procedures Manual. Standards for Acceptance of PCR DNA Records at NDIS. FBI Laboratory. May 1, 2016; 1.4: 37. https://www.fbi.gov/file-repository/ndis-operational-procedures-manual.pdf/view

31. Pizzamiglio M, Marino A, Coli A, Floris T, Garofano L. The use of mini-STRs on degraded DNA samples. International Congress Series 2006; 1288: 498-500. https://doi.org/10.1016/j.ics.2005.11.111

32. Ballantyne KN, Keerl V, Wollstein A, Choi Y, Zuniga SB, Ralf A, et al. A New Future of Forensic Y-chromosome Analysis: Rapidly Mutating Y-STRs for Differentiating Male Relatives and Paternal Lineages. Forensic Sci Int Genet 2011; 6.2: 208-18. https://doi.org/10.1016/j.fsigen.2011.04.017

Appendix

Table 1: Cycling Parameters for Fusion 5C and Fusion 6C.

Fusion 5C (12.5 µL)	Fusion 6C (25 µL)	Fusion 6C (12.5 µL)
96˚C for 1 minute	96˚C for 1 minute	96˚C for 1 minute
94˚C for 10 seconds 59˚C for 1 minute 72˚C for 30 seconds for 28 cycles	96˚C for 5 seconds 60˚C for 1 minute for 29 cycles	96˚C for 5 seconds 60˚C for 1 minute for 28 cycles
60˚C for 10 minutes 4˚C soak	60˚C for 10 minutes 4˚C soak	60˚C for 10 minutes 4˚C soak

Table 2: Limit of Detection Determination for Fusion 6C.

12s Full	MAX	LOD (MAX x 2)	Average
Blue	31.00	62.00	17.96
Green	29.00	58.00	23.61
Yellow	31.00	62.00	16.07
Red	31.00	62.00	22.25
Purple	30.00	60.00	15.82
24s Full	**MAX**	**LOD (MAX x 2)**	**Average**
Blue	27.00	54.00	17.46
Green	31.00	62.00	23.25
Yellow	51.00	102.00	17.71
Red	32.00	64.00	21.36
Purple	20.00	40.00	15.68
12s Half	**MAX**	**LOD (MAX x 2)**	**Average**
Blue	18.00	36.00	15.13
Green	29.00	58.00	21.63
Yellow	14.00	28.00	12.00
Red	26.00	52.00	20.00
Purple	16.00	32.00	13.50
24s Half	**MAX**	**LOD (MAX x 2)**	**Average**
Blue	18.00	36.00	15.22
Green	26.00	52.00	22.44
Yellow	15.00	30.00	12.67
Red	24.00	48.00	21.00
Purple	19.00	38.00	14.78

Table 3: Analytical Thresholds for Both 12 Seconds and 24 Seconds for Fusion 6C Full-Volume and Half-Volume Reactions.

Color Channel	Half Reaction Analytical Threshold (RFU)	Full Reaction Analytical Threshold (RFU)
Blue	40	65
Green	60	65
Yellow	40	102
Red	60	65
Purple	40	65

Table 4: Number of Overlapping Bins in Fusion 5C and Fusion 6C.

Color Channel	5C Overlapping Bins	6C Overlapping Bins	6C Overlapping Bins Including Three Additional Loci
Blue	121	142	182
Green	97	133	163
Yellow	142	94	119
Red	113	103	165
Purple	N/A	60	146

Table 5: Pull-up in Two-Person Mixtures.

	Fusion 5C	Sample	Fusion 6C
After GMIDX Analysis			
A	D8S1179 13→TH01 10.3 D19S433 13→D21S11 29	**07-574 Item 3**	TPOX 8→SE33 38
	None	**10-581 Item 3**	None
	TH01 9.3→D3S1358 9 D8S1179 12→D16S539 8.3 D8S1179 8→TH01 5.3	**10-581 Item 4**	D13S317 12 → D5S818 9
	D3S1358 14→D16S539 14	**11-584 Item 3**	None
	None	**12-571 Item 4**	D19S433 14 + D21S11 30 → D2S441 12 TPOX 8→SE33 38
	D8S1179 13→TH01 10.3	**08-574 Item 4**	None
After Human Analysis			
B	None	**07-574 Item 3**	TPOX 8→SE33 38
	None	**10-581 Item 3**	None
	TH01 9.3→D3S1358 9	**10-581 Item 4**	None
	D3S1358 14→D16S539 14	**11-584 Item 3**	None
	None	**12-571 Item 4**	TPOX 8→SE33 38
	None	**08-574 Item 4**	None

Table 6: Paired Two-Sample T-Test Comparing Autosomal Allele Counts for Fusion 5C and Fusion 6 for Two-Person Mixture Samples.

A. 12 Sec	*5C*	*6C*
Mean	42.583	42.063
Variance	478.163	436.188
Observations	48.000	48.000
Pearson Correlation	0.953	
Hypothesized Mean Difference	0.000	
df	47.000	
t Stat	0.547	
P(T<=t) one-tail	0.294	
t Critical one-tail	1.678	
P(T<=t) two-tail	0.587	
t Critical two-tail	2.012	
B. 24 Sec	*5C*	*6C*
Mean	46.000	46.500
Variance	499.132	497.915
Observations	54.000	54.000
Pearson Correlation	0.970	
Hypothesized Mean Difference	0.000	
df	53.000	
t Stat	-0.671	
P(T<=t) one-tail	0.253	
t Critical one-tail	1.674	
P(T<=t) two-tail	0.505	
t Critical two-tail	2.006	

Table 7: Paired Two-Sample T-Test for Comparing Autosomal Allele Counts for Fusion 5C and 6C for Three- and Four-Person Mixture Samples.

Three-Person	*5C*	*6C*
Mean	81.667	77.333
Variance	20.250	51.500
Observations	9.000	9.000
Pearson Correlation	0.197	
Hypothesized Mean Difference	0.000	
df	8.000	
t Stat	1.692	
P(T<=t) one-tail	0.065	
t Critical one-tail	1.860	
P(T<=t) two-tail	0.129	
t Critical two-tail	2.306	
Four-Person	*5C*	*6C*
Mean	94.917	97.917
Variance	29.538	18.265
Observations	12.000	12.000
Pearson Correlation	0.540	
Hypothesized Mean Difference	0.000	
df	11.000	
t Stat	-2.180	
P(T<=t) one-tail	0.026	
t Critical one-tail	1.796	
P(T<=t) two-tail	0.052	
t Critical two-tail	2.201	

Table 8: Autosomal Allele Counts for Three-Person Mixtures with Fusion 5C (left) and Fusion 6C (right).

Sample	Total Alleles Expected	Total Alleles Seen	Percent Profile
1	86	85	98.84
2	86	85	98.84
3	86	84	97.67
4	86	81	94.19
5	86	86	100.00
6	86	85	98.84
7	81	79	97.53
8	81	73	90.12
9	81	77	95.06

Sample	Total Alleles Expected	Total Alleles Seen	Percent Profile
1	86	67	77.91
2	86	84	97.67
3	86	86	100.00
4	86	72	83.72
5	86	71	82.56
6	86	86	100.00
7	81	81	100.00
8	81	72	88.89
9	81	77	95.06

Higher Percent Profile Than Other Kit
Lower Percent Profile Than Other Kit
Same Percent Profile

Table 9: Autosomal Allele Counts for Four-Person Mixtures with Fusion 5C (left) and Fusion 6 C (right).

Sample	Total Alleles Expected	Total Alleles Seen	Percent Profile
10	103	101	98.06
11	103	95	92.23
12	103	102	99.03
13	103	89	86.41
14	97	89	91.75
15	97	86	88.66
16	97	93	95.88
17	97	90	92.78
18	102	100	98.04
19	102	99	97.06
20	102	99	97.06
21	102	96	94.12

Sample	Total Alleles Expected	Total Alleles Seen	Percent Profile
10	103	95	92.23
11	103	98	95.15
12	103	100	97.09
13	103	103	100.00
14	97	94	96.91
15	97	91	93.81
16	97	94	96.91
17	97	93	95.88
18	102	102	100.00
19	102	102	100.00
20	102	101	99.02
21	102	102	100.00

Higher Percent Profile Than Other Kit
Lower Percent Profile Than Other Kit

Table 10: Paired Two Sample T-test for Means of Total Number of Alleles between Degraded Samples using Fusion 5C and Fusion 6C (One-Tail).

Condition	P-Value	Hypothesized Mean Difference	Significant Difference (α=0.05)
5C and 6C without SE33	0.319	0	No
5C and 6C with SE33	0.107	2	No

Table 11: Paired Two Sample T-test Performed on Each Locus Present between Fusion 5C and Fusion 6C with Significant Difference (One-Tail).

Loci with Significant Difference (α=0.05)	P-Value	Fusion 5C Mean (Average Detected Alleles)	Fusion 6C Mean (Average Detected Alleles)
D10S1248	0.0480	1.313	1.625
FGA	0.0270	1.438	1.750
DYS391	0.0002	0.313	0.875

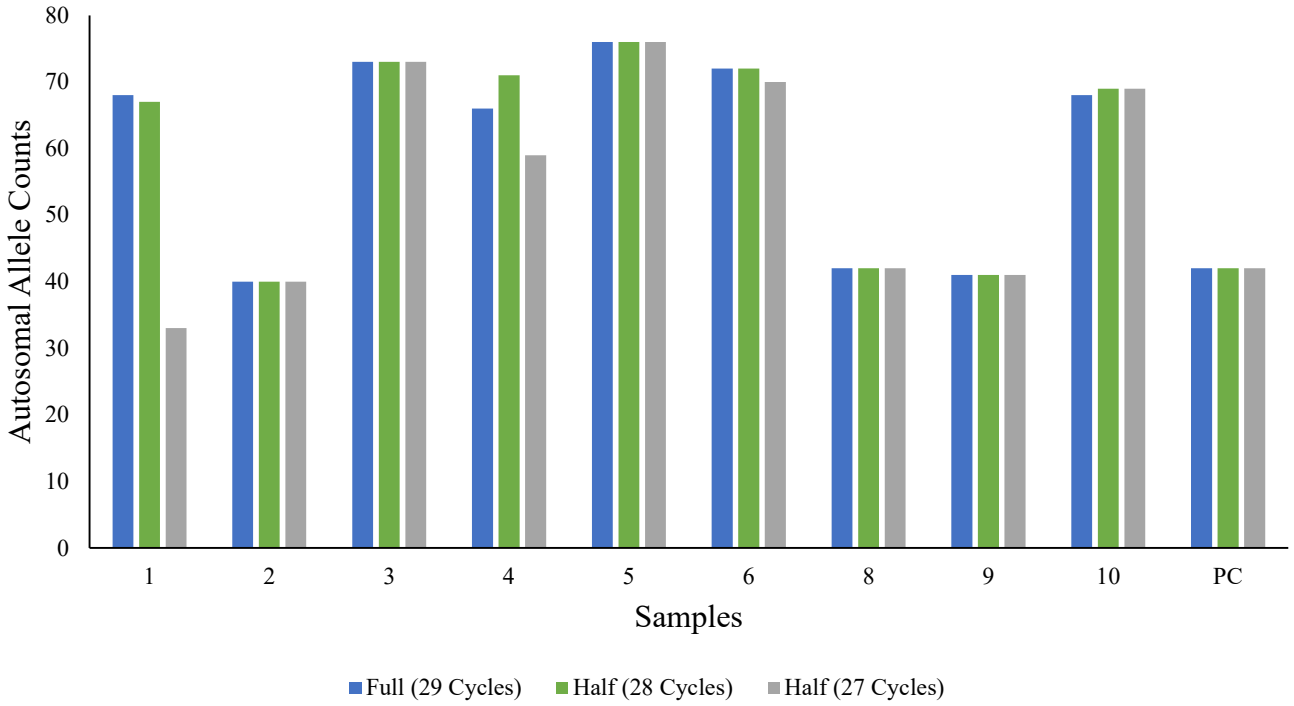

Figure 1: Comparison of Allele Counts from Fusion 6C Full- and Half-Volume Reactions. Full-volume reactions using a 29-cycle process were compared to half-volume reactions with 27- and 28- cycle processes (n=10).

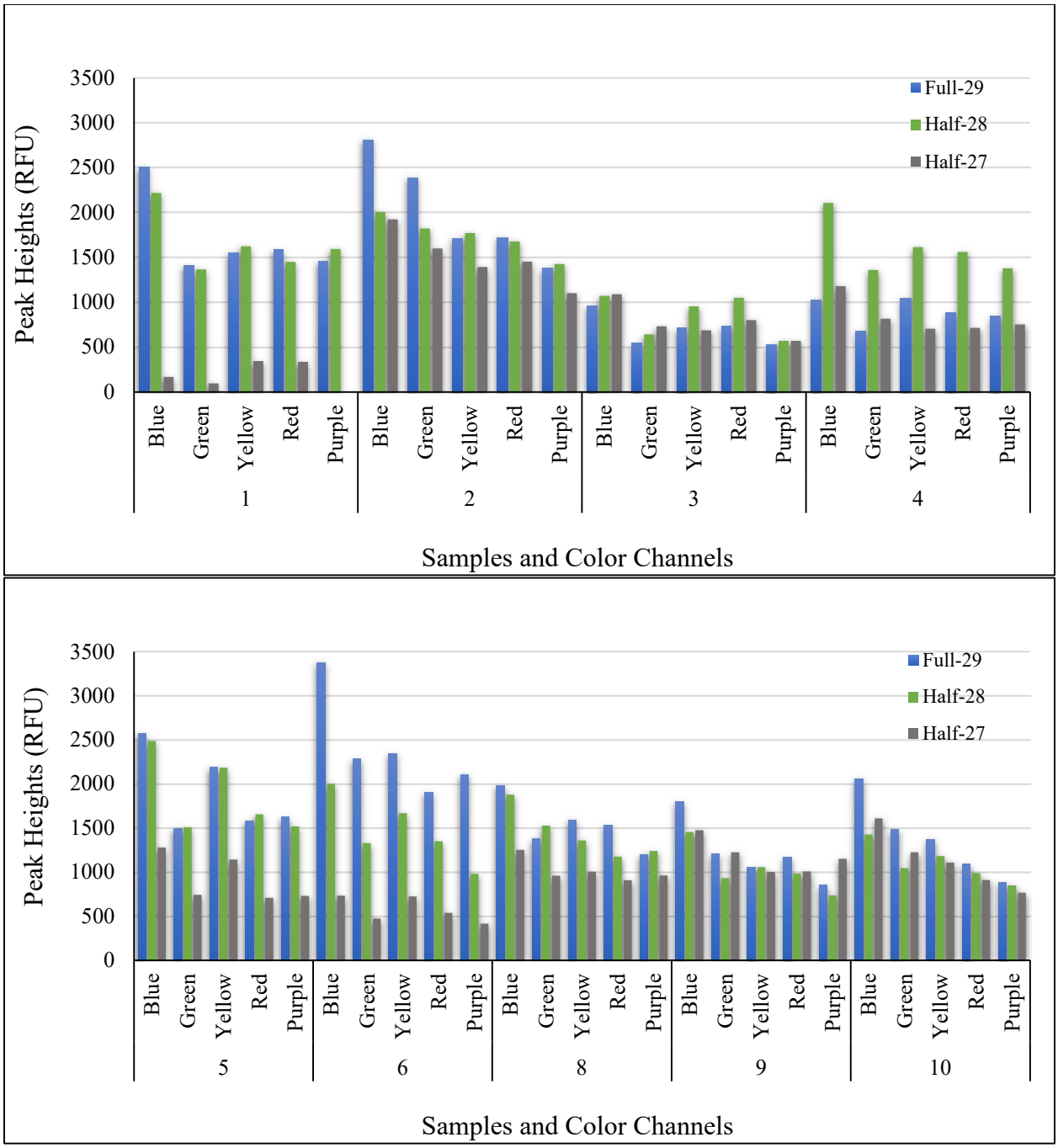

Figure 2: Comparison of Peaks Heights from Fusion 6C between Samples Amplified with a Full-Volume Reaction at 29 Cycles, and Half-Volume Reactions with 27- and 28 Cycles. Average peak heights across color channels were compared between the differing cycling protocols to assess which half-volume reaction was most similar to the full-volume.

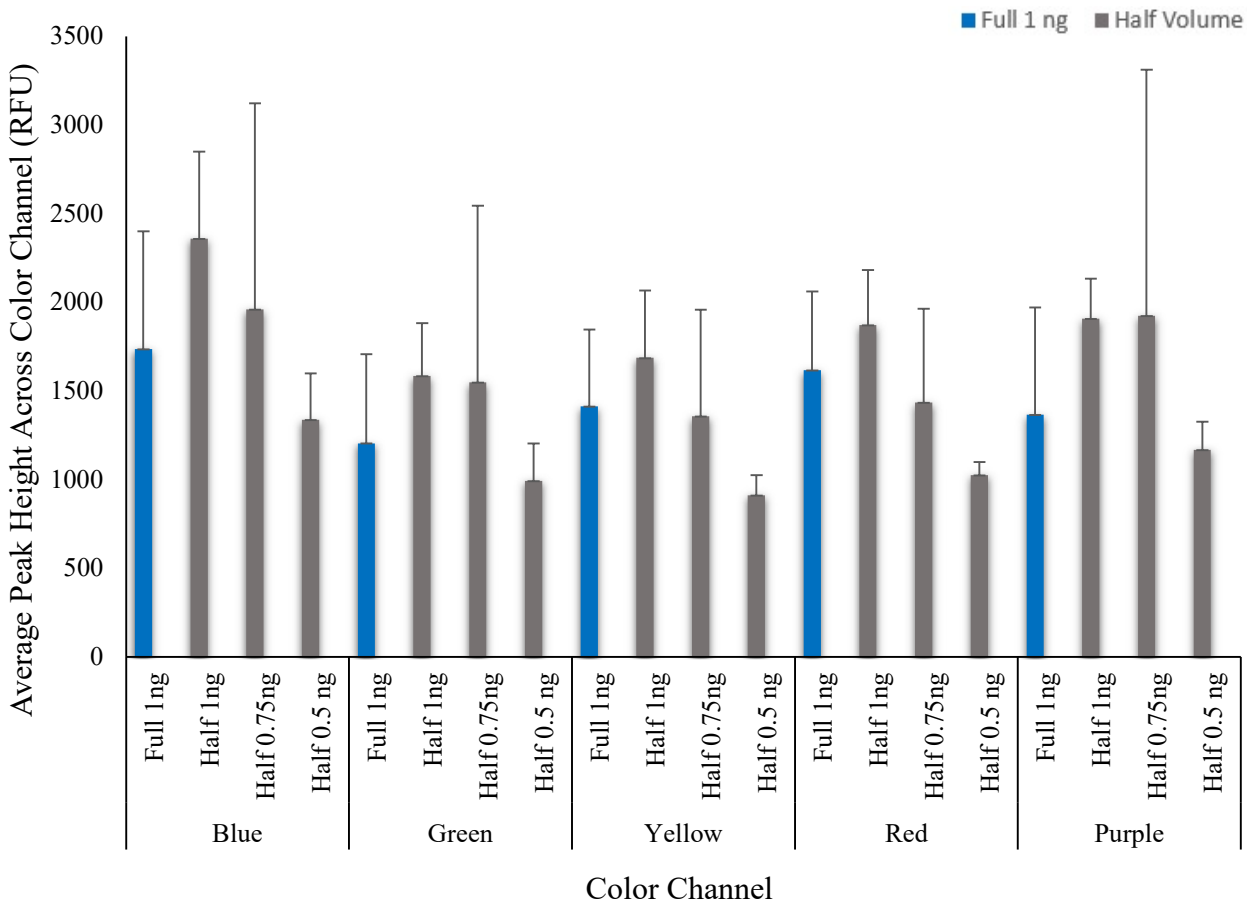

Figure 3: Average Peak Heights (RFU) Per Color Channel in Fusion 6C (12s Injection) Based Upon Template Quantity. This graph depicts a full-volume reaction with 1 ng, and three half-volume reactions with 1 ng, 0.75 ng, and 0.5 ng. Standard deviation is shown as error bars.

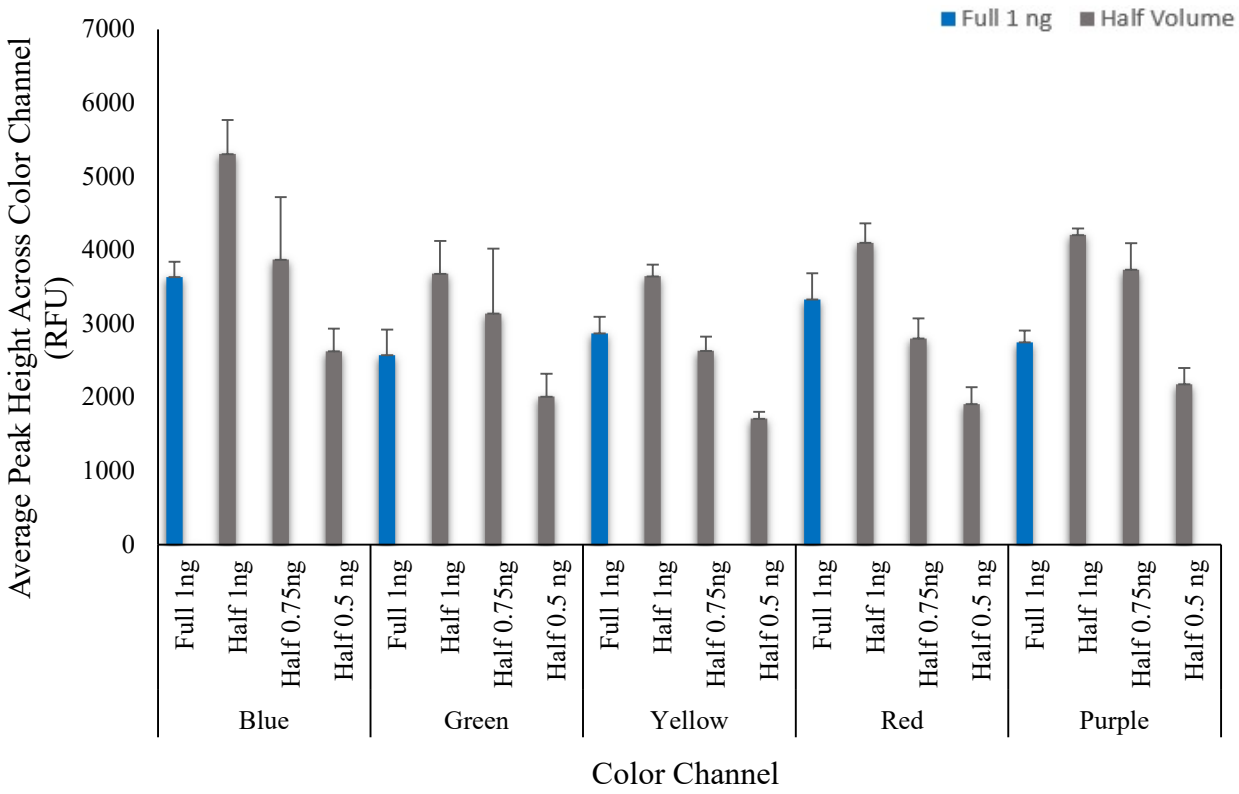

Figure 4: Average Peak Heights (RFU) Per Color Channel in Fusion 6C (24s Injection) Based Upon Template Quantity. This graph depicts a full-volume reaction with 1 ng, and three half-volume reactions with 1 ng, 0.75 ng, and 0.5 ng. Standard deviation is shown as error bars.

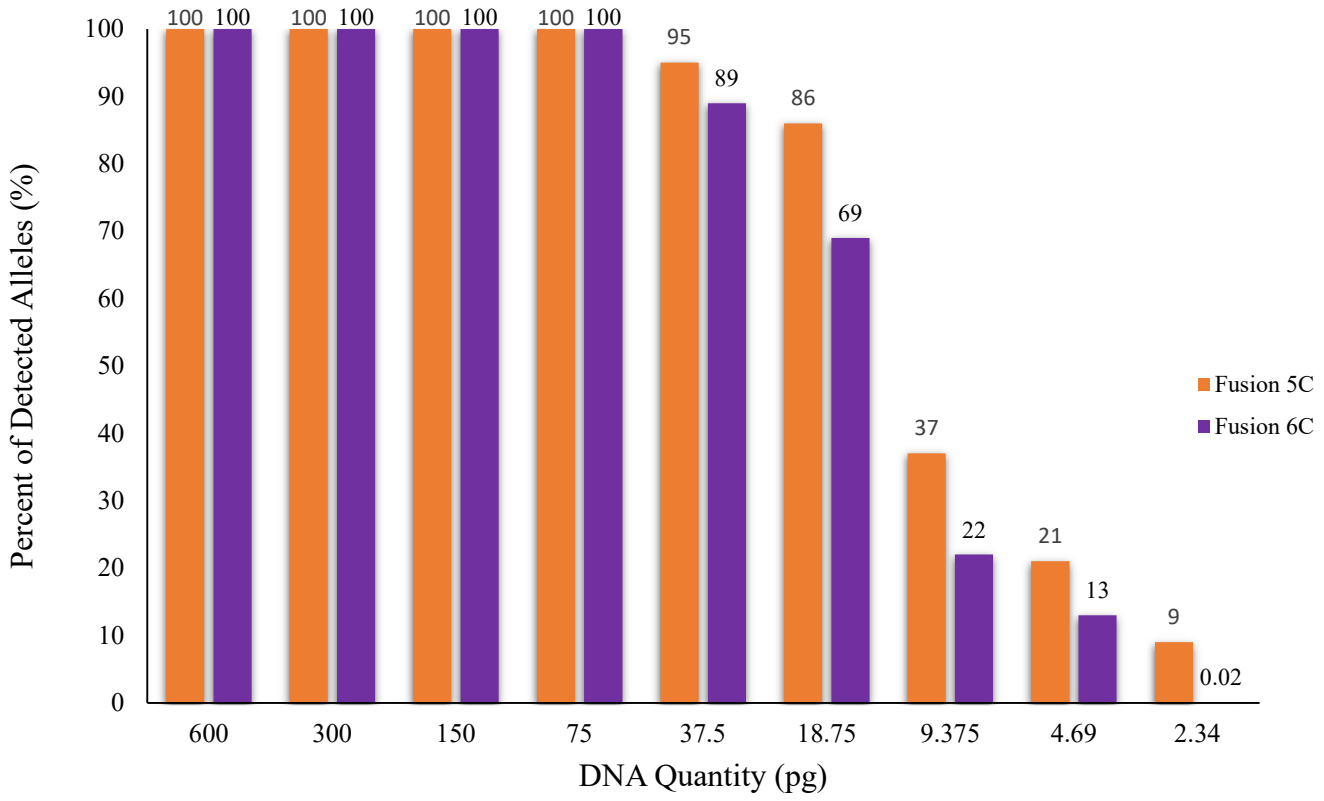

Figure 5: Percent Alleles Detected in Fusion 5C and 6C. Autosomal allele counts from Fusion 5C and Fusion 6C half-volume reactions were used to calculate percent of alleles detected (n=1).

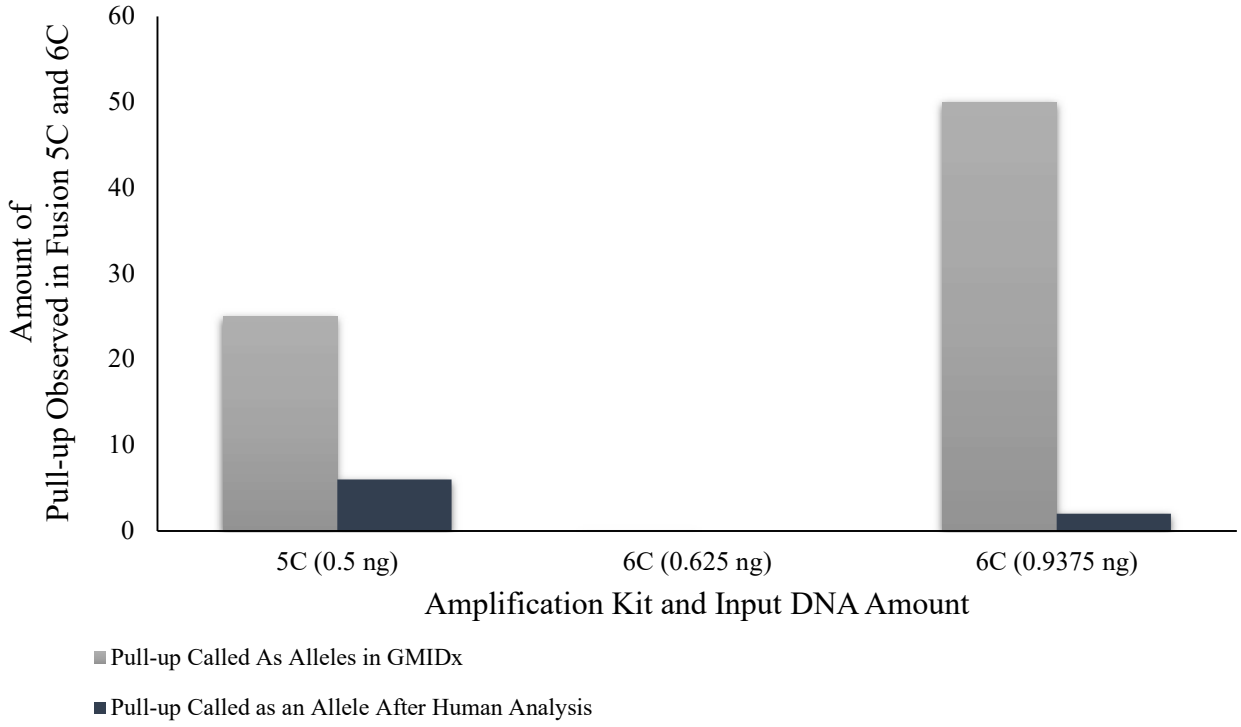

Figure 6: Number of Pull-Up Events and False Alleles Calls Using Fusion 5C and Fusion 6C in Three- and Four-Person Mixtures.

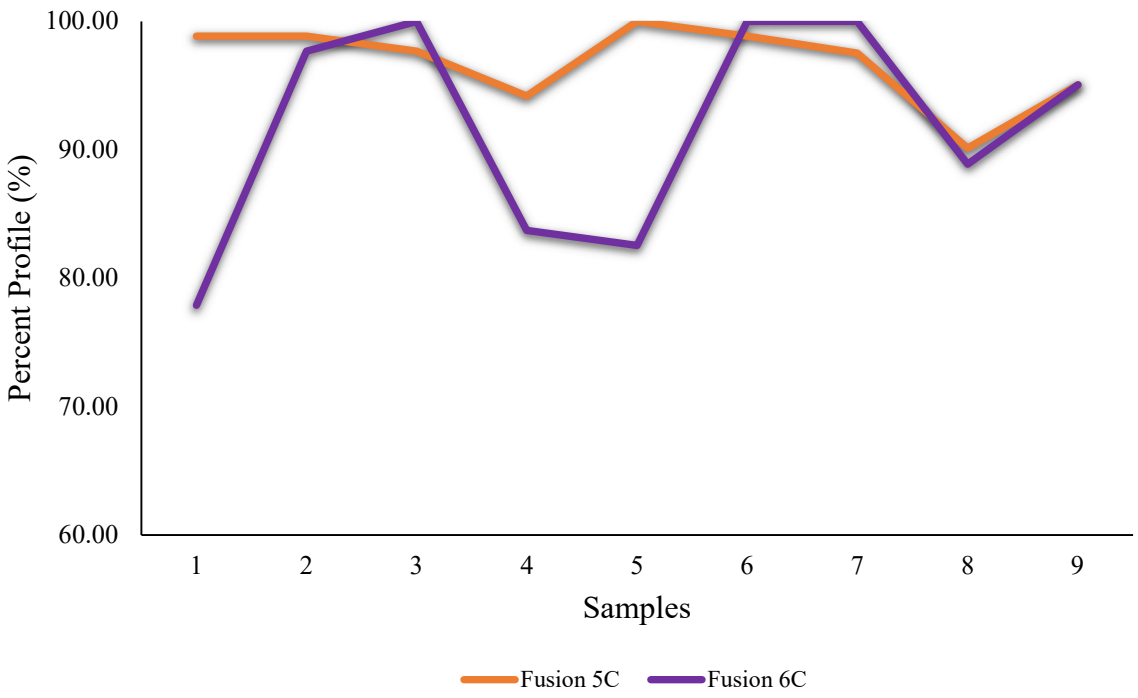

Figure 7: Percent Profiles of Three-Person Mixture Samples using Fusion 5C and Fusion 6C. In three-person mixtures, Fusion 5C was more consistent across samples and provided more allelic information in a higher number of samples than Fusion 6C.

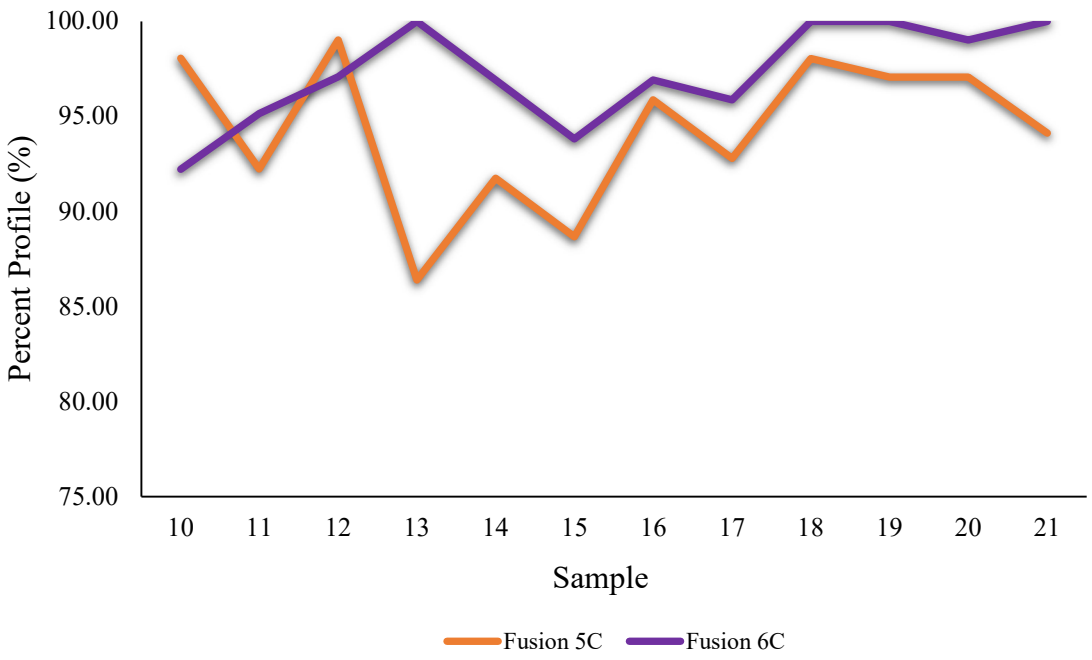

Figure 8: Percent Profile of Four-Person Mixture Samples Using Fusion 5C and Fusion 6C. Fusion 6C amplified samples were more consistent and provided more allelic information than Fusion 5C.

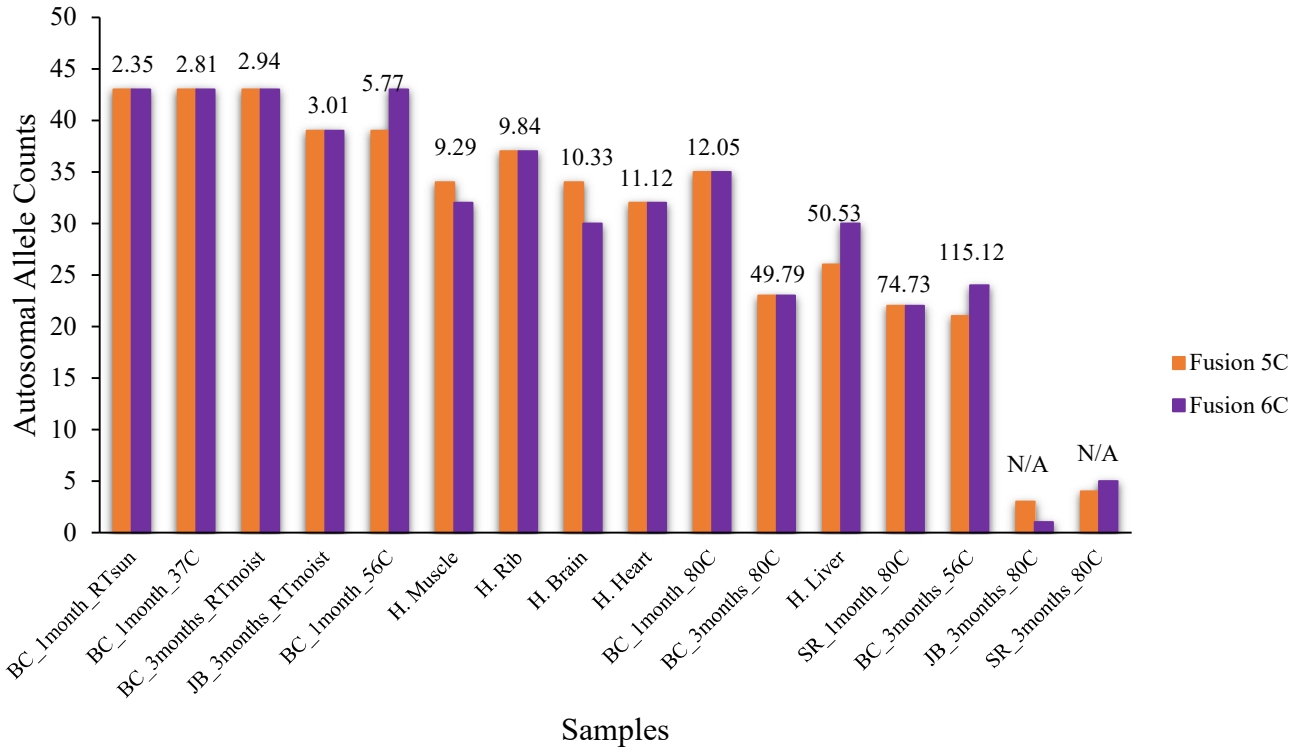

Figure 9: Degraded Sample Autosomal Allele Counts Without SE33 in Fusion 5C and Fusion 6C. The numbers above the bars indicate degradation values determined by Promega's PowerQuant® assay. As the degradation index increases, the allele counts of the sample decreases. Kits are providing similar allelic information, with Fusion 6C showing more observed alleles in 4 samples, and Fusion 5C showing more observed alleles in 3 samples.

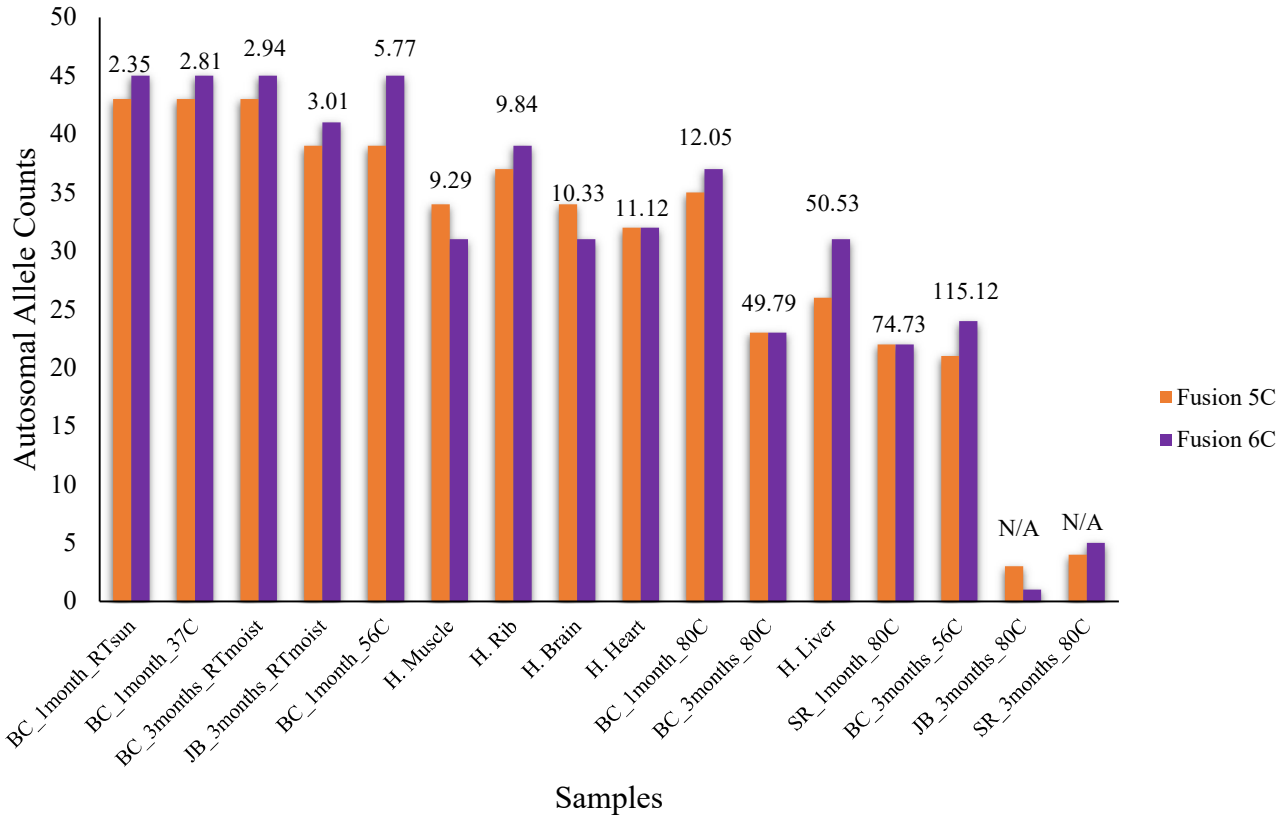

Figure 10: Degraded Sample Autosomal Allele Counts with SE33 in Fusion 5C and Fusion 6C (n=16). Samples amplified with Fusion 6C show higher allele counts for most of the samples.

Vita

Caitlin McCaughan was raised in New Kent, Virginia. She received her Bachelor of Science in Biological Sciences from Virginia Tech in 2019. She was the recipient of the 2018 Stacey Smith Research Excellence Award. Currently, Caitlin is pursuing a Master of Science in Forensic Science with a concentration in Forensic Biology from Virginia Commonwealth University. Caitlin is also employed at the Virginia Department of Forensic Science as an Administrative Specialist working on the Historical Case File Project.