COVER PAGE

National Institute of Justice
FY15 Research and Development in Forensic Science for Criminal Justice Purposes

Award 2015-MU-MU-K026
qPCR Genotype Determination and Mixture Detection Using High Resolution Melting Curve Analysis of STR Loci

Tracey Dawson Cruz, PhD
Professor, Department of Forensic Science
Virginia Commonwealth University
(804) 828-0642
tcdawson@vcu.edu

Andrea J. Publow, MFA, CRA
Dir. of Sponsored Programs – Gov't/NonProfit
Office of Research
Virginia Commonwealth University
(804) 828-6772
dirospa@vcu.edu

Final Draft Summary Overview

DUNS:
EIN:

Virginia Commonwealth University
P.O Box 980568
800 East Leigh Street, Suite 3200
Richmond, VA 23298-0658

VCU PT No. PT112125/FP2415
VCU Index No. 546883

Project/Grant Period: 1/1/2016-9/30/2018 (extended, no-cost)

Signature of Submitting Official

11/28/2018
Date

Introduction

The current method of generating a forensic DNA profile utilizes the PCR amplification of short tandem repeats (STRs) to allow for capillary electrophoresis (CE)- based detection of alleles at specific loci (1). The use of the PCR technique has enabled analysis of samples containing degraded or trace amounts of DNA, such as typing of DNA extracted from saliva on cigarette butts (2). Following the increase in sensitivity of analytical techniques is the subsequent increase in submission of 'touch DNA' samples that come from the transferal of skin cells that occur during contact with a surface (3). These touch DNA samples often contain low levels of template DNA (less than 100pg available for STR amplification) which becomes problematic during PCR as some target regions may be preferentially amplified over others. This can result in a number of undesirable effects that complicate profile interpretation such as allele drop-in, allele drop-out, and peak imbalances (4). Further, because these touch DNA sample originate from surfaces that may have been touched by numerous individuals, these samples are prone to mixtures – ie DNA present from more than one source contributor. When a mixture is present, along with low amounts of DNA, resulting data often includes one or more of the contributors' allele peaks falling below the analytical threshold, further confounding the profile interpretation process and often leading to "inconclusive" reporting.

In the current forensic laboratory workflow, both allele genotype assignment and mixture detection occur after DNA separation and detection (CE analysis) during the data review process when allele fragments are sized, allele values are assigned, and data quality evaluated. At this point, intra- and inter- locus imbalances and the presence of three or more peaks at multiple loci indicate to the examiner that multiple contributors are present in the DNA sample (5). Because this information is not available until the last step (end- point analysis), it is not possible to make earlier analytical adjustments to the protocols or workflow that may increase the likelihood of generating a profile with a distinguishable minor contributor. While reamplification of a low, mixed DNA sample may be possible, it is time consuming and risky – often providing little-to-no new information. Further, with low template or touch samples, the samples are more often consumed during initial testing leaving little remaining DNA for a second analysis. Additionally, no meaningful comparisons or conclusions can be made with respect to identification of the contributor of a DNA evidence sample until after the CE run (hours to overnight), import of the raw data into a genotyping software package, a careful data review by at least one trained examiner, STR profile interpretation and export, and formal reporting of the case conclusions – a process that can easily take weeks, from start (initial sample evaluation and serology) to finish (case report issued).

A pre-screening mixture detection assay that could also determine number of contributors (when mixtures are detected) and potentially provide early exclusionary information (for single-source samples, based on geno-group) earlier in the forensic DNA workflow would be useful to both the forensic DNA and investigative communities - particularly when sample consumption is a concern or when multiple surface swabs are available and could be potentially combined if a single contributor (of the same geno-group) can be indicated early (8). For example, the majority of the samples analyzed by the Bureau of Alcohol, Tobacco, Firearms, and Explosives (ATFE) are touch DNA samples collected from items such as guns, explosive debris, and ammunition (6). These often include multiple swabs from various areas of the evidence, for example, swabs from the trigger, safety lever, hand grips, slide, and/or hammer area are often collected as separate samples from a gun submitted for testing. Unfortunately, 50% of this type of evidence analyzed at the ATFE results in inconclusive results, low level data, and/or complicated mixtures with indistinguishable minor contributors (7). ATFE protocol currently dictates that low level DNA samples be concentrated down to 10µl and half of that total DNA yield be used for initial amplification. If a mixture is detected at the final CE/data analysis step and minor contributors are indistinguishable, there is no way to reamplify with more template, as only half of the DNA sample remains. However, if a pre-screening method were available early on in the

workflow to confirm number of contributors in a sample (single source or a mixture) then a request could be made for a deviation from the standard consumption policies to allow for more than half of the sample to be used in the initial amplification. By allowing for more of the sample to be used in the initial testing, there should be an increase in amplified product, thus increasing the likelihood that resulting allele peaks are above the analytical threshold and a major and minor contributor can be identified. Alternatively, if multiple sample swabs taken from the same evidence item can be determined early-on to be from a single contributor (sharing the same presumptive allele designations), examiners can more confidently combine DNA extracts from those items to increase the amount of template DNA available for STR amplification during the initial DNA testing, while avoiding the creation of accidental mixtures. Both scenarios (described above) include the use of screening information to redirect the DNA workflow in an effort to improve first round pass (success) rates associated with the testing of low level DNA and/or mixture samples, and would subsequently reduce retest rates. This, in turn, could save valuable examiner time as well as reduce consumable expenses. Lastly, any presumptive exclusionary typing information revealed during this early screening stage could provide early investigative leads that could help shape or redirect the path of the investigation without the long delays associated with the issuance of the final formal case report.

Objective

The overall objective of this project was to design an assay for mixture detection that could be multiplexed with the quantitation step of the forensic DNA workflow. The assay developed utilizes post-qPCR melt-curve analysis to detect the presence of double-stranded amplicon products from the targeted STRs (D5S818 and D18S51). The primary goal was to directly integrate this melt curve assay into existing commercially-available qPCR human DNA quantitation kits in order to accurately assign the sample to either a single-source geno-group or a geno-group that is typical of a mixed sample.

To achieve the goal noted above, the authors sought to explore two qPCR platforms, the Rotor-Gene® Q and the more commonly used ABI 7500, and several analytical approaches for the resulting melt curve data classification: 1) use of a commercially available principal component analysis (PCA)-based melt curve analysis software, 2) development and use of linear discriminate analysis (LDA) code written specifically for R statistical software, and 3) development and use of a novel support vector machines (SVM) software tool. After assessment of the tested qPCR platforms, selection of the best statistical approach, and integration of the melt curve assay into a commercially-available quantitation kit, the newly developed multiplex would need to be assessed for quantitation precision, geno-grouping concordance, and reproducibility.

In the initial proposal, the authors noted that, pending the outcome and success of this work, additional funding could be requested for the development of a web-based interface for an open-access online analysis tool that would allow for easy user import of melt curve data, selection of platform and kit settings, and automated classification (geno-grouping prediction) for both STR loci (D5S818, D18S51) as well as non-group classification (mixture detection) for determining if there are multiple DNA contributors or only a single contributor.

Materials and Methods

Sample Collection & STR Profile Generation

Buccal swab DNA extracts were utilized for this study and were collected according to a VCU-approved Institutional Review Board (IRB) protocol (HM20002931). Nearly 300 buccal samples were collected to identify a total of 10-20 samples that share genotypes for five to seven individual genotypes for both the D5S818 and D18S51 STR loci. The samples were extracted using a QIAcube liquid extraction robot (Qiagen) and the standard manufacturer's Buccal Swab Spin QIAcube Protocol

using QIAamp® DNA Blood Mini kit reagents (Qiagen) (9). STR profiles were developed by amplifying 1ng of DNA extract from each sample with the AmpFLSTR® Identifiler® PCR amplification kit (Life Technologies) on the GeneAmp 9600 thermal cycler (PerkinElmer, Waltham, MA). The 15µl reaction consisted of 5.7µl of PCR Reaction mix, 2µl of Primer set, 2.1µl Tris-EDTA (TE), 0.2µl of AmpliTaq™ Gold Polymerase (5U/µl) (Applied Biosystems, Foster City, California), and 5µl of template DNA. Thermal cycling parameters were: activation at 95°C for 11min followed by 28 cycles of 94°C denaturation for 60s, 59°C annealing for 60s, and 72° elongation for 60s, finished with a 60°C final extension for 90min. Amplified STR products were separated and detected on a 3130 Genetic Analyzer using a 36-cm capillary array (Applied Biosystems) and a 10s injection with an analytical threshold of 75 relative fluorescent units (RFUs). Each sample for CE analysis consisted of 0.1µl of GeneScan™ 500-LIZ™ size standard (Applied Biosystems) and 12µl of Hi-Di formamide (Life Technologies) diluent. The wells containing an allelic adder received 1µl and those containing samples included 1.5µl of amplified DNA. STR profiles were analyzed using GeneMapper ID™ software v4.1 (Applied Biosystems). Samples were sorted into known reference genotype groups based on the resulting D5S818 and D18S51 genotypes.

ABI 7500 – D5S818 & D18S51 Amplification & Melt Curve Detection

Before beginning testing on the ABI 7500 qPCR platform, both amplification reagent conditions and amplification/melt parameters had to be optimized for this instrument. To initially evaluate the ability of the ABI 7500 qPCR instrument to produce usable "dissociation" or melt curves, a small subset of samples representing three D5S818 genotypes [(11, 11), (11, 12), and (11, 13)] were amplified using the conditions established and described previously for the Rotor-Gene® Q instrument by Kuehnert et al. except that both sets of primers were included as a duplex for simultaneous amplification and melt (along with a decrease in water volume) (10,11). The amplification parameters, as well as reagent concentrations, selected for those studies are similar to those reported previously in the literature. Nicklas et al. found that a decrease in extension temperature from 72°C to 65°C, an increase of the primer concentration to 1 µM, and an increase of the MgCl2 to 3mM provided the best STR amplification curves and subsequent melts (12). Thus, the conditions for amplification used herein included a 38µl master mix composed of a 1X concentration of Taq Gold Buffer, 3mM MgCl2, 250µM dNTPs, 1µM of each forward and reverse primer, 1µM AmpliTaq Gold DNA polymerase (Applied Biosystems), and 5µM EvaGreen® intercalating dye (Biotium). Two microliters of template DNA were added to each reaction for a total reaction volume of 40µl. Primer sequences used for D5S818 amplification were (F) 5'-GGGTGATTTTCCTCTTTGGT-3' and (R) 5'-AACATTTGTATCTTTATCTGTATCCTTATTTAT-3'; primer sequences used for D18S51 amplification were (F) 5'-CAAACCCGACTACCAGCAAC-3' and (R) 5'-GAGCCATGTTCATGCCACTG-3 (12,13). The amplification cycling used consisted of an initial 10min 95°C denaturation followed by 45 cycles of: 95°C for 5s, 56 °C for 20s, and 65°C for 30s with fluorescence detected during the extension cycle. Following the amplification cycles, samples underwent a transition cycle consisting of 72°C for 2min, 95°C for 20s, 55°C for 20s and 56°C for 2min, after which the amplicons were melted. The melt parameters were altered to determine how to obtain the highest exportable resolution. Ultimately, a temperature range of 60-95°C using the "continuous" option (for 0.5% incremental increases in the temperature) were determined to be optimal as they gave the most product (highest fluorescence), exhibiting a plateau in the amplification curve by the end of the programmed cycling, while providing the highest resolution melt curves possible on the ABI 7500 platform. Consequently, these conditions were used for all subsequent studies on the ABI 7500.

To continue optimization of STR amplification and melt on the ABI 7500 instrument, a normalization study was performed to determine if DNA input quantity affects derivative melt curve peak heights. A total of 30 samples were quantified according to the manufacturer's protocol, but with half-volume reactions using the Investigator™ Quantiplex Kit (Qiagen) on the Rotor-Gene® Q

(Qiagen). Samples were subsequently diluted to 1ng/µl and melted with the same reaction conditions and amplification/melt parameters as described above. Data exported for comparison included fluorescence detected at each incremental melt temperature i.e. "melt cycle". LDA analysis was used to evaluate D5S818 genotype prediction accuracy as described below. Classification accuracy was compared to that obtained when template DNA inputs for amplification/melt were not normalized.

In addition to testing normalization, three samples were amplified and melted in duplicate using the conditions described above on two different days in order to examine both run-to-run (inter-run) and sample-to-sample (intra-run) reproducibility. For this study, the fluorescence of the primary melt curve peak and its corresponding melt cycle number were used for comparison of sample data within and between runs. Data sets were compared using a two-tailed students t-test (α=0.05).

Melt curve classification – LDA using R statistical software

Samples representing five to seven genotypes at both the D5S818 and D18S51 loci were amplified with melt curve detection using the ABI 7500 qPCR instrument (as described above) in order to evaluate the ability of the melt curve assay to predict genotypes. The samples were separated into "training" standards (8-10 of each genotype for each locus) and "validation" or unknown samples (≥10 of each genotype for each locus). The resulting raw melt curve data was exported using the ABI 7500 System detection software v2.0.6, and the morphological shapes of the melt curves along with the predominant peak cycle numbers and peak heights were examined.

Next, the derivative of fluorescence (dF) was calculated, derivative melt curves were generated, and the primary peaks and shoulder peaks were identified. The data from each sample at D5S818 and D18S51 were then summarized into their respective peak and shoulder peak temperatures (i.e. melt cycles) and peak heights. For D5S818 melt curves, the primary peak/shoulder peak temperatures and heights for up to three peaks were used, if observed. If only two peaks/shoulders were observed, the height at cycle 106 was used as the third data point (constant) (**Figure 1**). No sample had fewer than two observed peaks. For D18S51 melt curves, the primary peak/shoulder peak temperatures and heights for four peaks were used if observed. If only three peaks/shoulders were observed, the height at cycle 154 was used as the fourth data point (constant) (**Figure 1)**. No sample had fewer than three observed peaks. Samples evaluated as standards for each genotype were used to train the software for LDA classification using R statistical software (©The R Foundation, Vienna, Austria) (10,11,14). All other samples were used as experimental unknowns (validation samples). Initially, LDA was used to classify each unknown sample into a predicted genotype group. However, since the genotype prediction accuracy was low (<80%) prediction categories were narrowed by combining genotypes that had similar melt patterns. For this, misclassification rates for each genotype were calculated, the misclassification trends were identified, and several geno-grouping options were created for testing. The standard (training) samples were re-assigned to identified geno-groups, unknown samples reanalyzed, and LDA was used to classify each unknown sample into a predicted geno-group. Again, prediction accuracy rates were determined by calculating the percent of samples that correctly classified.

Rotor-Gene® Q – D5S818 & D18S51 Amplification & Melt Curve Detection

Samples were amplified for each STR locus (D5S818 and D18S51) separately on the Rotor-Gene® Q using the primer and amplification reactions initially established by Kuehnert (10) and as described above (for ABI 7500). Amplification and transition cycle parameters used were identical to those described above; following the transition cycle, the amplicons were melted by 0.1° incremental increases in temperature from 60° to 95° with each incremental step held for 2s. Fluorescent signal was detected throughout the melt in the high resolution melt curve detection channel.

Melt curve classification - Rotor-Gene® Q ScreenClust HRM®Software

For PCA analysis, melt curve data generated from each sample at both STR loci was separately analyzed using the Rotor-Gene® ScreenClust HRM® software (Qiagen) (15). For each run, the software package was trained utilizing the "Supervised Mode" to analyze the selected samples

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

(standards) as the training set for analysis. Each of standard samples used was placed into one of seven genotype groups (D5S818) or one of six genotype groups (D18S51) based on their known genotypes. For each genotype group, the software averages the curves of the known standard samples within that group and determines the variability of each sample in that group from the average. All additional samples analyzed within each run were then compared against each genotype group's average and was placed into a genotype group based on highest probability of belonging to that group and acceptable variability from the average (16). From the predicted clusters, confusion matrices were used to assess the software's accuracy for clustering together samples of the same genotype; this is denoted throughout as the predicted genotype accuracy of the software. From this, the percentage of misclassifications for each genotype was determined, trends were identified, and geno-grouping options were created. To subsequently evaluate accuracy of identified geno-groups, the standard (training) samples were re-assigned in the software, unknown samples reanalyzed, and the newly predicted clusters were assessed for accuracy, as indicated above.

Melt curve classification – LDA using R statistical software

Samples representing five to seven genotypes at both the D5S818 and D18S51 loci were amplified with melt curve detection using the Rotor-Gene® Q (as described above) in order to evaluate the ability of the melt curve assay to predict genotypes. The samples were separated into "training" standards and "validation" or unknown samples, as described above. The resulting raw melt curve data was exported, the derivative melt curves were generated, and the primary peaks and shoulder peaks were identified. The data from each sample at D5S818 and D18S51 were then summarized into their respective peak/shoulder peak temperatures/heights and analyzed using the LDA code in R statistical software, as described above (**Figure 1**). For D5S818 & D18S51, if only two or three peaks/shoulders (respectively) were observed in the melt curve, the height at 64.95°C was used as the additional data point (constant).

Integration of Melt Curve Assay into Quantiplex® & Testing

To evaluate the success of the melt curve assay within the qPCR-based quantification step of the forensic DNA workflow, the D5S818 and D18S51 primers and EvaGreen® dye (Biotium) were integrated into the Investigator™ Quantiplex kit (Qiagen) for the Rotor-Gene® Q (Qiagen) platform. First, alterations to reaction and amplification conditions were carefully evaluated to ensure that the resulting melt curves were consistent with what was expected based on previous studies. Ten samples from seven D5S818 genotypes and 10 samples seven D18S51 genotypes were amplified using Quantiplex® (Qiagen) reaction chemistry, but with STR primers and EvaGreen® dye (Biotium). Quantiplex® chemistry amplification with STR primers included 7.36µl of the Quantiplex® primer mix, 7.36µl of the Quantiplex® reaction mix, 0.16µl of the 100uM each of forward and reverse primer, and 0.8µl of the 5uM EvaGreen® (Biotium) intercalating dye. One microliter of template DNA was added to each reaction for a total reaction volume of 17.16µl. Each sample was amplified/melted twice, using two different amplification and melt cycling parameters. The first amplification and melt cycling parameters tested were the standard recommended Quantiplex® amplification conditions, but with a melt cycle added ("Quantiplex® amplification with melt"): 60s 95°C denaturation followed by 40 cycles of: 95°C for 1s and 60°C for 10s with an additional melt cycle 55-95°C at a 0.1°C incremental increase. The second amplification and melt cycling parameters tested ("Quantiplex® amplification with transition and melt") were the standard Quantiplex® amplification conditions, but with an added transition stage (72°C for 2min, 95°C for 20s, 55°C for 20s and 56°C for 2 min) between the amplification and melt (60-95°C at a 0.1° incremental increase). To examine melt curve reproducibility and determine if previously established prediction geno-groups could be used for the integrated assay, primary melt curve peak temperatures from each experimental group were compared to those previously obtained using optimized STR singleplex amplification and its noted inter-run variation (11).

In order to determine if the Quantiplex® amplicons themselves produce melt products when the transition and melt cycles were added to the amplification parameters an additional set of Quantiplex® standards (Qiagen) were amplified using the recommended reaction set-up (without STR primers or EvaGreen® dye (Biotium)) and the established amplification with transition and melt program. Resulting melt curves were compared to those obtained when standard reactions included the STR primers and EvaGreen® dye (Biotium).

In order to determine if alterations in Quantiplex® reaction chemistry (Qiagen) would affect resulting human DNA quantification estimates expected. QC measures and reported Quantiplex® inter-run variation were compared to those obtained when standards and samples were amplified with the Quantiplex® integrated STR melt curve assay (17). Variation in values between the integrated melt curve run and the run using recommended reaction conditions were compared to the inter-run variation reported by Qiagen using a student's t-test (α=0.05) (17). Further, sample human DNA quantification values obtained using standards whose reaction included the integrated melt curve assay were compared to values observed when standards did not include the integrated melt curve assay. Variation from expected values was calculated and compared using a one-way analysis of variance (ANOVA) test (α=0.05).

Finally, genotype and geno-group prediction accuracies of both single-source and mixture samples were determined using the integrated Quantiplex® melt curve quantitation assay. For this, 56 unknown samples and ten 2-person (1:1) mixtures of known genotypes were tested using the Quantiplex® amplification reaction with duplexed D5S818 and D18S51 primers and the added transition and melt program. The resulting melt curves from the single-source samples were analyzed using LDA in R (and standard sample data) as described above and genotype and geno-group prediction accuracies were determined. Next, in order to determine the ability of the assay to correctly identify mixtures versus single-source samples, the single-source samples were reanalyzed in R but with the addition of one added standard (training) group for prediction comparisons – a 1:1 mixture sample standard group (n=8). For genotyping and geno-grouping analysis, prediction accuracies were determined by calculating the percent of samples that correctly classified. For mixture screening, prediction accuracies were determined by noting the percent of samples that were correctly identified as single-source (regardless of predicted geno-grouping) versus the percent of mixtures that were correctly identified as a mixture. From this screening assay, mixtures were declared if either of the tested STR loci predicted as a mixture.

Development of a software tool for whole melt curve analysis

A functional data analysis approach was used to analyze the whole curve data by converting the data into a functional data space. In conjunction with machine learning techniques, such as LDA and Support Vector Machines (SVM), this approach provides a very flexible technique for describing the whole curve data (18). The whole curve data is projected into the functional data space, which represents the whole melt curve versus a prespecified subset of the feature space (specific curve characteristics, as described above). LDA, as described above, uses separating hyperplanes to discriminate between groups and SVM allows for separating curves to be used for discrimination. Further, SVM's use a localized kernel smoothing approach to classification.

For this work, the entire melt curve data from each tested sample (ie. "whole melt curve") was modeled using penalized splines to obtain a coefficient space to be used for classification. Data from both qPCR platforms (ABI 7500 qPCR platform and the Rotor-Gene® Q qPCR system) was analyzed. A cross-validation study was conducted on the whole curve data using three methods: LDA, SVM with Linear Basis functions, and SVM with Radial Basis functions (all techniques used penalized spline coefficients). Samples evaluated as standards for each genotype were used to train the software for classification using each method; all other samples were used as experimental unknowns (validation samples). Confusion matrices were generated and prediction accuracies were determined by calculating the percent of samples that correctly classified.

Results and Discussion

ABI 7500 Normalization & Reproducibility

A normalization study was conducted in order to determine if the concentration of the input DNA had any effect on the peak height ratios of the melt curves produced. Thirty samples were amplified and melted twice – once using a standard input of 2µL regardless of sample concentration (non-normalized) and again using a standardized input of 1ng across all samples (normalized). The LDA statistical approach was utilized to determine whether the genotype prediction accuracy was improved when all samples were amplified and melted using the same DNA input. Surprisingly, the prediction accuracy was best when the samples were non-normalized (**Table 1**). When the non-normalized standard samples were classified against themselves, a 100% genotype prediction accuracy rate was noted compared to only 73% when they were normalized. Although overall classification rates were lower when unknown samples were classified, non-normalized samples again had a higher genotype prediction accuracy rate than the normalized samples (40% vs. 26%). Based on these data, non-normalized DNA inputs were used for all other melt curve experiments described in this study. The finding that the amount of input DNA has no negative impact on the ability of the melt curve to accurately predict genotypes is crucial as ultimately, should HRM analysis be successfully incorporated into a quantitation assay, forensic samples would be amplified and melted without prior quantification.

The inter-run and intra-run melt curve reproducibility of the ABI 7500 platform was tested. When comparing the D5S818 melt curves produced from samples across two separate runs (inter-run), significant differences in both the primary peak fluorescence and the melt cycle numbers at which these peaks were produced were observed (**Figure 2,** p=0.0002073). However, there were no observable differences between samples run in duplicate on the same run (**Figure 2**, p=0.7848 and 0.2170 for run 1 and 2, respectively). These data show poor reproducibility across runs, but not within runs, suggesting that a melt curve assay developed for this platform (ABI 7500) may require that genotype standards be run with each plate.

Prediction Accuracies using Specific HRM Curve Characteristics & ScreenClust HRM®

Single Source, Singleplex HRM Analysis

When using the Rotor-Gene® Q ScreenClust HRM® software to predict D5S818 genotypes from Rotor-Gene® HRM data, samples were classified correctly only 23.77% of the time (**Table 2A**). When using the Rotor-Gene® Q ScreenClust HRM® software to predict D18S51 genotypes from HRM data, samples were classified correctly 40.38% of the time (**Table 2B**). In order to determine if prediction accuracies could be improved by grouping genotypes with similar melt patterns, geno-groups were generated and tested for both loci. Nine and four geno-grouping options (for D5S818 and D18S51, respectively) were initially created based on observed trends and classification rates (**Table 3**). For both loci, clustering the training and validation data into geno-groupings improved prediction accuracies. The highest rate of accuracy observed using established geno-groups were 46.6% and 65.38% (for D5S818 and D18S51, respectively). The top-performing geno-group options for both loci were assessed using the proposed alternate method (LDA) to allow for direct comparison of the two methods; in each case, the LDA method performed the approximately the same (D18S51) or better (D5S51) than the PCA-based Rotor-Gene® Q ScreenClust HRM® analysis (data not shown). While segregating the data into closely associated groups (‘geno-groups’) doubled the prediction accuracies generated, they fell short of the stated target (~85%).

Prediction Accuracies using Specific HRM Curve Characteristics & LDA

Single Source, Singleplex HRM Analysis

When using LDA in R software to predict D5S818 and D18S51 genotypes from Rotor-Gene® HRM data, samples were classified correctly 58.92% and 17.31% of the time (**Table 4**). Again, in

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

order to determine if prediction accuracies could be improved by grouping genotypes with similar melt patterns, geno-groups were generated and tested using the Rotor-Gene® HRM data for both loci. Twelve and four different geno-grouping options (for D5S818 and D18S51, respectively) were initially created (based on observed trends/classification rates) and tested, which again resulted in improvements in prediction accuracies (to 81.0% and 63.46% for D5S818 and D18S51, respectively) (**Table 4**).

In an effort to assess the performance of a more commonly-used qPCR platform, the above-mentioned studies were repeated using the same samples and LDA analysis in R software, but using the ABI 7500 qPCR platform for melt analysis. Using dissociation (ie. melt) data from ABI 7500, prediction accuracies were markedly reduced as compared to when Rotor-Gene® HRM data was used (**Table 4**). Although the ABI 7500 melt curve analysis generated a slightly higher rate of accurate genotype prediction for the D18S51 locus, this improvement was negligible.

Overall, the Rotor-Gene® Q qPCR platform out-performed the ABI 7500 for accurate classification of D5S818 and D18S51 melt curve data. This was not unexpected, as the Rotor-Gene® Q platform has a higher resolution melt curve capacity than the ABI 7500. Thus, all further efforts to integrate the STR melt curve assay into a quantification kit were focused on the Rotor-Gene® Q platform.

Mixture HRM Analysis using the Quantiplex® Integrated Assay

Initially, our goals were to test the integrated assay and ensure that the addition of STR primers and an HRM dye did not impede efficiency of the quantification or alter the subsequent melt curves produced. First, two different amplification parameters were tested with the integrated assay and the observed primary melt curve peaks were compared back to those obtained from the same samples tested using optimized singleplex STR reactions. The variation of the primary melt peak temperature observed was lowest using the Quantiplex® amplification parameters that include the transition and melt programs (0.045% variation for D5S818, 0.089% for D18S51, data not shown). Interestingly, this observed variation was even lower than what was previously noted from the inter-run variation observed when singleplex amplification and melt was performed using the original optimized reaction conditions (11). Further, examination of entire melt curves showed that samples amplified using the integrated melt curve Quantiplex® assay with the added transition cycle prior to melt produced curves which were (qualitatively) indistinguishable from those developed from singleplex STR amplification and melt using the amplification and reaction conditions that were optimized outside of the quantification kit (**Figure 3**). Additionally, no significant differences in primary melt peak temperatures were noted when the samples amplified using the integrated assay or the Quantiplex® amplification with the added transition cycle (prior to melt) were compared to the optimized singleplex STR amplification/reaction (p=0.8496 and 0.1895 for D5S818 and D18S51, respectively) (**Figure 4**). Finally, it should be noted that when the HRM dye (EvaGreen® (Biotium)) is added to the Quantiplex® amplification reaction using the transition and melt parameters (but no STR primers), an insignificant melt curve was produced outside of the range that we are observing (**Figure 5A**); the expected melt curves are only produced with the addition of the D5S818 and D18S51 primers and products produced therein include high melt peak heights in the expected STR D5S818 and D18S51 temperature ranges and minimal background noise (**Figure 5B**). This demonstrates that the Quantiplex® chemistry does not alter the melt curves produced or contribute any additional melt products to the integrated assay. Altogether, these results support the use of the added transition cycle prior to the melt program for improved reproducibility of the melt curve data using the integrated Quantiplex® melt curve assay.

In addition to studying the effects of the altered Quantiplex® reaction on the melt curves themselves, it was important to study the effects of the added reagents on the Quantiplex® kit's ability to accurately and reproducibly quantify DNA. The R^2 value obtained from the standard curves when STR primers and EvaGreen® dye were added to the standard samples remained within the

manufacturer's expected values (0.9926) (**Table 5**) (17). However, the slope was slightly lower than the expected acceptable range (-2.5459 v. expected range between -3.0 and -3.6) (**Table 5**). While this was unexpected, it is not problematic as lower standard curve slopes actually indicate that the PCR efficiency is higher than expected suggesting that there is greater than a 2-fold change for each new cycle in the amplification (17). A slight decrease in concentration variability was also noted when the concentration of standard samples that included the STR primers and EvaGreen® dye were compared to values obtained using the standard manufacturer's reaction versus the normal inter-run variation noted by our laboratory. Based on previous Quantiplex® internal validation data, an average inter-run difference of 0.5371 ng/µl and percent variation of 20.72% between duplicate standard samples is expected; however, standards amplified using the Quantiplex® reaction with STR primers/EvaGreen® dye showed a difference of only 0.7808 ng/µl (17.74%) (**Table 6**). When data from a set of unknown samples quantified using standards prepared using the manufacturer's recommended reaction were compared with values obtained when the same samples were quantified using the new Quantiplex® HRM integrated assay, no significant differences in concentrations were observed (p= 0.7685, data not shown). This indicates that modifications to the chemistry and reaction parameters introduced with the Quantiplex® HRM integrated assay do not alter the expected human DNA concentration values produced.

In order to evaluate the effectiveness of the newly-developed Quantiplex® HRM integrated assay, 56 single-source samples and 10 mixture samples were analyzed. Initially, the ability of the integrated assay to accurately assign D5S818 and D18S51 geno-groups to single-source (unknown) samples was assessed. The previously established best-performing geno-groupings were used; this included for D5S818 [Group 1: (10,11) (11,13); Group 2: (11,11) (12,12); Group 3: (11,12) (12,13) (13,13)] and for D18S51 [Group 1: (12,14) (12,15); Group 2: (12,16); Group 3: (13,14) (13,16) (14,15)], but an additional Group 4 was added to account for the additional genotype (12,13) that was analyzed in this study (11). Overall, 35 of 56 single-source samples classified correctly for D5S818 and 30 of 56 single-source samples classified correctly for D18S51 (**Table 7**). Taken altogether, these data indicate that the integration of both STR loci into the quantitation assay, along with the addition of a standard sample category for mixture classification, did not reduce the accuracy of these noted geno-group predictions. However, given the relatively low geno-group prediction accuracies reported using this assay, it is not recommended for use as an early screening tool for reporting potential exclusions from single-source forensic samples.

Ultimately, the goal of this assay was to develop a method that could be used in conjunction with qPCR-based DNA quantitation as a *mixture* screening tool. Thus, the final experiment sought to determine the ability of the newly-developed Quantiplex® HRM integrated assay to accurately predict whether an unknown sample is a single-source or mixture sample. Fortunately, when melt curve data from the integrated quantitation assay were analyzed separately for each STR locus tested, mixtures were accurately predicted for 100% of mixture samples tested at the D5S818 locus and for 80% of mixture samples tested at the D18S51 locus (**Table 8**). However, this assay was designed as a duplex assay to increase the probability of an accurate determination of single-source versus mixture status for forensic samples. Thus, the determination of a mixture would only require that one STR locus (of the two tested) be predicted as a mixture sample. With this in mind, the integrated assay demonstrated the ability to accurately identify a mixture for 100% of the mixture samples tested (**Table 9**). Conversely, single-source samples were predicted as such in 92.86% of samples tested (**Table 9**). For this data set, two single-source samples misclassified as mixtures using the D5S818 melt data and two misclassified as mixtures using the D18S51 melt data. Taking all tested samples into account (56 single-source and 10 mixture samples), the Quantiplex® HRM integrated assay was able to properly distinguish between single-source and mixture samples in 62 of 66 samples tested for an overall accuracy rate of 93.94% (**Table 9**). Based on these data, we are confident that this assay provides a viable mixture screening assay for forensic DNA samples when characteristic features of the melt curves generated are used.

Prediction Accuracies using Whole Curve Analysis using LDA & SVM

Single Source, Singleplex HRM Analysis

The entire melt curve data set from each tested sample was modeled using penalized splines to obtain a coefficient space to be used for classification (**Figure 6**). Each curve generated includes an average penalized spline, which captures the full features of the melt curves. The splines form an averaged melt curve that can is then used in the classification process. Data from both qPCR platforms (ABI 7500 qPCR platform and the Rotor-Gene® Q qPCR system) were analyzed using three methods: LDA, SVM with Linear Basis functions, and SVM with Radial Basis functions (**Table 10**). SVM with Radial Basis functions performed best for melt data generated from the Rotor-Gene® Q for D5S818 with genotype classifying accurately at rate of 74.87% when unknowns were compared to the training dataset (**Table 11**). However, for D18S51 genotype classifications using Rotor-Gene® Q data, SVM with Linear Basis functions performed best, but only correctly classified sample genotypes 26.92% of the time. This lowered rate of correct classification for the D18S51 locus is not unexpected and reflects the same trends noted in the data detailed above. Also as expected and reflected in the data above, the ABI7500 was outperformed by the Rotor-Gene® Q for genotype classification. For melt data generated on the ABI7500 qPCR platform, SVM with linear basis functions performed best for the D5S818 data (50.26% genotype accuracy) (**Table 11**) while SVM with radial basis functions performed better for the D18S51 data (21.16% genotype accuracy). It should be noted that all SVM methods had 100% correct classification rates to the training set (comparing to self).

Mixture HRM Analysis using the Quantiplex® Integrated Assay

In order to evaluate the effectiveness of the newly-developed Quantiplex® HRM integrated assay, Rotor-Gene® Q whole curve data from 56 single-source samples and 10 mixture samples were analyzed. In this study, for each STR locus tested, only the best statistical classification method identified for whole curve analysis (as determined above). Initially, the ability of the integrated assay to accurately assign D5S818 and D18S51 genotypes for single-source samples was assessed. Overall, 32 of 56 single-source samples (57.14%) classified correctly for D5S818 and 24 of 56 single-source samples (42.86%) classified correctly for D18S51 (**Table 12**). While these genotype prediction accuracies are equivalent to or better than what was observed when only characteristic features of the curve were analyzed, they are slightly lower than the geno-grouping accuracies observed when only characteristic features of the curve were analyzed. In either case, neither approach produced prediction accuracies that would justify the use of this assay for the early identification of exclusionary results from single-source forensic samples.

As the ultimate goal of this work was to develop an assay that could accurately identify mixture versus single-source samples, the final experiment sought to determine if the use of whole curve data would improve upon the ability of the newly-developed Quantiplex® HRM integrated assay to accurately predict whether an unknown sample is a single-source or mixture sample. Using the SVM with Radial Basis functions to analyze the D5S818 data, 100% of the mixtures tested were accurately predicted as such; likewise, using the SVM with Linear Basis functions to analyze the D18S51 data, 100% of mixture samples tested were accurately predicted as such (**Table 12**). Conversely, single-source samples were predicted as such in 92.8% of samples tested (**Table 13**). For this data set, three single-source samples misclassified as mixtures using the D5S818 melt data and four misclassified as mixtures using the D18S51 melt data; however, this assumes that only one locus be required to be called as a "mixture" for the sample classification to be mixture (as noted in the studies detailed above). The use of whole curve data notably improved the prediction capacity of mixtures for D18S51 (to 100%), as with D5S818, suggesting that this assumption may need to be reevaluated in future studies. Taking all tested samples into account (56 single-source and 10 mixture samples), the Quantiplex® HRM integrated assay was able to properly distinguish between single-source and mixture samples in 59 of 66 samples tested for an overall accuracy rate of 89.39% (**Table 13**). Based

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

on these data, we are confident that this assay also provides a viable mixture screening assay for forensic DNA samples when melt data from the entire sample curve is used.

Conclusions & Future Implications

Overall, this work provides a qPCR-based integrated HRM + quantitation assay that can provide an analyst with indication of a mixed forensic sample early in the forensic DNA workflow. Having this information can arm the analyst with more information earlier, as well as providing them with greater confidence in combining sample DNA extracts for subsequent STR amplification and analysis. The availability of a screening assay for early mixture detection would help alleviate many of the challenges and stresses associated with mixture and/or low level touch DNA testing faced in forensic crime. Currently, forensic DNA units do not have a way to accurately determine if a mixture is present prior to the CE analysis stage. If more information were available early on in the analytical workflow, protocols could be adjusted to allow for combination of single-source, low-level surface swabs prior to STR amplification - allowing for easier detection of low level contributors. Ultimately, this assay could result, most importantly, in a time savings for forensic DNA labs by reducing the manual time needed for low level DNA sample retesting and minimizing sample consumption concerns for low level samples.

The solution proposed in this work was to develop an STR-based melt curve assay that could be incorporated into the qPCR quantification step in the forensic DNA workflow. Prior to evaluating the ability of STR melt curves to predict sample characteristics, it was important to determine if sample DNA input would affect the melt curves produced, since samples would ultimately be amplified and melted prior to availability of known quantitation values. Data from these studies showed that normalizing the amount of input DNA did not change any attributes of the produced melt curves and thus, advanced knowledge of sample quantitation would not be needed. This was a crucial early finding. Unfortunately, inter-run variation was found to be higher than expected suggesting the need for amplification/melt of known standards with every run. However, the reproducibility study was conducted on the ABI 7500 platform only, which does not have high resolution melt capacity. With a higher resolution instrument, run-to-run variation would likely be minimized. Future studies would need to re-assess the reproducibility of this assay on the Rotor-Gene® Q, the qPCR platform ultimately used for the Quantiplex® HRM integrated assay.

The first major goal of this study was to evaluate the ability of two qPCR platforms to produce melt curves that could be used to accurately predict STR genotypes or geno-groups of unknown samples using *only* key characteristic features of the curves. Initially, two statistical methods for classification were used. While the methods were similar in their abilities to accurately predict genotypes or geno-groups from D18S51 melt curve data, the LDA method (in R Software) substantially outperformed the PCA-based method (ScreenClust HRM® software) using D5S818 melt curve data. As expected, data from the higher resolution instrument (Rotor-Gene® Q) did consistently provide higher genotype and geno-group prediction accuracies than data from the ABI 7500 platform. Further, geno-grouping standard (training) samples (whose genotypes were morphologically similar) resulted in improved prediction accuracies regardless of locus tested or qPCR instrument used; geno-group predictions for D5S818 exceeded 81% when melt curve data from the Rotor-Gene® Q was analyzed using LDA. However, neither genotype prediction accuracies nor geno-group prediction accuracies using key characteristic curve data were high enough to reach the pre-determined desired rate of 85%. Based on all of these findings, further attempts to integrate this STR melt curve assay into an existing quantification kit were completed on the Rotor-Gene® Q platform using the HRM channel and Qiagen's Investigator™ Quantiplex quantification kit along with an LDA-based analysis. Data presented herein shows that integration of the D5S818 and D18S51 STR primers and EvaGreen® dye into the Quantiplex kit chemistry along with the addition of a back-end transition cycle and melt

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

process altered neither the expected melt curves generated, the quality of the quantitation standard curves, nor the expected concentration values of unknown samples. As predicted, genotype prediction accuracies of unknowns remained lower than the target 85%. More importantly, however, our data reveals that this newly designed Quantiplex® HRM integrated assay is able to accurately distinguish between single-source and mixture samples ~94% of the time when only limited key characteristic features of the melt curve data are used.

The second major goal of this work was to determine if use of the entire melt curve data set could improve the ability of the assay to accurately predict genotypes and to identify mixtures (versus single-source samples). Unfortunately, traditional LDA analysis is limited in its ability to evaluate large, complex amounts of data (such as the entire exported melt curve data set). Consequently, it was necessary to evaluate the whole curve data using machine learning techniques (incorporating both LDA and SVM) that employ different algorithms and have the ability to model data and classify large datasets based on previously trained observations. As with the previous analysis of limited curve data, we found that the use of the higher resolution Rotor-Gene® Q platform allowed for more accurate predictions of genotype when the entire melt curve dataset was used versus use of the entire melt curve data from the ABI 7500 (regardless of classification method). Additionally, our data show that the use of SVM-based algorithms consistently achieved higher classification rates than an LDA-based method when the entire melt curve datasets were used, regardless of STR locus or qPCR instrument. However, the SVM algorithm that provided the highest genotype prediction accuracies differed between the two STR loci; the whole melt curve assay was most successful in predicting D5S818 genotypes when the SVM with Radial Basis functions was used while the SVM with Linear Basis functions worked best for analysis of D18S51 data. When these top performing methods were used to analyze whole curve data from the integrated Quantiplex® HRM assay, genotyping predictions again fell below the targeted 85% accuracy, however, the ability of the assay to identify single source versus mixture samples reached 100% for *each* individual locus tested, suggesting that single source classification may need only one locus to classify as mixture when whole curve data and SVM analysis is used (versus both as determined by the studies using limited curve data and LDA).

Overall, this work has successfully produced a qPCR-based melt curve assay for the prescreening identification of mixtures and the assay has been demonstrated to be viable when integrated into a commercial quantification assay. Implementation of this assay into a forensic DNA laboratory will provide the analyst with more information about their evidentiary samples without the need for *any* additional steps in the workflow. However, there are several considerations that must be addressed prior to crime lab implementation. First, the reproducibility of the integrated melt curve assay must be thoroughly evaluated on all potential qPCR platforms. Further, as only 2-person, 1:1 mixtures from reference DNA samples were used for testing herein, it will be necessary to further demonstrate performance across a spectrum of mixture ratios, using mixtures with a greater number of contributors (ie. >2-person mixtures), and using DNA from more compromised, forensically-relevant samples. Given the tremendous success of this assay for identification of single-source versus mixture samples, it may also be fruitful to assess the final integrated assay on the ABI 7500, a qPCR platform which is more commonly used in forensic laboratories. Lastly, it should be noted, that training and validation data sets were used for only single cross validations and that SVM typically requires larger data sets for training and validation than what was able to be generated in this study. Also, SVM data is best utilized for the storage of large reference sample (standards) data sets for long-term application and use. Consequently, future studies would also be needed for the generation of larger D5S818/D18S51 datasets as well as for examination of a 10-fold cross validation SVM approach. In addition, the use of the machine learning technique of Artificial Neural Networks could be explored for whole curve classification as well. Ultimately, we aim to incorporate the best methods into a single web-based tool to facilitate broader access of the melt curve SVM database and to provide an easy-to-use, free, on-line tool for quick assessment of melt curve data. The reference database could be

packaged in R and posted to the Comprehensive R Archive network (CRAN) or a free-downloadable plug-in could be made available through Bitbucket or Github. A user interface could be designed to allow for simple upload of the melt curve file with user selection of the loci tested and run criteria (qPCR platform, kit, etc.). An easily exportable report could provide single-source or mixture identification for each sample along with presumptive genotyping data for the tested loci for identified single-source samples.

References

1. Kimpton CP, Gill P, Walton A, Urquhart A, Millican ES, Adams M. Automated DNA profiling employing multiplex amplification of short tandem repeat loci. PCR Methods Appl. 1993 Aug;3(1):13-22.
2. Hochmeister MN, Budowle B, Jung J, Borer UV, Comey CT, Dirnhofer R. PCR-based typing of DNA extracted from cigarette butts. Int J Legal Med. 1991;104(4):229-33.
3. Aditya S, Sharma AK, Bhattacharyya CN, Chaudhuri K. Generating STR profile from "Touch DNA". J Forensic Leg Med 2011;18:295-298.
4. Gill P. Application of low copy number DNA profiling. Croat Med J. 2001 Jun;42(3):229-32.
5. Butler JM. Forensic DNA Typing. Elsevier Academic Press: New York, 2005.
6. Bureau of Alcohol, Tobacco, Firearms, and Explosives. "National Response Team". US Department of Justice. Accessed September 10, 2012.
7. Balding DJ, Buckleton J. Interpreting low template DNA profiles. Forensic Sci Int Genet 2009;4:1-10.
8. Nguyen Q, McKinney J, Johnson DJ, Roberts KA, Hardy WR. STR melting curve analysis as a genetic screening tool for crime scene samples. J Forensic Sci. 2012 Jul;57(4):887-99.
9. Qiagen. QIAamp DNA Blood Mini-Buccal Swab Protocol version 3. April 2010.
10. Kuehnert K. Development of a Genotype Prediction Screening Assay Using High Resolution Melting Curve Analysis of the STR Loci D5S818 and D18S51. MS thesis. Virginia Commonwealth University, 2015.
11. Cloudy, D. Discrimination of D5S818 and D18S51 STR genotypes using LDA & PCA evaluation of high resolution melt curve data from the Qiagen Rotor-Gene® Q qPCR Instrument. MS thesis. Virginia Commonwealth University, 2017.
12. Nicklas JA, Noreault-Conti T, Buel E. Development of a fast, simple profiling method for sample screening using high resolution melting (HRM) of STRS. J Forensic Sci. 2012 Mar;57(2):478-89.
13. Kline MC, Hill CR, Decker AE, Butler, JM. STR sequence analysis for characterizing normal, variant, and null alleles. Forensic Sci Int Genet. 2011;5(4):329-32.
14. R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
15. Qiagen. Rotor-Gene® Screen Clust HRM® Software: User Guide. Sept 2009.
16. Qiagen. Rotor-Gene® Q User Manual Version 2. Feb 2012.
17. Qiagen. Investigator Quantiplex Kit Validation Report. July 2011.
18. Lee Y, Lin Y, Wahba, G. Multicategory Support Vector Machines. J American Stat Assoc. 2004; 99(465): 67.

Tables & Figures:

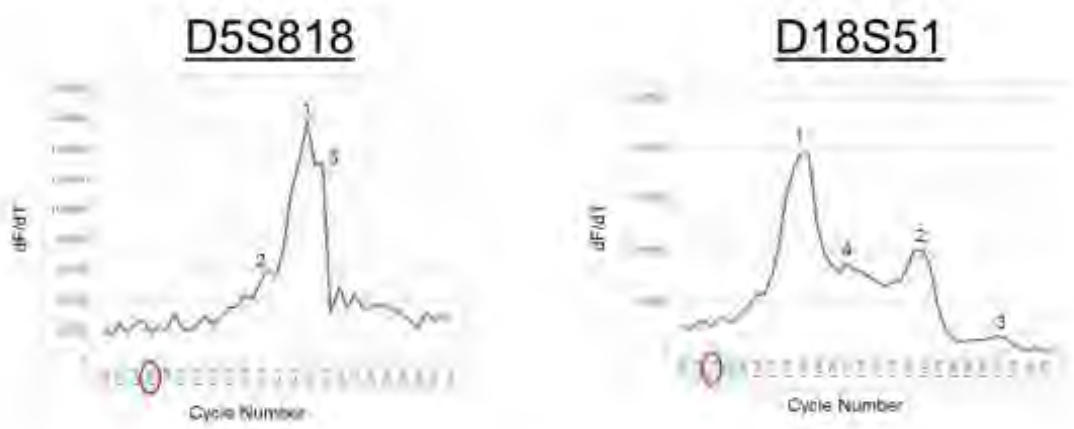


Figure 1. Melt curve classification using key curve characteristics. Initially, only key curve features from each locus (numbered) were used for melt curve analysis for LDA classification. For D5S818, the height at cycle number 106 from the ABI 7500 data (or 64.95°C for Rotor-Gene®Q) was used as the third data point if a third peak was not observed. For D18S51, the height at cycle number 154 from the ABI 7500 data (or 64.95°C for Rotor-Gene®Q) was used as the fourth data point if a fourth peak was not observed.

Table 1. Normalization study classification accuracy for of D5S818 genotypes using ABI 7500 melt curve data using and LDA.

	Standards* v. Self	**Unknowns^ v. Standards**
Normalized	73.0%	26.0%
Non-Normalized	100.0%	40.0%

*n=19 samples

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

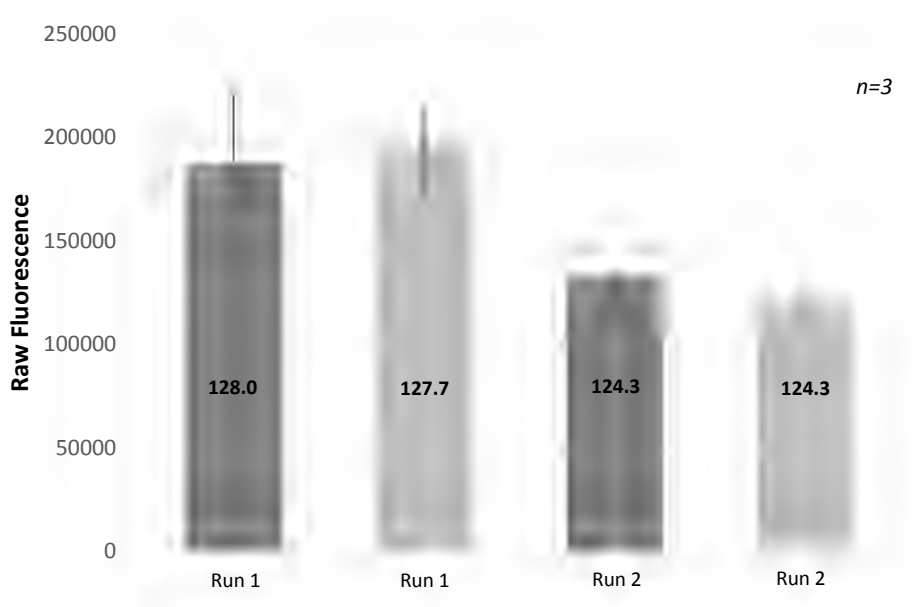


Figure 4. The inter-run and intra-run fluorescence variation of the primary D5S818 peak from ABI 7500. Numbers inside bars indicate the average melt cycle number of the primary D5S818 peak. Differences noted between duplicate sample values within run are not significant (p=0.7848 for run 1 and p=*0*.2170 for run 2). However, differences between runs were (p=0.00020703).

Table 2. Classification of Genotypes using HRM data and Principal Component Analysis (PCA) based Rotor-Gene®Q ScreenClust HRM® software A) for the D5S818 STR and B) for the D18S51 STR locus.

A)

D5S818 Predicted Genotypes (PCA)								
Known Genotypes n= 206		(10,11)	(11,11)	(11,12)	(11,13)	(12,12)	(12,13)	(13,13)
	(10,11)	**5**	1	1	3	1	0	2
	(11,11)	4	**14**	3	1	7	1	2
	(11,12)	7	0	**10**	9	14	10	17
	(11,13)	3	9	3	**3**	2	6	8
	(12,12)	1	7	0	1	**7**	4	1
	(12,13)	1	1	7	1	5	**5**	7
	(13,13)	0	1	0	0	4	2	**5**
Accuracy Rate: 23.77%								

B)

D18S51 Predicted Genotypes (PCA)							
Known Genotypes		(12,14)	(12,15)	(12,16)	(13,14)	(13,16)	(14,15)
	(12,14)	**7**	3	2	5	1	2
	(12,15)	6	**6**	4	2	2	0

	(12,16)	2	2	**8**	1	1	0
	(13,14)	4	1	1	**7**	6	1
	(13,16)	0	1	0	1	**7**	5
	(14,15)	1	1	0	1	6	**7**
Accuracy Rate: 40.38%							

Table 3. Geno-groupings created to evaluate HRM analysis of the D5S818 & D18S51 using Rotor-Gene® Q ScreenClust HRM® software. **only top four performing groupings shown*

A) D5S818 Geno-groupings*

Option 1 Groupings	Group 1	(10,11), (11,13)	
	Group 2	(11,11), (12,12)	Accuracy Rate: 46.6%
	Group 3	(11,12), (12,13), (13,13)	
Option 9 Groupings	Group 1	(11,11), (13,13)	
	Group 2	(11,12), (12,12), (12,13)	Accuracy Rate: 46.1%
	Group 3	(10,11), (11,13)	
Option 2 Groupings	Group 1	(10,11), (11,12)	
	Group 2	(11,11), (12,12)	Accuracy Rate: 45.6%
	Group 3	(11,13), (12,13), (13,13)	
Option 3 Groupings	Group 1	(10,11), (11,11)	
	Group 2	(11,12), (12,12)	Accuracy Rate: 42.2%
	Group 3	(11,13), (12,13), (13,13)	
Option 6 Groupings	Group 1	(10,11), (12,13), (13,13)	
	Group 2	(11,11), (11,13)	Accuracy Rate: 41.7%
	Group 3	(11,12), (12,12)	

B) D18S51 Geno-groupings*

Option 1 Groupings	Group 1	(12,14), (12,15), (13,14)	
	Group 2	(12,16)	Accuracy: 65.38%
	Group 3	(13,16), (14,15)	
Option 2 Groupings	Group 1	(12,14), (13,14)	
	Group 2	(12,15), (12,16)	Accuracy: 63.46%
	Group 3	(13,16), (14,15)	
Option 4 Groupings	Group 1	(12,14), (12,15)	
	Group 2	(12,16)	Accuracy: 49.04%
	Group 3	(13,14), (13,16), (14,15)	
Option 3 Groupings	Group 1	(12,14), (12,15)	
	Group 2	(12,16)	Accuracy: 42.31%
	Group 3	(13,14), (13,16)	

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

Table 4. Classification accuracy of D5S818 & D18S51 genotypes and geno-groupings using ABI 7500 and Rotor-Gene® Q melt curve data and LDA.

	D5S818 Classification Accuracy (%)		D18S51 Classification Accuracy (%)	
	Genotype	Geno-group	Genotype	Geno-group
ABI 7500	23.94% n=188	65.40% n=188	18.33% n=60	55.00% n=60
Rotor-Gene® Q*	58.92% n=185	81.00% n=185	17.31% n=52	63.46% n=52

*Data previously reported (30)

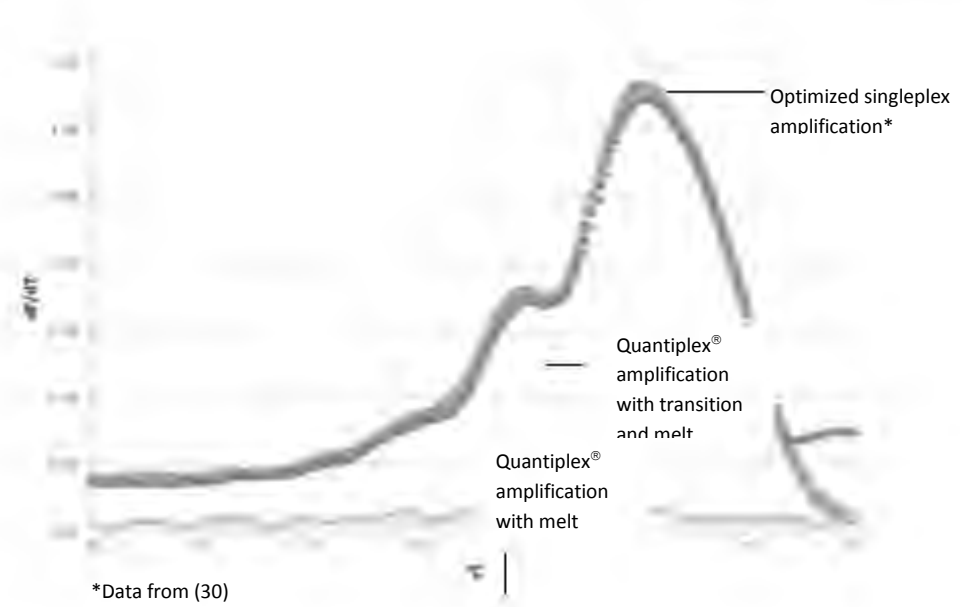


Figure 3. D5S818 melt curve for a single sample using three different amplification/melt parameters on the Rotor-Gene Q®. dF/dT represents change in fluorescence level (positive or negative) with respect to per unit change (increase) in temperature. The optimized singleplex amplification and Quantiplex® amplification with transition and melt were very similar in both fluorescence, primary peak temperature, and overall patterns. However, the Quantiplex® amplification with melt had substantially less fluorescence and was not distinguishable from background noise.

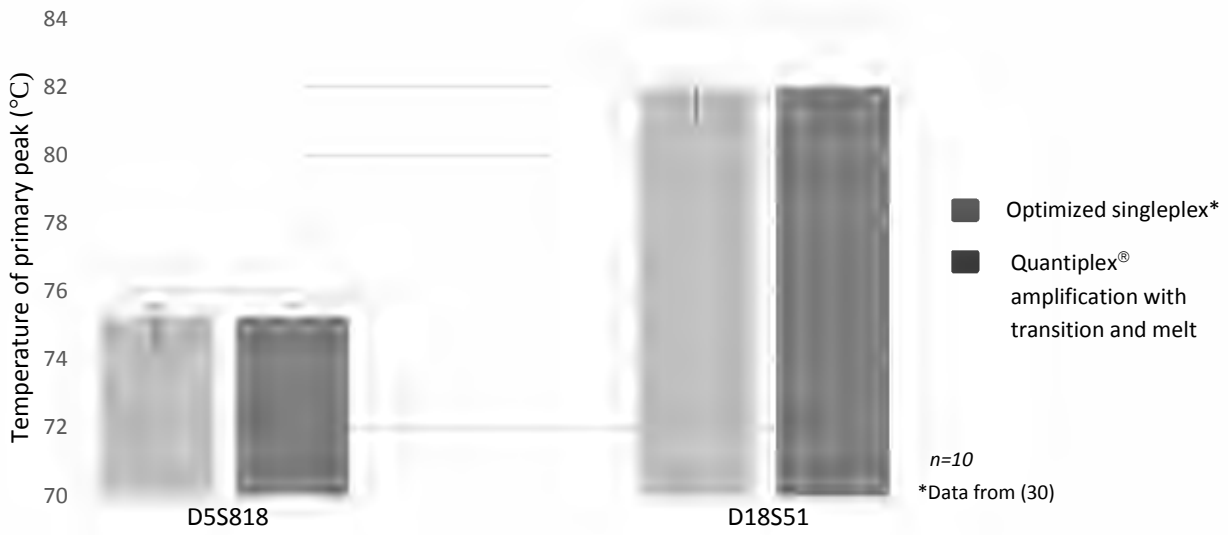
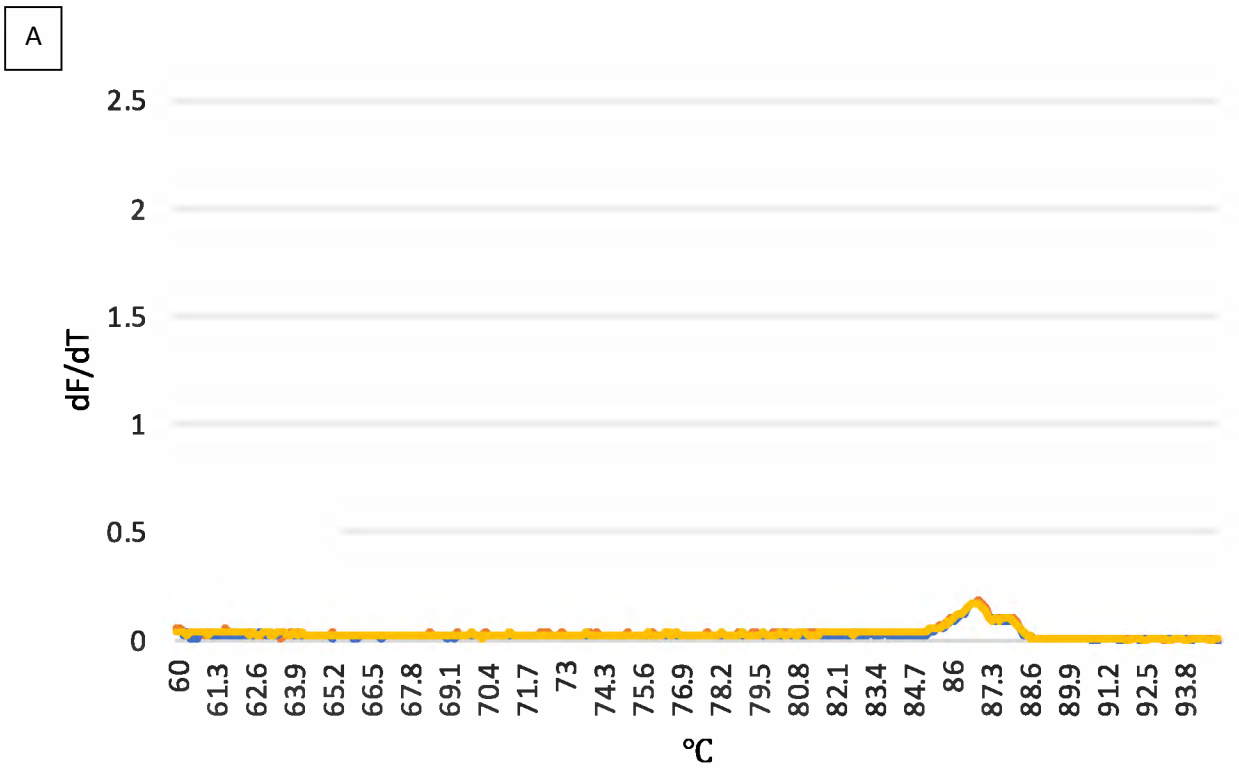


Figure 4. Primary peak temperature comparison of the Quantiplex® amplification with transition and melt to the optimized singleplex reaction. For the D5S818 locus there is no significant difference between the two amplification and melt parameters (p=0.8496). Similarly, for D18D51 locus there is no significant difference between the two amplification and melt parameters (p=0.1895).



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

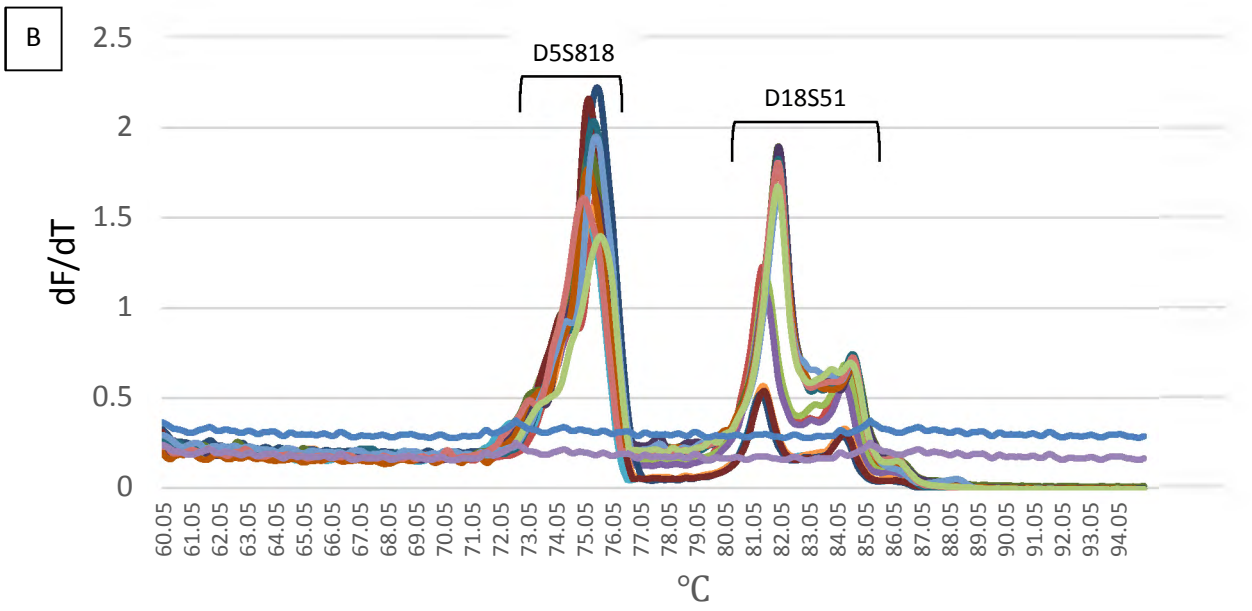


Figure 5. Standard melt curves using Quantiplex® amplification with transition and melt parameters. Quantiplex® standard samples (with EvaGreen® dye and without STR primers) normally produce an insignificant melt curve (A). However, the addition of EvaGreen® dye and STR primers to Quantiplex® standard samples allows for high melt curve peak heights and minimal background noise in the expected temperature range (B).

Table 5. Standard curve QC measures from Quantiplex® runs on the Rotor-Gene® Q.

	Quantiplex® reaction and amplification*	Quantiplex® reaction with STR primers & amplification with transition & melt
Slope	-3.0 – -3.6	-2.5459
R^2 value	>0.9900	0.9926

*Values reported by Qiagen (35)

Table 6. Human DNA quantification inter-run concentration variation of standard samples using Quantiplex® on the Rotor-Gene® Q. n=16

	Quantiplex® reaction and amplification*	Quantiplex® reaction with STR primers & amplification with transition & melt◊
Average Difference between runs	0.5371 ng/µl	0.7808 ng/µl
% Variation between runs	20.72%	17.74%

*based on multiple runs in Dawson-Cruz laboratory ◊ as compared to standard Quantiplex® and amplification run

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

Table 7. Classification accuracy of the Quantiplex®HRM Integrated assay for both STR loci tested using limited characteristic curve features, the Rotor-Gene® Q, & LDA.

	D5S818 Actual				
		Geno-group 1	Geno-group 2	Geno-group 3	Mixture
Predicted	Geno-group 1	**6**	0	2	0
	Geno-group 2	1	**23**	9	0
	Geno-group 3	4	3	**6**	0
	Mixture	0	1	1	**10**
Overall Accuracy Rate: **68.18%**					

n=66

	D18S51 Actual					
		Geno-group 1	Geno-group 2	Geno-group 3	Geno-group 4	Mixture
Predicted	Geno-group 1	**8**	3	1	1	0
	Geno-group 2	0	**0**	1	0	0
	Geno-group 3	9	3	**20**	0	0
	Geno-group 4	0	0	4	**5**	2
	Mixture	0	0	0	2	**8**
Overall Accuracy Rate: **62.12%**						

n=66

Table 8. Mixture prediction accuracy of the Quantiplex®HRM Integrated assay for both STR loci tested using limited characteristic curve features, the Rotor-Gene® Q, & LDA

Mixture Prediction			
	D5S818	**D18S51**	**Overall**
Mixture 1	Y	Y	Y
Mixture 2	Y	Y	Y
Mixture 3	Y	Y	Y
Mixture 4	Y	Y	Y
Mixture 5	Y	Y	Y
Mixture 6	Y	Y	Y
Mixture 7	Y	Y	Y
Mixture 8	Y	N	Y
Mixture 9	Y	Y	Y
Mixture 10	Y	N	Y
Accuracy			**100%**

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

Table 9. Single-source v. Mixture prediction accuracies of the Quantiplex®HRM Integrated assay for both STR loci tested using limited characteristic curve features, the Rotor-Gene® Q, & LDA

	D5S818 (%)	**D18S51** (%)	**Combined accuracy** (%)
Single-source n=56	96.43	96.43	92.86
Mixtures n=10	100.0	80.00	100.0
		Overall Accuracy (%):	**93.94**

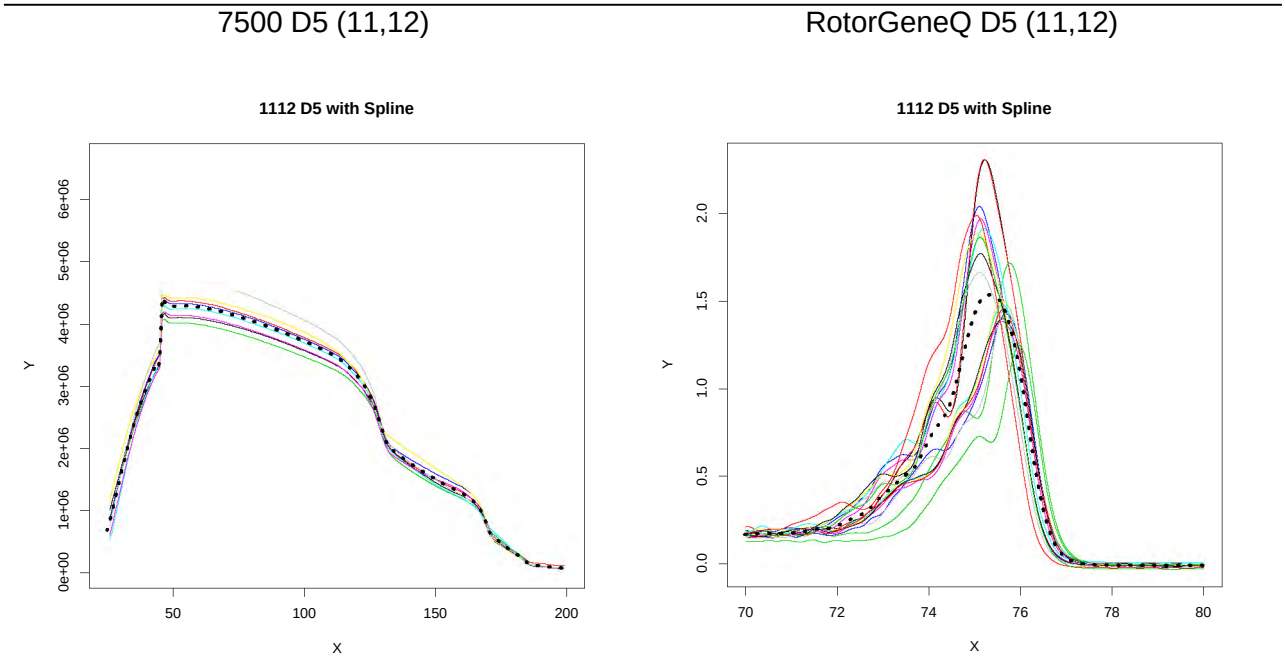


Figure 6: Whole melt curves for D5S818 (11,12) using both ABI 7500 and Rotor-Gene® Q data. The dotted line is a fitted spline used for classification methods.

Table 10: Classification accuracies for whole curve analysis for both ABI 7500 and Rotor-Gene® Q for three machine learning classification techniques: Linear Discriminant Analysis (LDA), Support Vector Machine with Linear Basis functions (SVM-Linear) and Support Vector Machine with Radial Basis functions (SVM-Radial). The best method for each locus tested is denoted in bold.

	ABI7500		**RotorGeneQ**	
Technique	D5	D18	D5	D18
LDA	43.39%	9.52%	66.31%	13.46%
SVM-Linear	**50.26%**	16.93%	69.52%	**26.92%**
SVM-Radial	40.21%	**21.16%**	**74.87%**	9.62%

Table 11: Best performing classification accuracies for both STR loci tested using whole melt curve data and the ABI 7500 (SVM-Linear, 50.26% accuracy) and the Rotor-Gene® Q (SVM-Radial, 74.87% accuracy). Items in bold are correctly classified.

ABI7500

	Predicted Genotype						
	10,11	11,11	11,12	11,13	12,12	12,13	13,13
10,11	**5**	0	0	2	0	0	3
11,11	0	**10**	9	0	5	2	3
11,12	2	9	**44**	4	2	4	13
11,13	1	0	1	**16**	6	3	1
12,12	0	0	4	1	**8**	1	4
12,13	0	0	0	5	1	**10**	6
13,13	0	0	0	2	0	0	**2**

RotorGeneQ

	Predicted Genotype						
	10,11	11,11	11,12	11,13	12,12	12,13	13,13
10,11	**9**	0	3	1	0	0	0
11,11	0	**33**	3	0	2	0	1
11,12	0	0	**34**	1	9	0	4
11,13	0	0	4	**29**	0	0	0
12,12	1	1	1	0	**11**	1	0
12,13	0	2	3	1	8	**16**	0
13,13	0	0	0	0	1	0	**8**

Table 12. Classification accuracy of the Quantiplex®HRM Integrated assay for both STR loci tested using whole melt curve data, the Rotor-Gene® Q, & the best SVM classification technique.

D5S818 Predicted Genotypes									
Actual		10,11	11,11	11,12	11,13	12,12	12,13	13,13	Mixture
	10,11	**3**	0	0	1	0	0	0	0
	11,11	0	**3**	1	0	1	1	0	0
	11,12	0	2	**19**	1	0	0	1	0
	11,13	0	0	1	**4**	0	0	0	1
	12,12	0	0	2	1	**0**	0	0	1
	12,13	0	0	3	2	0	**2**	0	1
	13,13	0	0	4	0	0	0	**1**	0
	Mixture	0	0	0	0	0	0	0	**10**
Overall Accuracy Rate: **63.64%**									

D18S51 Predicted Genotypes									
Actual		12,13	12,14	12,15	12,16	13,14	13,16	14,15	Mixture
	12,13	**7**	1	0	0	0	0	0	0
	12,14	3	**3**	2	0	0	0	0	0
	12,15	1	0	**6**	1	0	0	1	0
	12,16	1	0	2	**1**	0	1	1	0
	13,14	5	0	0	0	**1**	2	4	0
	13,16	0	0	0	0	0	**2**	3	2
	14,15	0	0	0	0	0	0	**4**	2
	Mixture	0	0	0	0	0	0	0	**10**
Overall Accuracy Rate: **51.50%**									

Table 13. Single-source v. Mixture prediction accuracies of the Quantiplex®HRM Integrated assay for both STR loci tested using whole melt curve data, the Rotor-Gene® Q, & the best SVM classification technique

	D5S818 (%)	**D18S51** (%)	**Combined accuracy** (%)
Single-source n=56	94.64	92.86	87.50
Mixtures n=10	100.0	100.00	100.0
Overall Accuracy (%):			**89.39**