ORIGINAL ARTICLE

Validation of the Applied Biosystems RapidHIT ID instrument and ACE GlobalFiler Express sample cartridge

Jennifer Churchill Cihlar[1,2] · Kapema Bupe Kapema[1] · Bruce Budowle[1,2]

Abstract

Rapid DNA platforms are fully automated systems capable of processing DNA from biological samples and interpreting the results in approximately 90 minutes with minimal human intervention. With a greater reliance on the system than on the analyst, validation data are especially needed to define the performance and limitations of commercially available Rapid DNA systems. Thus, validation studies of a Rapid DNA workflow consisting of the Applied Biosystems RapidHIT ID Instrument and RapidLINK software with a focus on the ACE GlobalFiler Express Sample Cartridge and reference buccal swabs were performed in accordance with Scientific Working Group on DNA Analysis Methods Validation Guidelines. These validation studies included assessments of sensitivity, contamination, concordance, reproducibility and repeatability, stability, inhibition, mixtures, sample reprocessing, precision, and first-pass success rate. Overall, the current Applied Biosystems RapidHIT ID Instrument with the ACE GlobalFiler Express sample cartridge was found to be a reliable tool for generation of STR profiles from reference-type buccal swabs.

Keywords Rapid DNA · STR · Validation · SWGDAM · ACE cartridge

Introduction

Short tandem repeat (STR) typing is routinely used to associate or exclude individuals as possible contributors of biological evidence [1–6]. Forensic laboratories currently use capillary electrophoresis (CE)-based technologies and workflows to generate STR profiles from biological evidence. However, forensic scientists have continually worked towards reducing the amount of time and labor needed to generate these STR profiles (e.g., updates and developments in automation, direct amplification, polymerase chain reaction (PCR) protocols, etc.) [7–11] as the demand on forensic resources, including trained laboratory personnel, continues to increase while reduced turnaround times are sought. Rapid DNA Systems offer the forensic community a fully automated system capable of processing DNA from

a biological sample and interpreting the results in approximately 90 min with minimal human intervention.

Rapid DNA Systems can be placed in both laboratory and non-laboratory environments that each play a role in investigations, prosecutions, and the subsequent care of survivors, such as booking stations [12–15], customs and border patrol stations [16], the site of mass disasters [17, 18], or hospitals for tracking biopsy samples. The benefit of Rapid DNA Systems, with placement in non-traditional laboratory settings, is the potential to decrease the turn-around time from sample collection through sample processing. However, operation of Rapid DNA Systems in non-traditional laboratory settings relies on operators with widely varying backgrounds and not necessarily having the training and education of scientists. Because of the greater reliance on the system in such settings, data are needed to define the performance and limitations of commercially available Rapid DNA systems.

Rapid DNA Systems collectively include a Rapid DNA instrument, the PCR STR typing kit/Rapid DNA cartridge, and an expert system software all used to develop and interpret an STR profile [19]. Two primary Rapid DNA analysis platforms currently exist and are being evaluated and employed by forensic laboratories and police agencies [13–15, 18, 20–34]. One is the ANDE Rapid DNA

✉ Jennifer Churchill Cihlar
Jennifer.Cihlar@unthsc.edu

1 Center for Human Identification, University of North Texas Health Science Center, Fort Worth, TX, USA

2 Department of Microbiology, Immunology, and Genetics, University of North Texas Health Science Center, Fort Worth, TX, USA

Table 1 Description of each validation study performed on the RapidHIT ID instrument during this study

Study	Number of Samples in Study	Swab Type	Sample Cartridge	SWGDAM Validation Guideline Addressed
Concordance Study	51	Puritan	ACE	4.3
	10	Puritan	ACE Positive Control Cartridge	4.3
Contamination Study	10	Puritan	ACE	4.5
	8	Puritan	ACE Negative Control Cartridge	4.5
Sensitivity (Swipe) Study	20	Puritan	ACE	4.2
	20	Copan	ACE	4.2
Repeatability and Reproducibility Study	27	Puritan	ACE	4.3.1 & 4.3.2
Swab Reprocessing Study	24	Puritan	ACE	4.3.1 & 4.3.2
Stability Study	44	Puritan	ACE	4.1
Inhibition Study	18	Puritan	ACE	4.1
Mixture Study	21	Puritan	ACE	4.4

Identification System (ANDE Corporation, Longmont, CO, USA), which was formerly known as the NetBio DNAS-can Rapid DNA Analysis System [35–41]. The other is the Applied Biosystems RapidHIT ID System (Thermo Fisher Scientific, Waltham, MA, USA), which was formerly a product of IntegenX (Pleasanton, CA, USA) [42–46]. Previous changes/updates to the RapidHIT ID System instrumentation are the use of different PCR STR typing kits [22–24] and a single sample analysis design [14, 22]. In the past few years, the RapidHIT ID instrument has been modified again compared with the previous versions. These modifications included configuration of the primary cartridge and other hardware, such as the heated optics and thermal cycling (personal communication), modifications to thermal cycling parameters and available software tools (personal communication), introduction of an additional (RapidINTEL) sample cartridge [26, 33, 47], and changes to the thresholds used when making genotype calls [48]. Thus, with these material modifications, the reliability and efficacy of the current RapidHIT ID System need to be evaluated.

Validation studies of a Rapid DNA workflow consisting of the Applied Biosystems RapidHIT ID Instrument (Thermo Fisher Scientific) and RapidLINK software (Thermo Fisher Scientific) with a focus on the ACE GlobalFiler Express Sample Cartridge (Thermo Fisher Scientific) and reference buccal swabs were performed in accordance with the Scientific Working Group on DNA Analysis Methods (SWGDAM) Validation Guidelines [49]. This Rapid DNA workflow was evaluated for concordance with traditionally used methodologies, sensitivity, contamination, reproducibility and repeatability, swab stability, effect of common inhibitors, mixture detection, swab re-analysis, and precision. Performance metrics such as genotyping success, first-pass success rate, peak heights, and peak height ratios were collected from each of these studies, collated, and are discussed in detail herein. Each of these studies provides

data that support the reliability and efficacy of the Applied Biosystems RapidHIT ID Instrument, ACE GlobalFiler Express Sample Cartridge, and RapidLINK software workflow for generation of STR profiles from reference-type buccal swabs.

Methods and materials

Buccal swab collection

Samples were anonymized and collected in accordance with methods approved by the Institutional Review Board for the University of North Texas Health Science Center in Fort Worth, TX, USA. Specific collection parameters (e.g., swab type and number of swipes) varied between studies and are discussed in greater detail below. The Puritan 3″ Sterile Standard Cotton Swab with Semi-Flexible Polystyrene Handle (C/N: 25–803 2PC; Puritan, Guilford, ME) and 4N6FLOQSwabs Genetics (C/N: 4,479,433; Copan, Brescia, Italy) were the swabs used in this study. For the purpose of this study, a "swipe" was deemed to be one motion from the top of the cheek to the bottom of the cheek. Due to constraints during the COVID-19 pandemic, all swabs were self-collected during the course of this study.

Concordance data from 3500 genetic analyzer workflow

DNA was extracted from the buccal swabs using the Prep-Filer Forensic DNA extraction kit (Thermo Fisher Scientific) as per manufacturer's guidelines [50]. The extracted DNA was quantified by qPCR using the Quantifiler Trio DNA Quantification kit (Thermo Fisher Scientific) [51] on a 7500 real-time PCR system (Thermo Fisher Scientific) following manufacturer's guidelines. Extracted DNAs
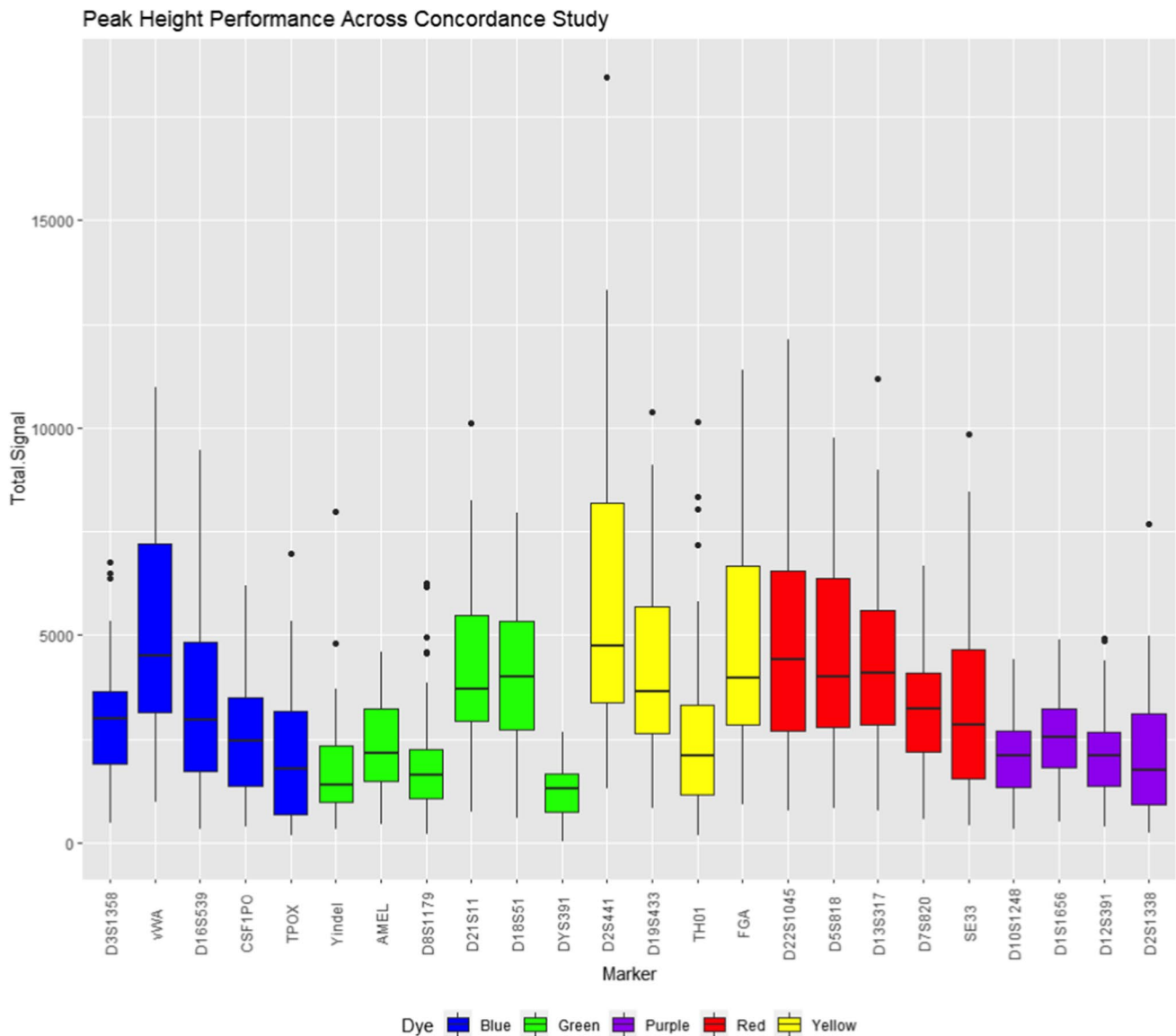
Fig. 1 Boxplots illustrating Total Signal observed for each locus from the 61 samples typed by the RapidHIT ID System in the Concordance Study. The horizontal bar within the box represents the median for the associated data. The "whiskers" or vertical bars on the box- plot represent 1.5 * interquartile range. The individually drawn points (black dots) represent points that extend beyond 1.5*interquartile range

were amplified with the GlobalFiler Express PCR Ampli- fication Kit (Thermo Fisher Scientific) [52] on a Verti® Thermal Cycler (Thermo Fisher Scientific). The PCR con- tained 6μL each of the GlobalFiler Express Master Mix and GlobalFiler® Express Primer set and 3μL of sample lysate (or Molecular Biology Grade H_2O for negative control reactions) in a final volume of 15μL. A total of one nano- gram was used for input DNA amounts. Thermal cycling parameters were: enzyme activation at 95 °C for 1 minute; 26 cycles of denaturation at 94 °C for 3 seconds and anneal- ing/extension at 60 °C for 30 seconds; followed by a final extension step at 60 °C for 8 minutes and final hold at 4 °C.

The amplified products were subjected to electrophoresis on a 3500xL Genetic Analyzer (Thermo Fisher Scientific) according to the GlobalFiler Express PCR Amplification Kit User Guide [52]. Raw data were analyzed using GeneMap- per™ *ID-X* Software v1.6 (Thermo Fisher Scientific), with allele peak threshold calls set at 175 relative fluorescence units (RFU).

Validation data from RapidHIT ID system

A total of 253 samples (including positive and negative con- trols) were processed in a series of studies designed to evaluate
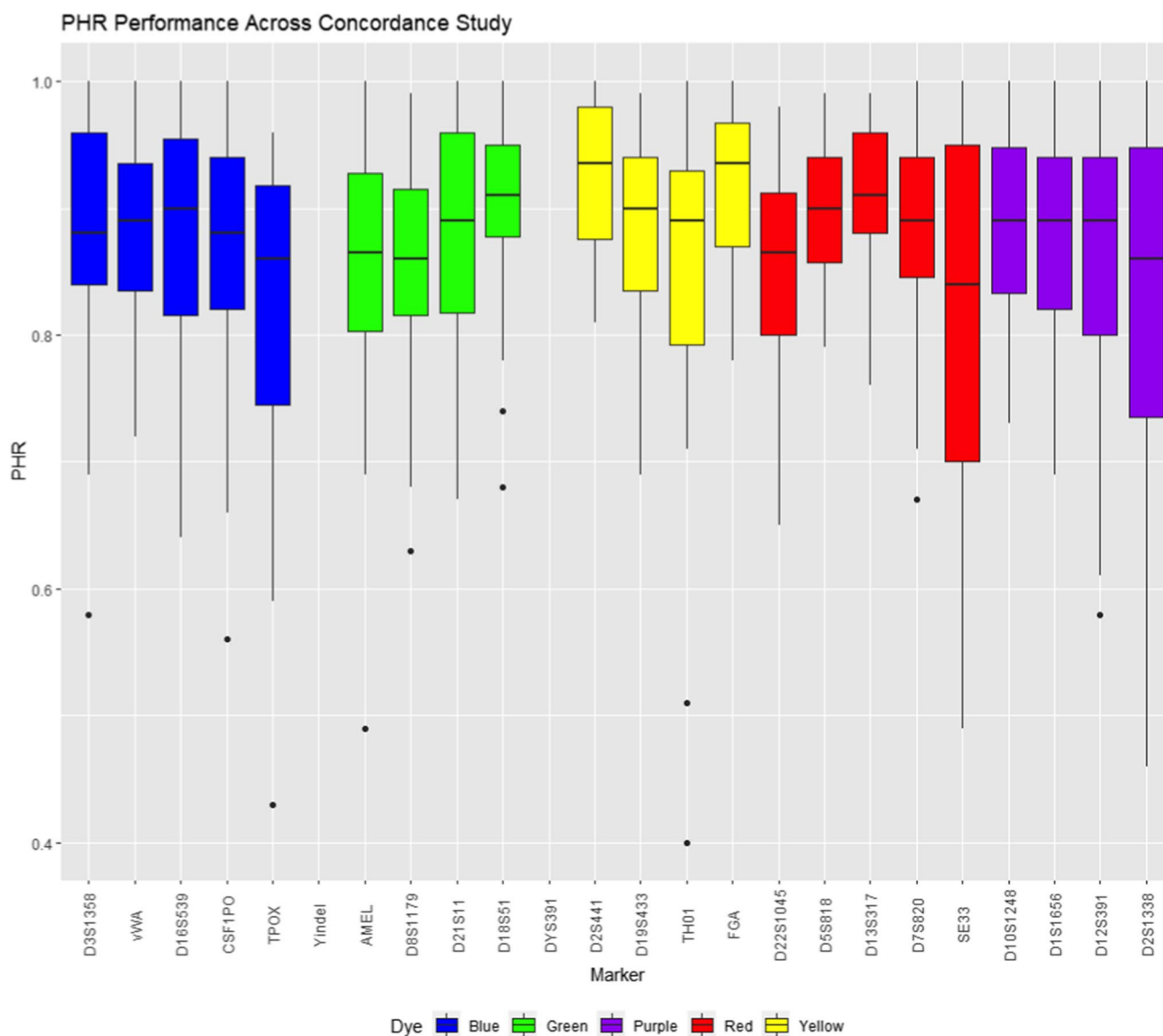
Fig. 2 Boxplots illustrating PHRs observed for each locus from the 61 samples typed by the RapidHIT ID System in the Concordance Study. The horizontal bar within the box represents the median for the associated data. The "whiskers" or vertical bars on the boxplot represent 1.5 * interquartile range. The individually drawn points (black dots) represent points that extend beyond 1.5*interquartile range

the performance of the Applied Biosystems RapidHIT ID Instrument (RapidHIT ID Software v1.3.1; Expert System v1.0) with the ACE GlobalFiler Express Sample Cartridge (C/N: A41831) and RapidLINK software v1.1.5 workflow (Table 1). All manufacturer recommendations were followed when operating the RapidHIT ID instrument [53]. Due to the length of time dedicated to this study (including a brief shutdown during the COVID-19 pandemic), multiple sample cartridge and primary cartridge lots were used. Any lot-specific observations encountered in the study are included in the Results section. At the completion of each Rapid run, 1) the cartridge was removed; 2) the swab was allowed to dry; 3) the

cartridge and associated swab were stored at room temperature until the completion of the study. This process allowed for re-runs or re-analyses, if needed, of the original swab.

Concordance study

For the Concordance Study, 102 Puritan 3'' Sterile Standard Cotton swabs were self-collected from 51 individuals. Two swabs were self-collected at a time (one from each side of the mouth) using six swipes each. One swab from each individual was used for the traditional CE workflow on the 3500 Genetic Analyzer, and one swab was run on the RapidHIT

Table 2 Details on discordant allele calls from profiles generated by the RapidHIT ID System workflow

Sample	Notes
C3	D19S433 stutter peak called and labeled with IMB quality flag
C5	Allele called at DYS391 due to pull-up artifact in female sample. Allele did not receive a quality flag
C8	Three pull-up artifacts were called and labeled with quality flags (two labeled OB and one labeled OL). Additionally, a stutter peak at D2S441 was called and labeled with an OB and IMB quality flag
C25	SE33 stutter peak called and labeled with an IMB and PL quality flag
C26	SE33 allele labeled with OB quality flag
C27	SE33 allele labeled with OB quality flag
C31	D19S433 stutter peak called and labeled with IMB and PL quality flag
C37	D2S441 stutter peak called and labeled with OB and IMB quality flag
C50	D2S441 stutter peak called and labeled with OB and IMB quality flag
C59	D2S441 stutter peak called and labeled with OB and IMB quality flag. D19S433 stutter peak called and labeled with OB and IMB quality flag

ID System to allow for an assessment of concordance. Swabs were stored at ambient conditions prior to loading into an ACE GlobalFiler Express Sample Cartridge. Additionally, ten ACE Positive Control Cartridges were run throughout the course of the study.

Contamination study

For the Contamination Study, eight ACE Negative Control Cartridges were run throughout the course of the study. Additionally, ten unused Puritan 3" Sterile Standard Cotton swabs were placed inside ACE GlobalFiler Express Sample Cartridges and run interspersed every five runs during the Concordance Study.

Sensitivity study

For the Sensitivity Study, ten Puritan 3" Sterile Standard Cotton swabs and ten 4N6FLOQSwabs Genetics were self-collected from two individuals. Two swabs were self-collected at a time (one from each side of the mouth) with 3 hours between collections completed on the same day. Each swab collection had a variable number of swipes to generate a dilution series relevant to reference swab collection protocols. This "dilution series" ranged from two to ten swipes in two swipe intervals and was generated for both swab types. Swabs were stored at ambient conditions prior to loading into ACE GlobalFiler Express Sample Cartridges.

Reproducibility and repeatability study

For the Reproducibility and Repeatability Study, 27 Puritan 3" Sterile Standard Cotton swabs were self-collected by one individual for a study design divided into two categories. For Category One, a series of three swabs were collected on three different days: Swab 1) ten swipes on the left side of the mouth; Swab 2) ten swipes on the right side of the mouth; Swab 3) five swipes on the left side of the mouth and five swipes on the right side of the mouth. Each of these swabs was processed on the RapidHIT ID instrument the same day as collection. This nine-swab experiment was repeated once a new primary cartridge was installed on the same instrument. For Category Two, three swabs were collected (a minimum of 3 hours apart) on three different days (Day 1: ten swipes on the left side of the mouth, Day 2: ten swipes on the right side of the mouth; Day 3: five swipes on the left side of the mouth and five swipes on the right side of the mouth) where each of the three swabs was run on a different RapidHIT ID instrument. Swabs were stored at ambient conditions prior to loading into ACE GlobalFiler Express Sample Cartridges.

Swab reprocessing study

For the Swab Reprocessing Study, experiments were placed in four different categories: 1) immediate reprocessing on the RapidHIT ID instrument, 2) store overnight before reprocessing on the RapidHIT ID instrument, 3) store overnight before reprocessing on the 3500 Genetic Analyzer instrument, 4) store for at least two months before reprocessing on the 3500 Genetic Analyzer. Three individuals self-collected four swabs on four different days, where Puritan 3" Sterile Standard Cotton swabs were used to collect one swab per day by the individual using ten swipes each. Swabs were stored at ambient conditions prior to loading into ACE GlobalFiler Express Sample Cartridges or being processed by the 3500 Genetic Analyzer workflow. For Category 1, a swab from each of the three individuals was processed with the RapidHIT ID System, immediately removed upon completion of the run, loaded into a new sample cartridge, and reprocessed on the RapidHIT ID System two additional
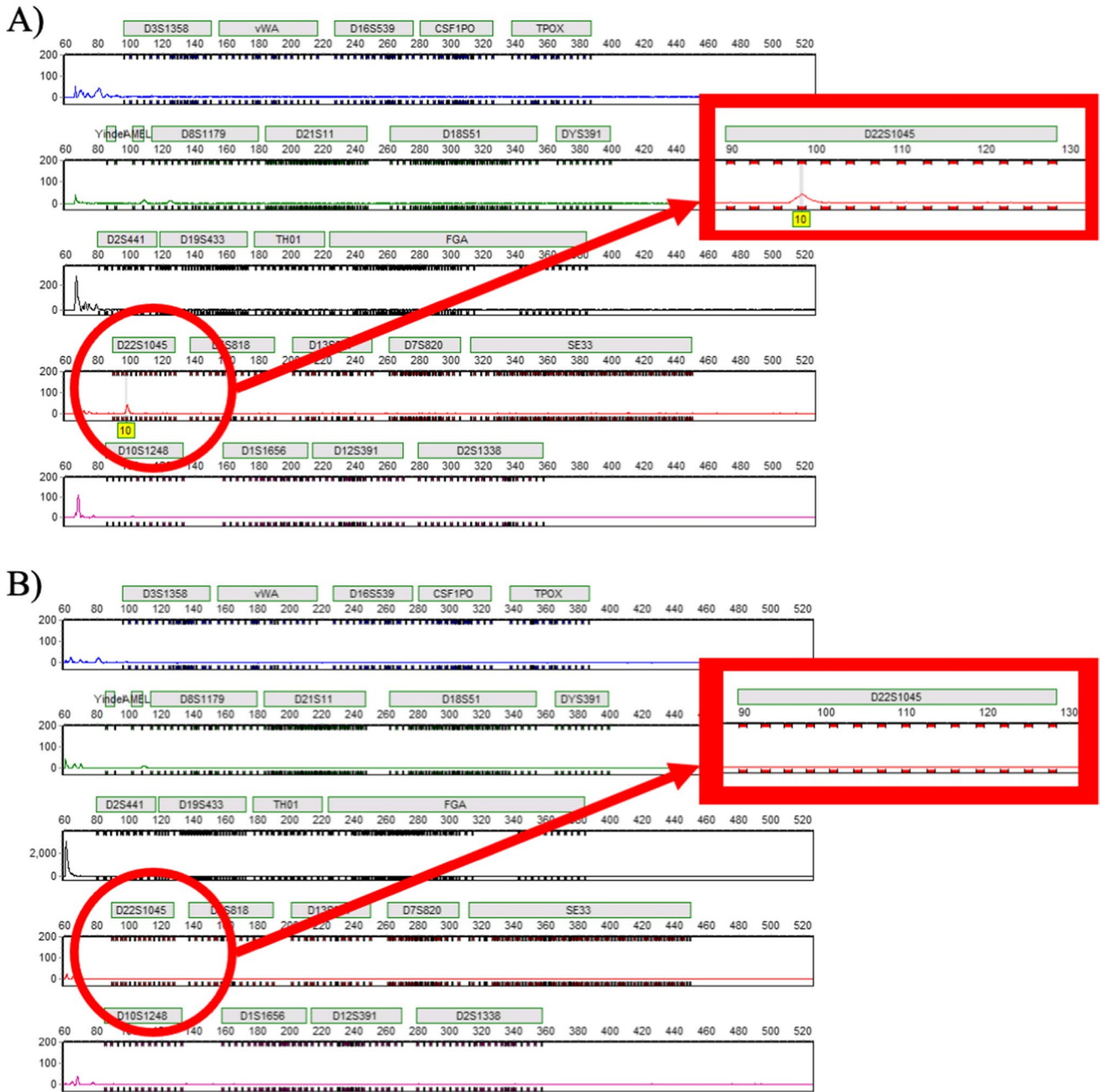
Fig. 3 Electropherograms from one of the Contamination Study samples where one peak was called and labeled with a quality flag (**A**) but was not present when the same swab was rerun with a new lot of ACE GlobalFiler Express Cartridges (**B**)

times. For Category 2, a swab from each of the three individuals was processed with the RapidHIT ID System, removed from the instrument upon completion of the run, allowed to dry at ambient conditions overnight, loaded into a new sample cartridge, and reprocessed on the RapidHIT ID System two additional times. For Category 3, a swab from each of the three individuals was processed with the RapidHIT ID System, removed from the instrument upon completion

of the run, and allowed to dry at ambient conditions overnight before processing with the 3500 Genetic Analyzer workflow. For Category 4, a swab from each of the three individuals was processed with the RapidHIT ID System, removed from the instrument upon completion of the run, stored under ambient conditions for at least two months, and then processed with the 3500 Genetic Analyzer workflow.
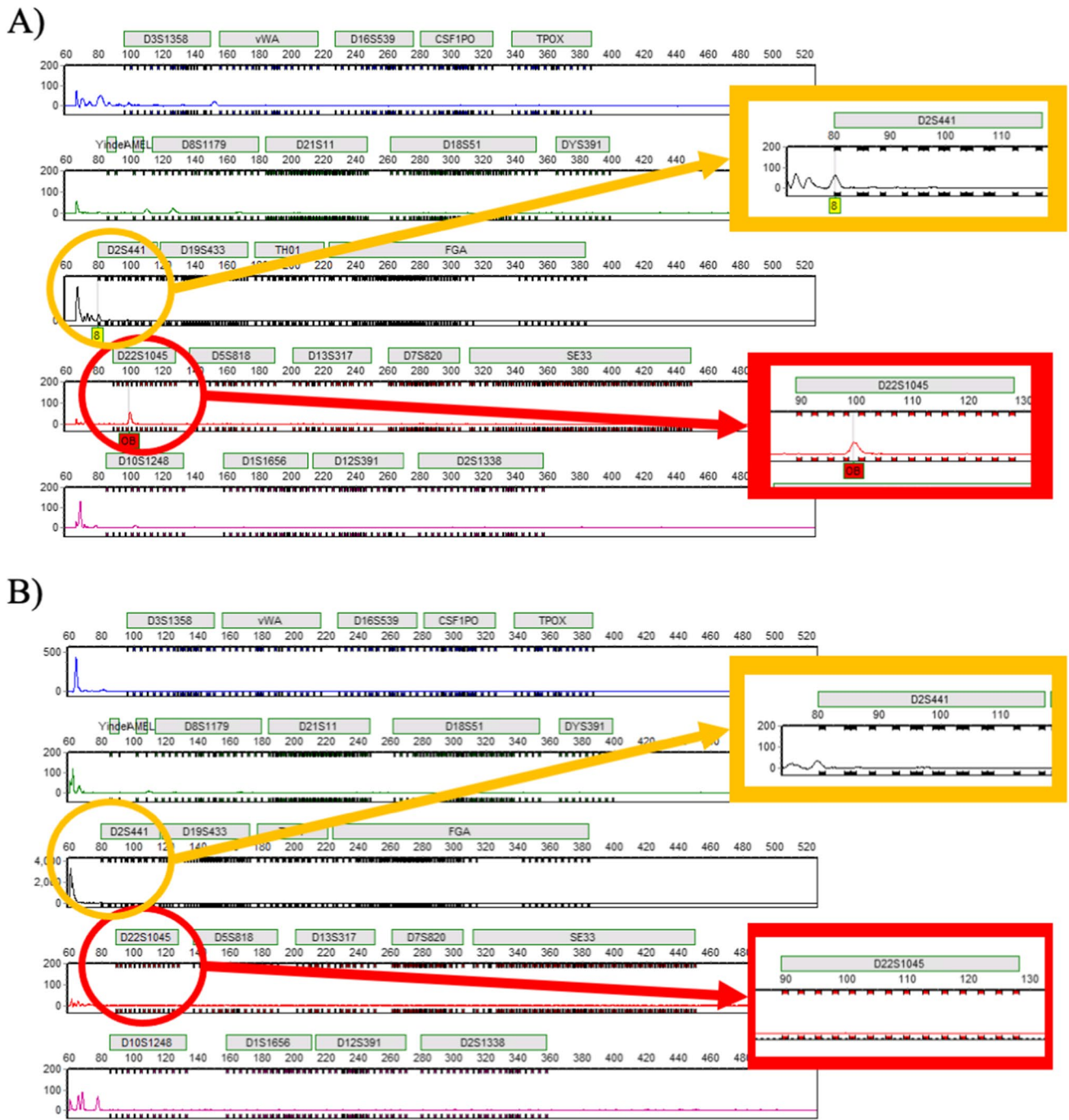
Fig. 4 Electropherograms from two ACE Negative Control Cartridges with differing lot numbers that were run back-to-back. The first electropherogram (**A**) exhibited the same peak seen in the Contamination Study samples with the same lot number and a second called peak, likely the result of unused primer sequences. The second electropherogram (**B**) had zero called peaks with the RapidHIT ID System

Stability study

For the Stability Study, 14 Puritan 3" Sterile Standard Cotton swabs (two swabs at a time using a different side of the mouth for each swab at six swipes each) were self-collected by two individuals. If more than one set of swabs was self-collected on the same day, a minimum of 3 hours was allowed to elapse between collections. The swabs were maintained under ambient conditions
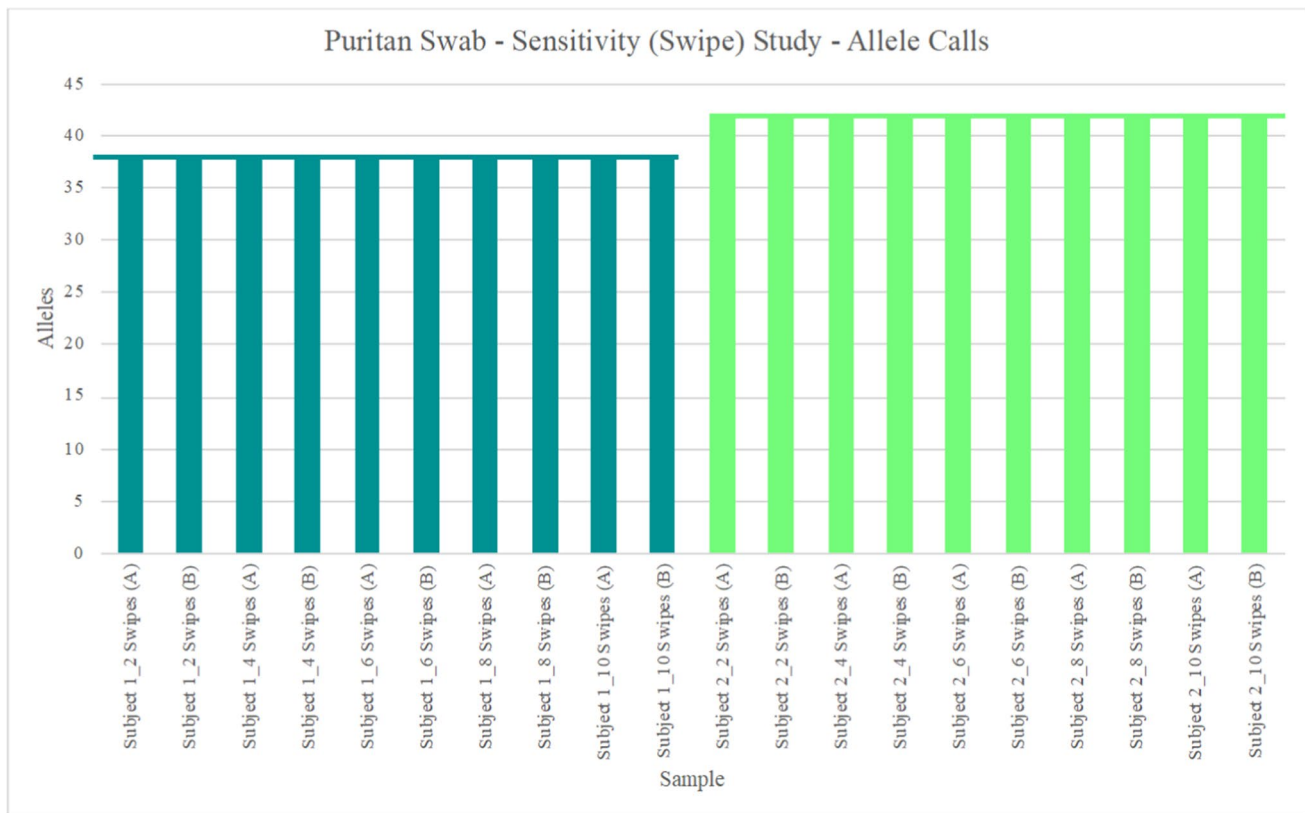
Fig. 5 The number of alleles called after primary analysis with the RapidHIT ID System for the Puritan sensitivity series compared to the total number of possible alleles for each swab (horizontal line)

and analyzed by loading into ACE GlobalFiler Express Sample Cartridges after reaching zero to 12 weeks storage time in two week intervals. The zero week time point was used as the baseline data.

Inhibition study

For the Inhibition Study, two swabs were self-collected by three individuals per inhibitor, which included drinking a cup of black coffee, gargling with Cool Mint Listerine mouthwash, and smoking a cigarette, immediately after consuming the coffee or mouthwash or smoking the cigarette. Puritan 3" Sterile Standard Cotton swabs were used to collect one swab from each side of the individual's mouth using six swipes each. Prior to drinking, gargling, or smoking the inhibitor, two clean, baseline swabs were provided by the same individuals for the Concordance Study. Swabs were stored at ambient conditions prior to loading into ACE GlobalFiler Express Sample Cartridges.

Mixture study

For the Mixture Study, saliva was self-collected from two individuals and used to generate the single-source and mixed swabs in triplicate. Saliva from the two contributors was mixed at ratios of 1:1, 1:4, 4:1, 1:9, and 9:1. A total of 50 µL of saliva was applied to each (single-source and mixed) Puritan 3" Sterile Standard Cotton swab. Swabs were stored at ambient conditions prior to loading into ACE GlobalFiler Express Sample Cartridges.

Precision study

Data generated in each of the studies described above were used to assess ladder selection and precision.

Data analysis

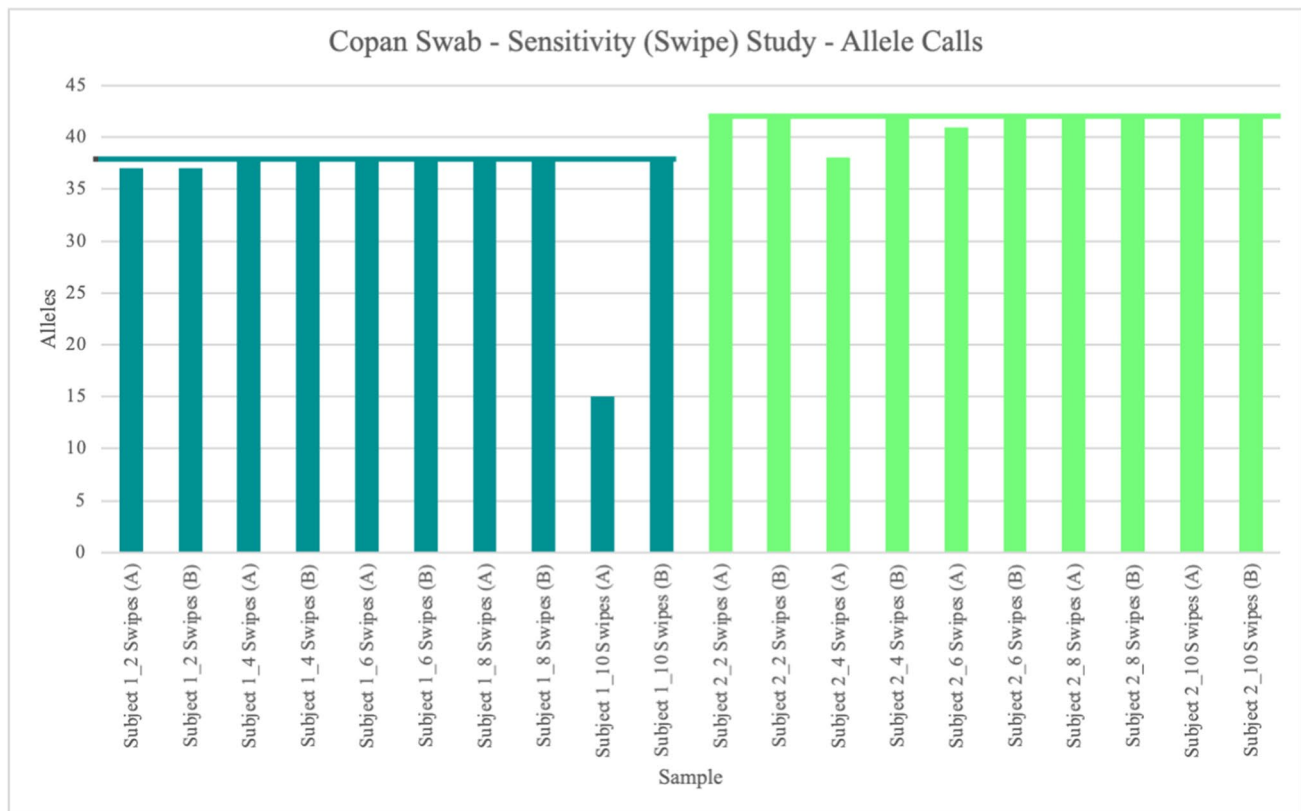The RapidLINK Software v1.1.5 was used throughout the study to manage the RapidHIT ID runs, instrumentation,

Fig. 6 The number of alleles called after primary analysis with the RapidHIT ID System for the Copan sensitivity series compared to the total number of possible alleles for each swab (horizontal line)

and consumables [54]. Performance metrics, including concordance, first-pass success rate, peak heights, and peak height ratios (PHRs; height of smaller peak/height of larger peak = PHR), were collected and collated from each of the validation studies. Allele calls, peak heights, PHRs, and the presence of any quality flags (abbreviations and descriptions provided in Supplementary Table 1) were pulled from the peak table output file after primary analysis by the GeneMarker HID STR Human Identity Software installed with the RapidLINK Software (HIDAutoLite v2.9.5; SoftGenetics, State College, PA, USA). "Total Signal" was calculated by summing the peak heights of the alleles from each locus. Thresholds and quality flag parameters used to make allele calls were developed and defined by Thermo Fisher Scientific [48]. Current CE workflows use a multi-capillary system (e.g., 3500 Genetic Analyzer) making it possible to run the samples and an allelic ladder simultaneously and subsequently reducing run-to-run migration variation for sample genotyping. The RapidHIT ID instrument employs a single-capillary and, thus, cannot run

an allelic ladder alongside the sample being analyzed. Therefore, a library of allelic ladders, which were run under technically separate migration conditions, was employed to size and identify alleles [53]. Each sample was genotyped using each ladder in the library and the local ladder allowing ladders to be ranked for selection based on the sum of differences between the sample allele peaks and candidate ladder allele peaks. The size difference between the sample allele peaks and selected ladder allele peaks was obtained from the peak table output file and evaluated in the Precision Study. First pass success was defined as the RapidHIT ID System's ability to generate a complete STR profile with no quality flags from the analysis software (noted by a "green check"). Any swab that did not generate a "green check" after primary analysis underwent secondary manual analysis (or Modified Rapid DNA Analysis) by a trained/qualified scientist. Data handling and statistical analyses were performed with Excel (Microsoft Corporation, Redmond, WA, USA) and R v3.3.2 and RStudio v1.0.136 (R Core Team, 2016).
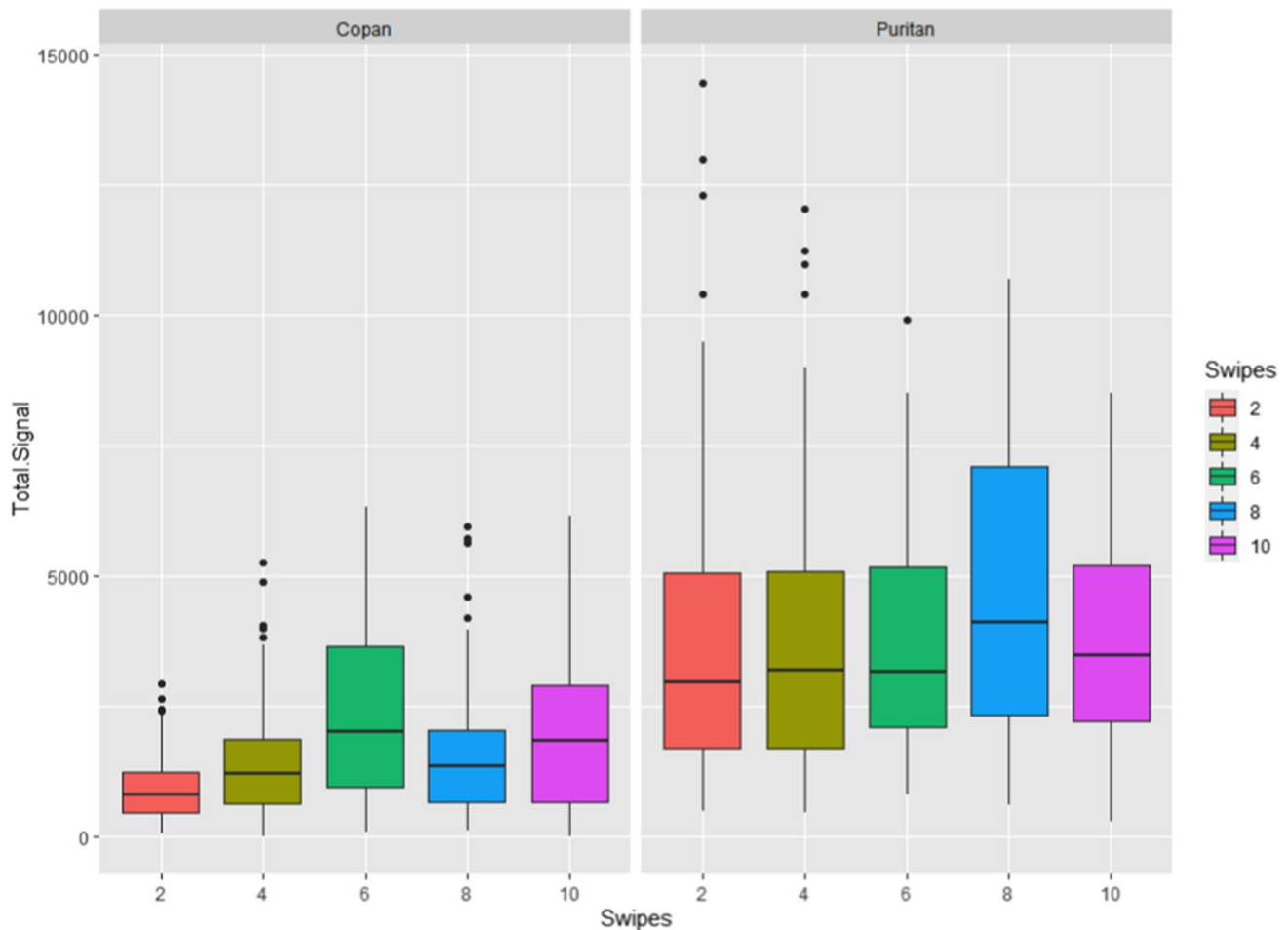
Fig. 7 Boxplots illustrating Total Signal observed for each swipe number (grouped by swab type) from the 40 swabs typed by the RapidHIT ID System in the Sensitivity Study. The horizontal bar within the box represents the median for the associated data. The "whiskers" or vertical bars on the boxplot represent 1.5 * interquartile range. The individually drawn points (black dots) represent points that extend beyond 1.5*interquartile range

Results and discussion

Results are presented from a set of validation studies, performed in accordance with SWGDAM Validation Guidelines for Forensic DNA Analysis Methods [49], to assess the Applied Biosystems RapidHIT ID Instrument, ACE GlobalFiler Express Sample Cartridge, and RapidLINK software workflow's performance. This set of validation studies included assessments of concordance with traditionally used methodologies, sensitivity, contamination, reproducibility and repeatability, swab stability, the effect of common inhibitors, mixture detection, swab re-analysis, and precision. An overall assessment of the workflow's efficacy and reliability for reference buccal swabs is described below.

Concordance study

The purpose of the Concordance Study was to compare the genotype calls made by the RapidHIT ID System to the genotype calls generated by a traditional CE workflow (i.e., 3500 Genetic Analyzer workflow). A total of 61 samples were included in this study, which included 51 Concordance Study samples and ten positive control cartridges (Table 1). Data from these 61 samples were also used to evaluate locus-specific performance for the Total Signal and PHR performance metrics.

Locus-specific performance metrics for Total Signal and PHR are illustrated in Figs. 1 and 2. The two haploid loci (DYS391 and Yindel) had an average Total Signal of 1,232 ($\pm$668) RFUs and 1,982 ($\pm$1,648) RFUs,
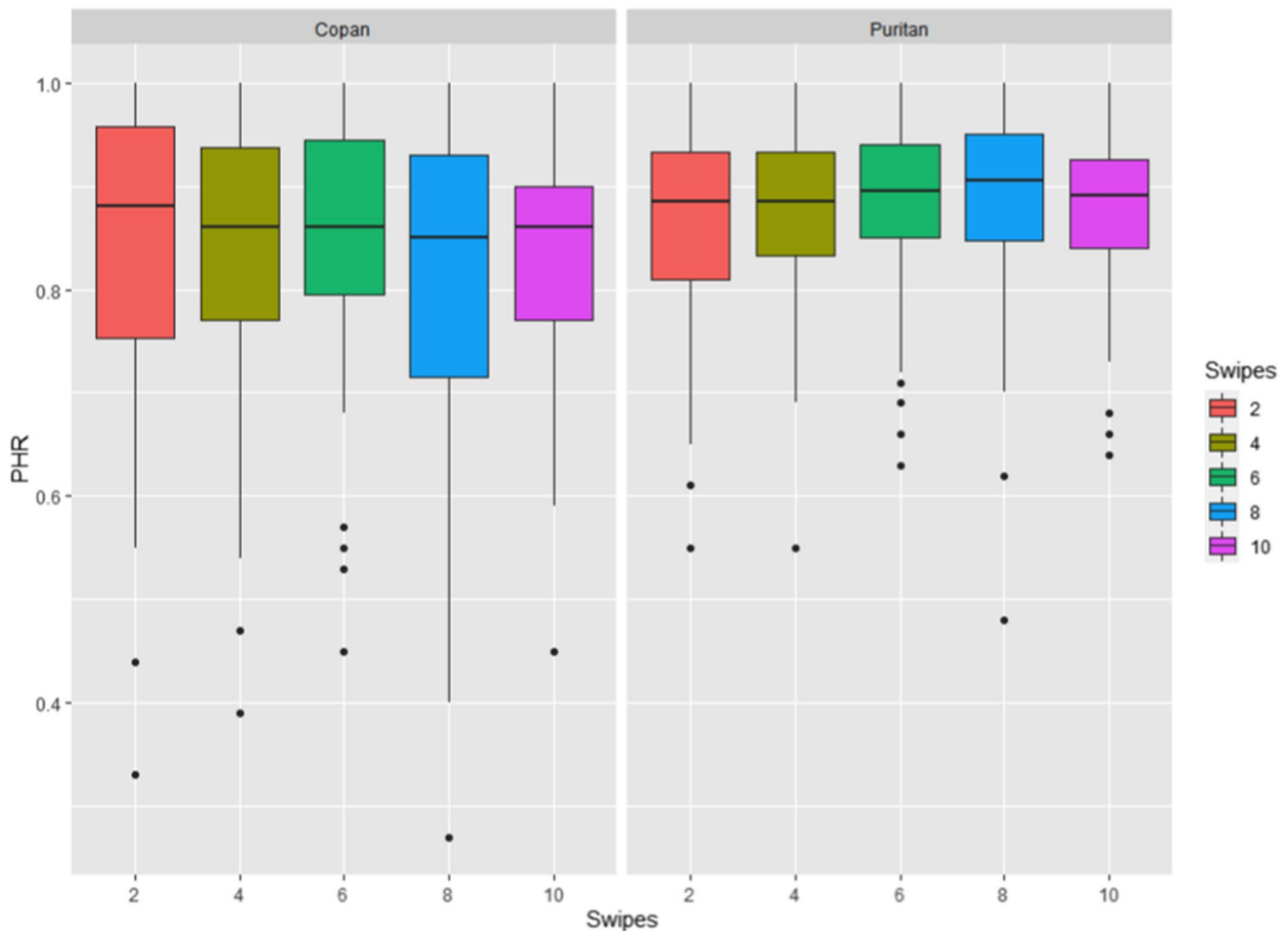
Fig. 8 Boxplots illustrating PHRs observed for each swipe number (grouped by swab type) from the 40 swabs typed by the RapidHIT ID System in the Sensitivity Study. The horizontal bar within the box represents the median for the associated data. The "whiskers" or vertical bars on the boxplot represent 1.5 * interquartile range. The individually drawn points (black dots) represent points that extend beyond 1.5*interquartile range

respectively. The average Total Signal of the diploid loci ranged from 1,931 ($\pm$ 1,380) RFUs at D8S1179 to 5,776 ($\pm$ 3,408) RFUs at D2S441. The average PHR for the diploid loci ranged from 0.82 ($\pm$ 0.14) for SE33 to 0.92 ($\pm$ 0.06) for D2S441. Average PHR was lower among the larger loci (e.g., SE33), a trend that was similar to that of previous studies examining multiple versions of the RapidHIT ID System [13, 42, 44, 48].

When analyzing the ten positive control cartridges included throughout the validation study, there were no quality flags generated for any of the 430 alleles called in the positive control samples. Further, each of the ten positive control runs produced profiles concordant with each other. For the 51 concordance samples collected

for this study, STR profiles were produced by the 3500 Genetic Analyzer workflow and the RapidHIT ID System. When evaluating the FGA locus in two of the profiles (C5 and C34) generated by the 3500 Genetic Analyzer workflow, both alleles were present for these two heterozygous samples. However, the peak heights were below the 175 RFU threshold used for making allele calls. Any discordances between the profiles generated by the two workflows for these 51 samples are detailed in Table 2. The initial Rapid runs for C6 and C45 did not produce a profile due to a "size call failed" error. These two swabs were immediately re-run in a new sample cartridge, producing a full, concordant profile. A total of 2,045 allele peaks were called in the
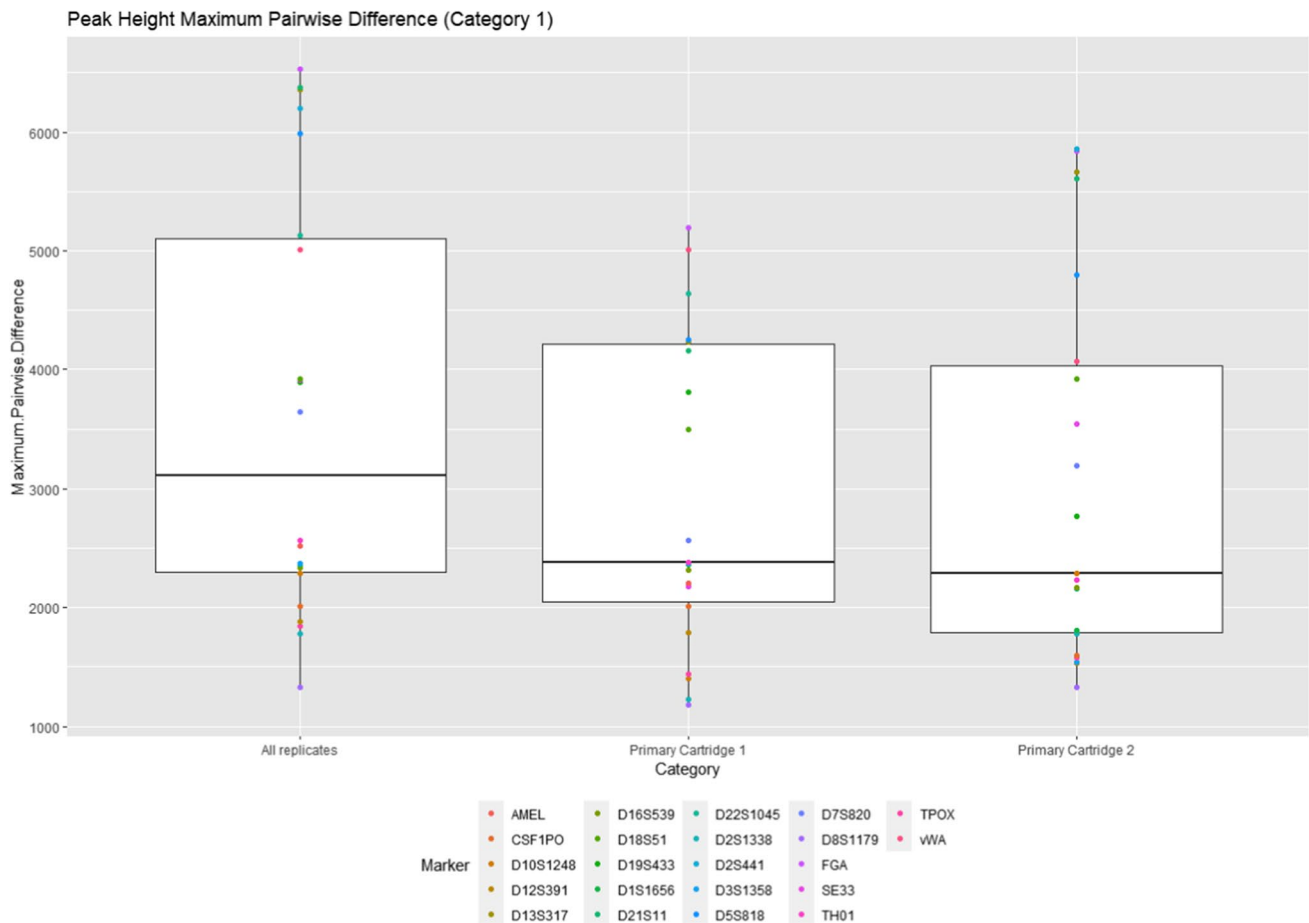
Fig. 9 Maximum pairwise difference in Total Signal across the 396 loci genotyped in the 18 swabs processed for Category 1 of the Reproducibility and Repeatability Study grouped by primary cartridge usage. The overlayed scatter plot shows the maximum pairwise difference for each locus. The horizontal bar within the box represents the median for the associated data. The "whiskers" or vertical bars on the boxplot represent 1.5 * interquartile range

51 STR profiles generated by the RapidHIT ID System. Fourteen of these allele calls (0.68%) were discordant with the corresponding profiles generated by the 3500 Genetic Analyzer Workflow (Table 2). Thirteen of these discordant allele calls (93%) were labeled with a quality flag by the RapidHIT ID System (first-pass success rate for the study as a whole is discussed in a separate section below). The one discordant allele call that did not receive a quality flag came from a female sample (C5) that displayed an allele called at DYS391 which was due to pull-up (Supplementary Fig. 1); this artifact was not present in this sample's profile from the 3500 Genetic Analyzer workflow. Three of the flagged discordant results were also from pull-up artifacts in sample C8 (Table 2; Supplementary Fig. 2). Eight of the flagged discordant results were from stutter peaks that labeled for review (Table 2). The remaining two flagged discordant results were alleles in samples C26 and C27 at the SE33 locus that were labeled

with an OB quality flag instead of an allele name. The un-called allele in sample C26 was the micro-variant allele18.3, while the un-called allele in sample C27 was a 31.2 allele. Modified Rapid DNA analysis (or secondary manual review) by a trained/qualified scientist correctly called 12 of the 14 quality flagged alleles and the one discordant allele call that did not receive a quality flag, increasing the total number of full, concordant profiles generated. The remaining two allele calls with OB quality flags (0.1% = 2/2033 (now that 12 peaks were identified as stutter peaks or artifacts)) would have likely required reanalysis of the swab, a new swab, or use of a differing ladder to confirm the true allele call.

Contamination study

The contamination study was performed to detect potential run-to-run contamination or carry-over. For this
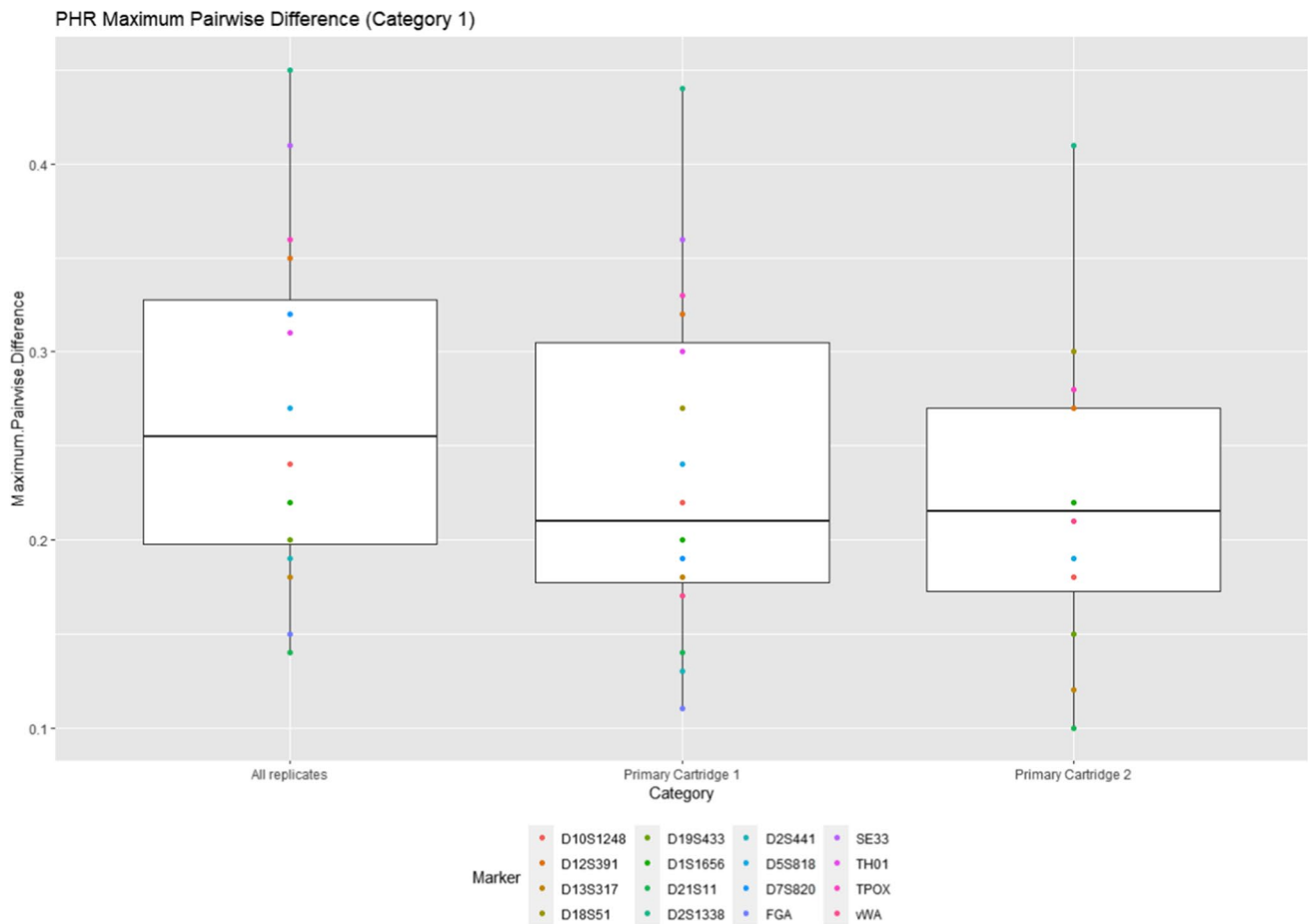
Fig. 10 Maximum pairwise difference in PHR across the 396 loci genotyped in the 18 swabs processed for Category 1 of the Reproducibility and Repeatability Study grouped by primary cartridge usage. The overlayed scatter plot shows the maximum pairwise differ-ence for each locus. The horizontal bar within the box represents the median for the associated data. The "whiskers" or vertical bars on the boxplot represent 1.5 * interquartile range

study, ten unused swabs were run interspersed through-out the swabs run for the Concordance study. While a "yellow check" was expected for the Contamination Study samples as an unused swab was placed in an ACE GlobalFiler Express Sample Cartridge, five of the ten unused, processed swabs also had one peak called and labeled with a quality flag at the D22S1045 locus (Fig. 3A). Four of these peaks were labeled with an IHO quality flag, and the fifth peak was labeled with an OB and IHO quality flag. This labeled peak was the same for each of the five swabs despite running different Concord-ance Study swabs in between each unused swab. This peak also differed from the instrument operator's geno-type at the D22S1045 locus. This peak was not present in the second half of the Contamination Study samples, which were run with a different lot number of the ACE GlobalFiler Express Sample Cartridge. This observation and being a single peak suggested that it was a lot-spe-cific artifact from this lot of cartridges. To verify, the

previously processed swabs exhibiting an artifact were removed from the sample cartridge and reprocessed in a new sample cartridge of a different lot. When pro-cessing the exact same swab in a new sample cartridge, there were no peaks observed at the D22S1045 locus (Fig. 3B). Subsequently, this peak was reported to the manufacturer who confirmed it was a dye artifact (per-sonal communication).

Eight ACE Negative Control Cartridges also were run throughout the validation study. Seven of these negative controls received "green checks" and had zero alleles called after primary analysis with the RapidHIT ID System. The eighth negative control received a "red X" with two allele calls (Fig. 4A). The peak at the D2S441 locus appeared to be the likely result of unused primer sequences. The peak at the D22S1045 locus was the same peak seen in the five Contamination Study swabs described above. These two allele calls differed from the instrument operator's genotype. However, the lot number
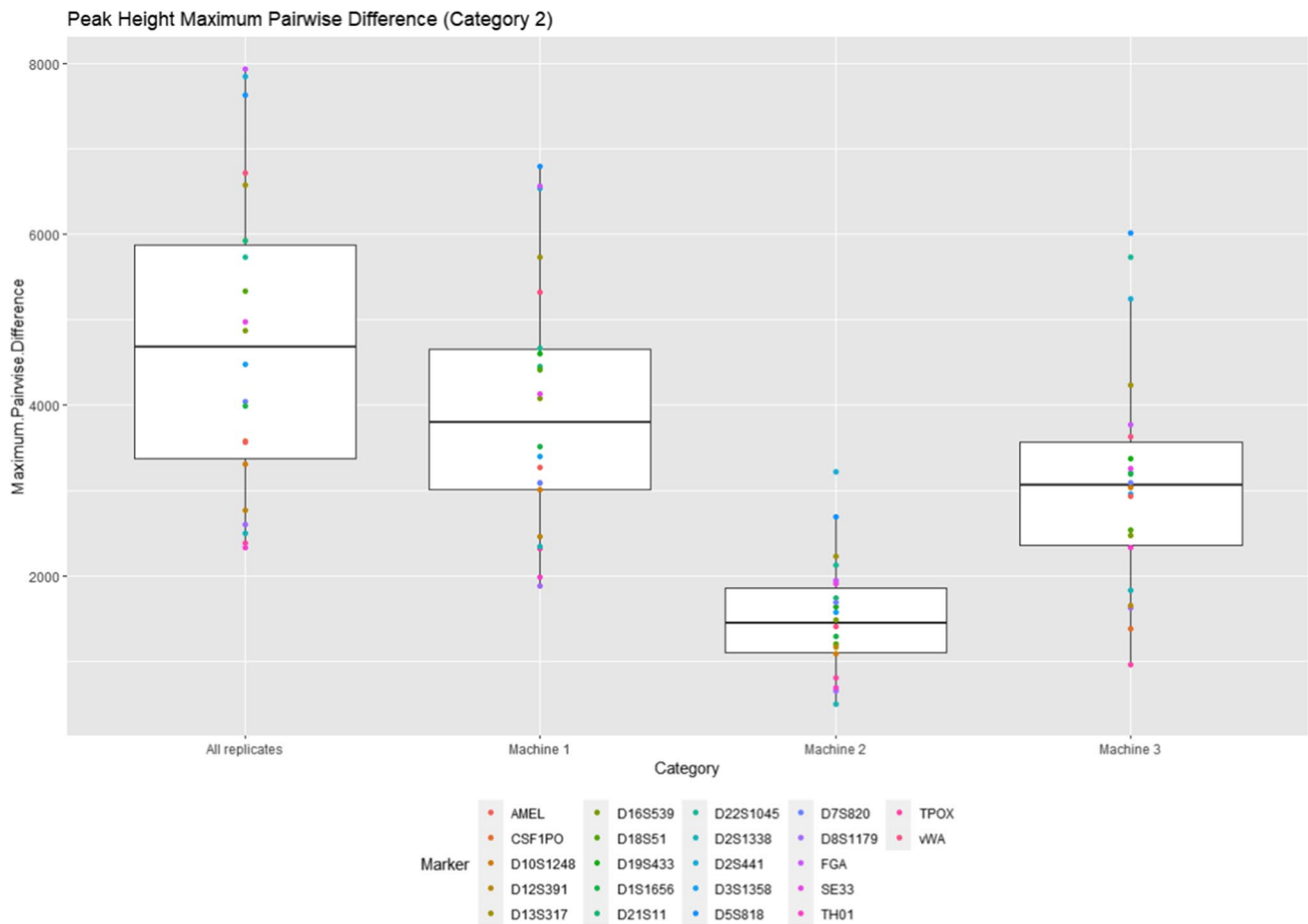
Peak Height Maximum Pairwise Difference (Category 2)

Fig. 11 Maximum pairwise difference in Total Signal across the 198 loci genotyped in the 9 swabs processed for Category 2 of the Reproducibility and Repeatability Study grouped by instrument usage. The overlayed scatter plot shows the maximum pairwise difference for each locus. The horizontal bar within the box represents the median for the associated data. The "whiskers" or vertical bars on the boxplot represent 1.5 * interquartile range

for the ACE Negative Control Cartridge was the same as the lot number from the five Contamination Study swabs discussed above. To further confirm these two peaks were not the result of any cartridge leakage or run-to-run contamination, another ACE Negative Control Cartridge from a different lot was run immediately afterward (Fig. 4B). This negative control run received a "green check" with zero called alleles, further supporting a lot-specific artifact.

While no carry-over or run-to-run contamination was observed in this study, the additional value of detecting any potential artifacts or lot-specific effects when processing negative controls was highlighted. Similar results (i.e., detection of primer flare peaks, dye blobs, artifacts, "anomaly peaks", etc.) have been documented in previous studies [14, 22, 42]. For this study, these artifacts did not affect the ability to accurately genotype the DNA positive samples but were more readily visible in the negative control samples.

Sensitivity study

The purpose of this study was to determine the reliability of results from a range of input DNA quantities. Collection protocols for reference swabs (and the number of swipes) can vary from user to user and at times is not well-defined. For this study, a swipe was defined as one motion from the top of the cheek to the bottom of the cheek. Since the workflow for the RapidHIT ID System does not require a DNA quantification step, swipe number was used as a surrogate for DNA input values in this study. Thus, swabs collected with a varying number of swipes ranging from two to ten swipes in two swipe intervals were assessed in this study.

For the Puritan swab sensitivity series, each replicate for both subjects at each number of swipes (n = 4 per number of swipes) produced complete profiles (Fig. 5) with no quality flags and were concordant with each processed replicate. For the Copan swab sensitivity series, six of the 20 processed swabs received a quality flag and,
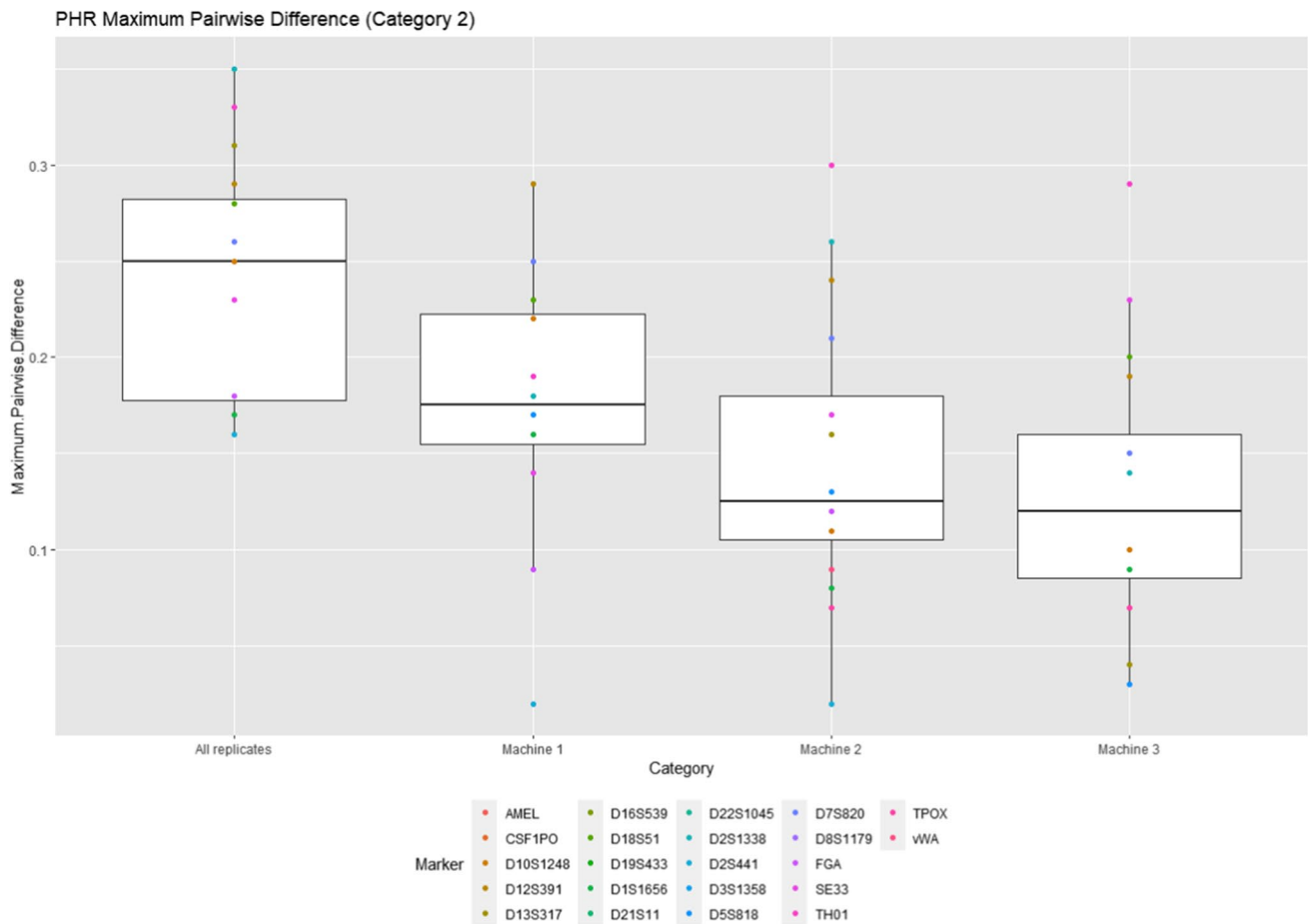
Fig. 12 Maximum pairwise difference in PHR across the 198 loci genotyped in the 9 swabs processed for Category 2 of the Reproducibility and Repeatability Study grouped by instrument usage. The overlayed scatter plot shows the maximum pairwise difference for each locus. The horizontal bar within the box represents the median for the associated data. The "whiskers" or vertical bars on the boxplot represent 1.5 * interquartile range

subsequently, a "yellow check." The number of called alleles compared to the total number of possible alleles for each swab in the Copan sensitivity series is illustrated in Fig. 6. For the two swipe swabs, both replicates of Sample 1 had 37 of 38 alleles called. Allele drop-out was present at the TPOX locus with the called TPOX allele also labeled with an IHO quality flag. The larger allele at the D18S51 locus for replicate B (Sample 1's two swipe replicate B) also was labeled with an IMB quality flag. For the four swipe swabs, locus drop-out (TPOX and SE33) was present in Sample 2's four swipe replicate A, resulting in 38 out of 42 alleles called. The larger allele at the D2S1338 locus for this sample also was labeled with an IMB quality flag. For the six swipe swabs, allele drop-out at the TPOX locus was present in Sample 2's six swipe replicate A, resulting in 41 out of 42 alleles called. The called TPOX allele was labeled with an IHO quality flag, and a stutter peak was called and labeled with an IMB and PL quality flag at the D12S391 locus. For

the eight swipe swabs, full profiles were generated for each swab. However, Sample 1's eight swipe replicate B swab had the larger allele at the D2S1338 locus labeled with a IMB quality flag, prompting manual review of the profile. For the ten swipe swabs, Sample 1's ten swipe replicate A had 15 out of 38 alleles called, with the allele calls at the D8S1179 and D13S317 loci receiving an IHO quality flag. Based on the results and allele calls from the entire sensitivity series, Sample 1's ten swipe replicate A swab is likely an outlier with potential problems during sample collection.

Complete, concordant profiles were generated for 75% of the two swipe swabs, 88% of the four swipe swabs, 88% of the six swipe swabs, 100% of the eight swipe swabs, and 100% (excluding outlier) of the ten swipe swabs. The Total Signal and PHR performance metrics were evaluated to see if similar trends were observed (Figs. 7 and 8). For this study, the Total Signal from each locus of each profile was grouped by swipe
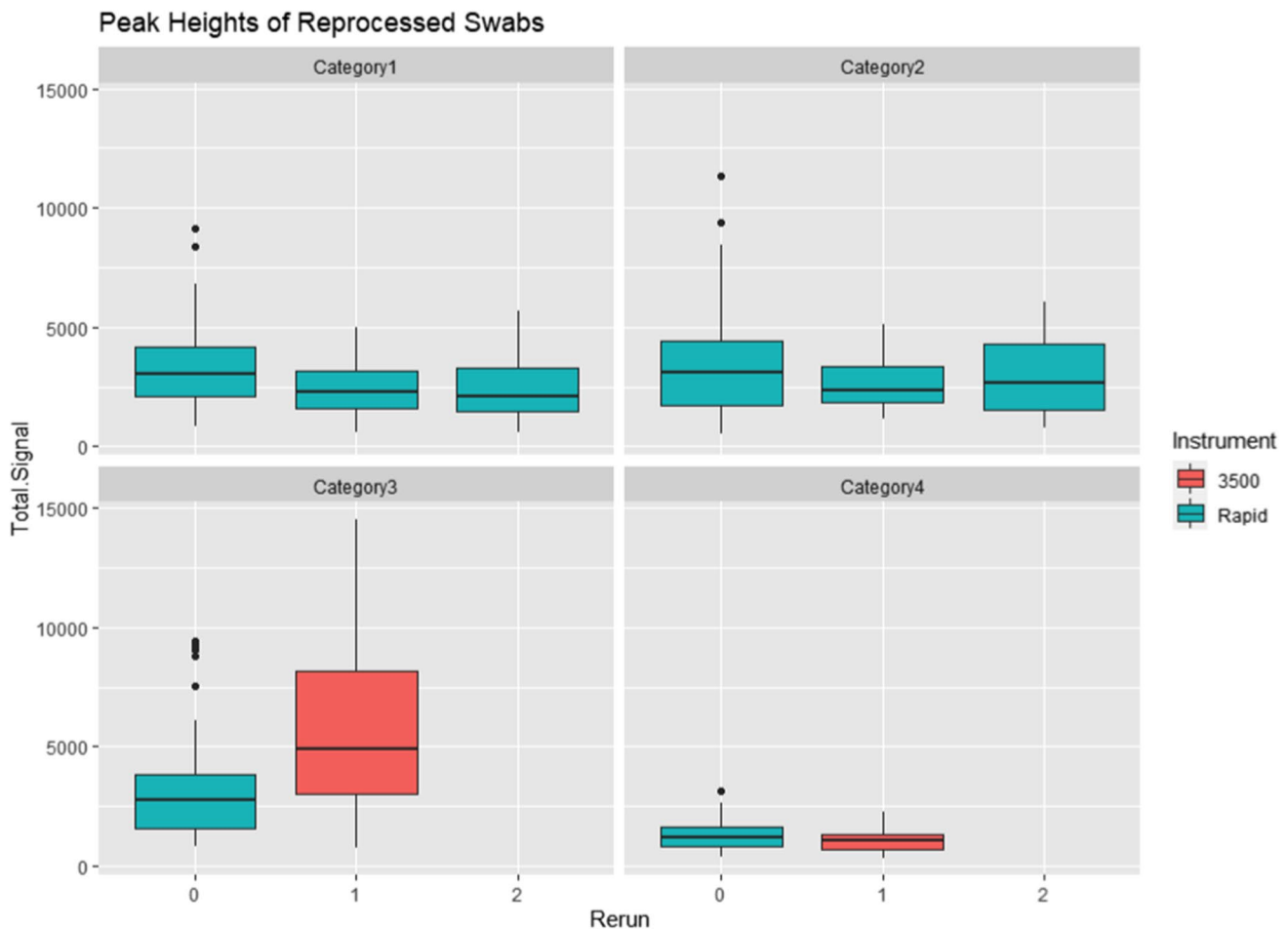
Fig. 13 Boxplots illustrating Total Signal observed for the genotyped loci from swabs processed (Rerun 0) and reprocessed (Rerun 1 or 2) on the RapidHIT ID System or 3500 Genetic Analyzer in the Swab Reprocessing Study. Results are grouped into their experimental categories which detail the instrument and manner in which swabs were reprocessed. The horizontal bar within the box represents the median for the associated data. The "whiskers" or vertical bars on the box-plot represent 1.5 * interquartile range. The individually drawn points (black dots) represent points that extend beyond 1.5*interquartile range

number and averaged. For the Puritan swabs, the average Total Signal ranged from 3,825 ($\pm$ 2,909) RFUs for the two swipe swabs to 3,720 ($\pm$ 2,049) RFUs for the ten swipe swabs, and the average PHR ranged from 0.86 ($\pm$ 0.10) for the two swipe swabs to 0.88 ($\pm$ 0.08) for the ten swipe swabs. For the Copan swabs, the average Total Signal ranged from 915 ($\pm$ 638) RFUs for the two swipe swabs to 1,982 ($\pm$ 1,644) RFUs for the ten swipe swabs, and the average PHR ranged from 0.84 ($\pm$ 0.14) for the two swipe swabs to 0.83 ($\pm$ 0.11) for the ten swipe swabs. Loci such as TPOX and SE33 are the largest loci in their respective dye channels and showed the lowest performance in both Total Signal and PHR in their dye channel (Figs. 1 and 2). These results align with the performance and drop-out seen in the Sensitivity Study as allele and locus drop-out occurred only in the TPOX and SE33 loci.

Repeatability and reproducibility study

The purpose of the Repeatability and Reproducibility Study was to assess the variability of profiles generated from different primary cartridges (Category 1) and RapidHIT ID instruments (Category 2). A total of 27 swabs were processed on the RapidHIT ID System in this study. Complete profiles that were concordant with each of the processed replicates were produced for 26 of the swabs (96%) and generated a "green check". Machine 2's Day 2 swab produced an SRI2 quality flag for the profile, indicating an issue arose during sizing that could affect accurate allele calling. This swab was immediately rerun in a new sample cartridge and produced a complete, concordant profile.

To expand on the extent of variability across different primary cartridges or different RapidHIT ID instruments, the
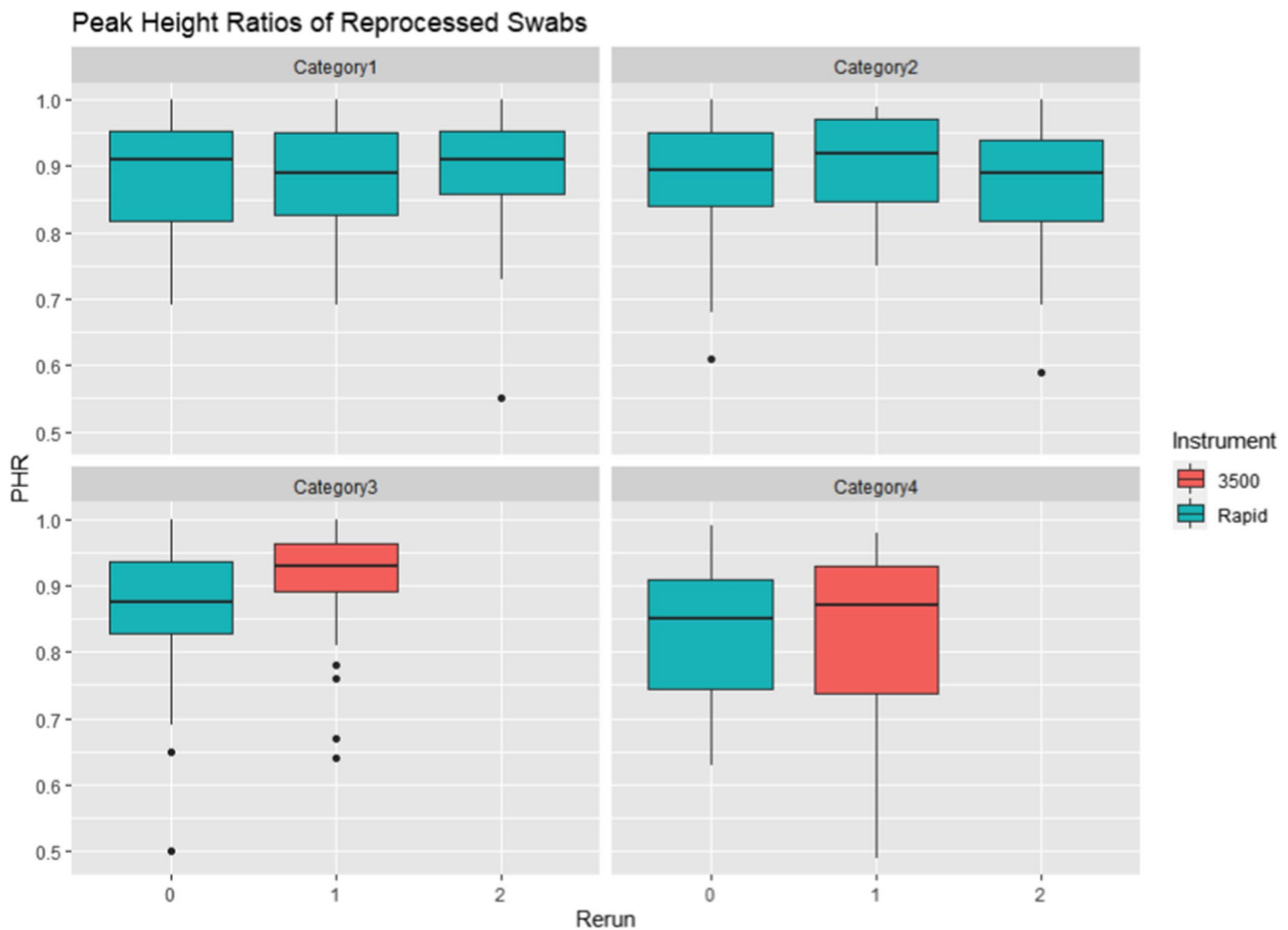
Fig. 14 Boxplots illustrating PHR observed for the genotyped loci from swabs processed (Rerun 0) and reprocessed (Rerun 1 or 2) on the RapidHIT ID System or 3500 Genetic Analyzer in the Swab Reprocessing Study. Results are grouped into their experimental categories which detail the instrument and manner in which swabs were reprocessed. The horizontal bar within the box represents the median for the associated data. The "whiskers" or vertical bars on the boxplot represent 1.5 * interquartile range. The individually drawn points (black dots) represent points that extend beyond 1.5*interquartile range

Total Signal and PHR for the 594 loci that were genotyped across these 27 swabs were used to calculate the maximum pairwise difference (maximum for each locus – minimum for each locus) in Total Signal and PHR across the replicates in each category (Figs. 9, 10, 11, 12). For Category 1, 396 loci were genotyped in the 18 Category 1 swabs. The average maximum pairwise difference in Total Signal was 3,648 ($\pm$1,780) RFUs for Category 1 (Fig. 9). The average and spread of this metric were lower when only comparing replicates run on the same primary cartridge as the average maximum pairwise difference was 2,964 ($\pm$1,343) RFUs and 3,071 ($\pm$1,582) RFUs for Primary Cartridge 1 and Primary Cartridge 2, respectively. For Category 1, the average maximum pairwise difference in PHR for the heterozygous loci was 0.27 ($\pm$0.09) (Fig. 10). The average of this metric

was lower when comparing replicates processed on the same primary cartridge where the average maximum pairwise difference for PHR was 0.24 ($\pm$0.09) and 0.22 ($\pm$0.08) for Primary Cartridge 1 and Primary Cartridge 2, respectively. For Category 2, 198 loci were genotyped in the nine Category 2 swabs. The average maximum pairwise difference in Total Signal was 4,729 ($\pm$1,811) RFUs for Category 2 (Fig. 11). The average and spread of this metric were lower when only comparing replicates run on the same instrument as the average maximum pairwise difference was 3,981 ($\pm$1,496) RFUs, 1,522 ($\pm$665) RFUs, and 3,113 ($\pm$1,323) RFUs for Machine 1, Machine 2, and Machine 3, respectively. For Category 2, the average maximum pairwise difference in PHR for the heterozygous loci was 0.24 ($\pm$0.06) (Fig. 12). The average of this metric was lower when only comparing
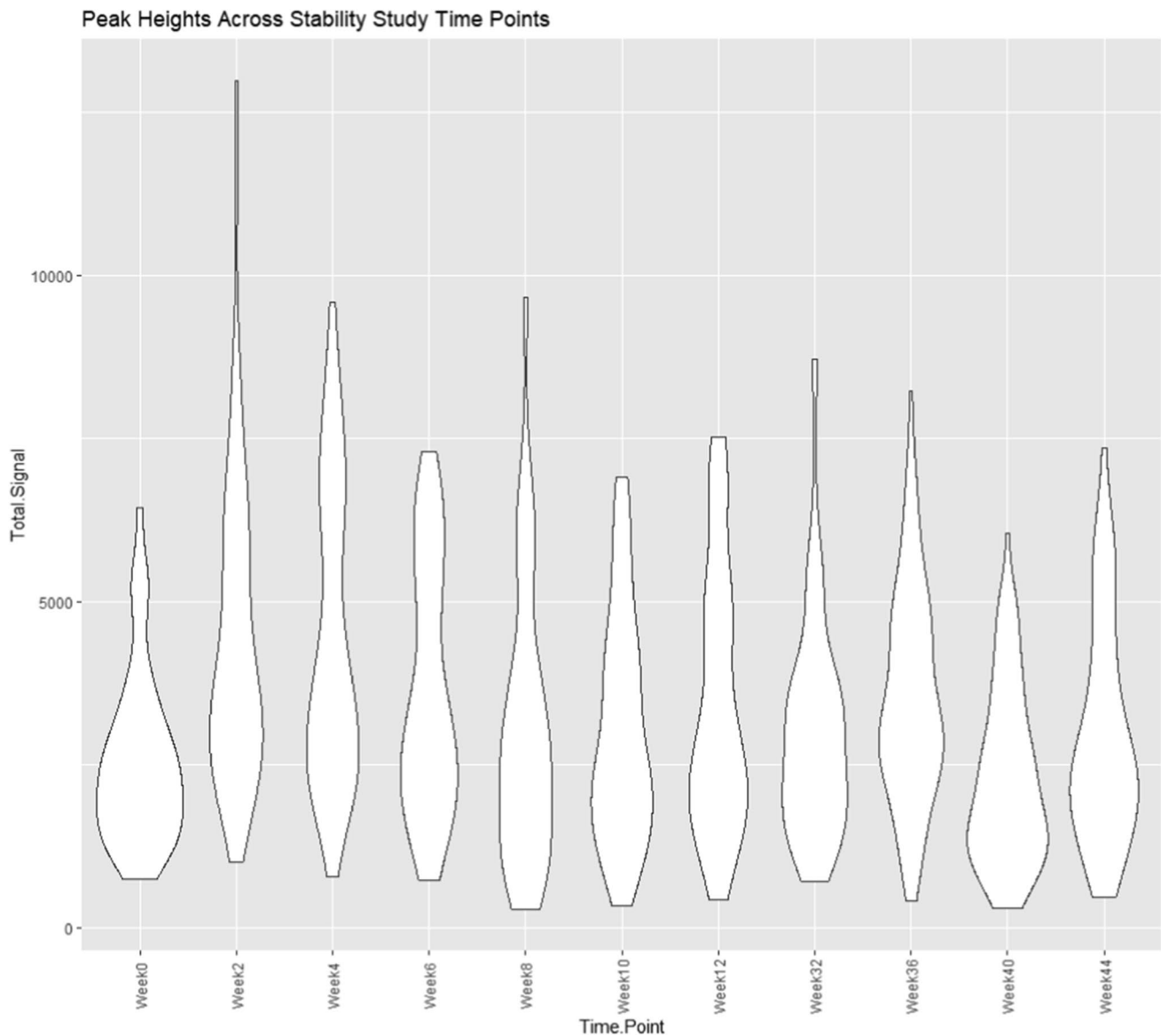
Fig. 15 The distribution of Total Signal for each genotyped locus at each evaluated time point is illustrated with a violin plot

replicates run on the same instrument as the average maximum pairwise difference was 0.17 (±0.07), 0.14 (±0.08), and 0.13 (±0.07) for Machine 1, Machine 2, and Machine 3, respectively. The decrease in variability of the Total Signal and PHR performance metrics when evaluating profiles generated on the same primary cartridge or instrument compared to differing primary cartridges or instruments was expected. However, this variability in performance for Total Signal and PHR, which was not substantial, did not impact the RapidHIT ID System's ability to generate complete and concordant profiles when different primary cartridges were used on the same instrument or when different RapidHIT ID instruments were used.

The overlayed scatter plot in Figs. 9 to 12 shows the specific maximum pairwise difference for each locus, allowing for visualization of the most variable loci per metric. The amount of DNA collected with each swab can and will vary even with uniform collection parameters, likely accounting for some of the variability seen in the performance of each locus. Despite this limitation in sample collection, for PHR (Figs. 10 and 12), the same group of loci repeatedly show up in the uppermost quartile of each boxplot (e.g., D2S1338, SE33, TPOX, THO1). The higher variability in PHR of these loci across the replicates in this study corresponded well with the elevated stochastic thresholds put in place by the manufacturer for these loci [48].
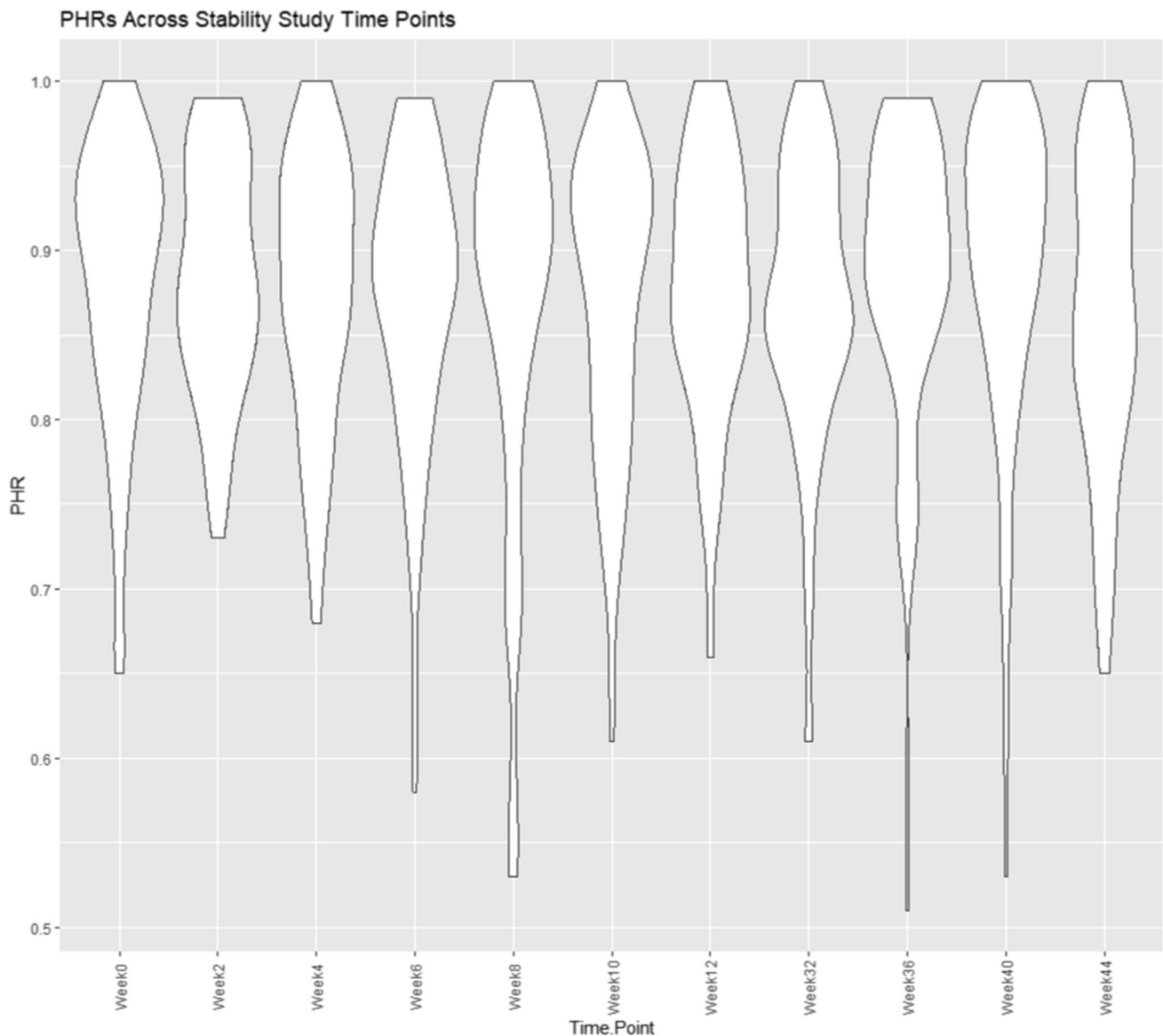
Fig. 16 The distribution of PHR for each genotyped locus at each evaluated time point is illustrated with a violin plot

Swab reprocessing study

With the RapidHIT ID System, the processed swab can be removed from the ACE GlobalFiler Express Cartridge. Thus, the purpose of the Swab Reprocessing Study was to assess the potential for reanalysis of the same swab with the RapidHIT ID System workflow and/or the 3500 Genetic Analyzer workflow. Experiments for this study were grouped into categories based on the instrument and manner in which the swabs were reprocessed, resulting in a total of 30 STR profiles that were evaluated.

Each STR profile produced in this study (on both the RapidHIT ID and 3500 Genetic Analyzer instrument) produced concordant results with the exception of two alleles.

For Subject 3 in Category 4 (long-term storage), the reprocessing run on the 3500 Genetic Analyzer displayed a 20 allele (153 RFU peak height) at the FGA locus that fell below the 175 RFU threshold (Supplementary Fig. 3). As both alleles of this heterozygous locus were present, the 20 allele was manually called upon secondary review of the data. For the initial run on the RapidHIT ID instrument for Subject 2 in Category 2 (overnight dry before reprocessing), a 6 allele was called, without a quality flag, at the DYS391 locus of this female sample (Supplementary Fig. 4). Additional replicates of this swab from the Reprocessing Study did not include the allele call at the DYS391 locus, and the RapidHIT ID instrument operator was ruled out as a potential source of contamination. As discussed above in the
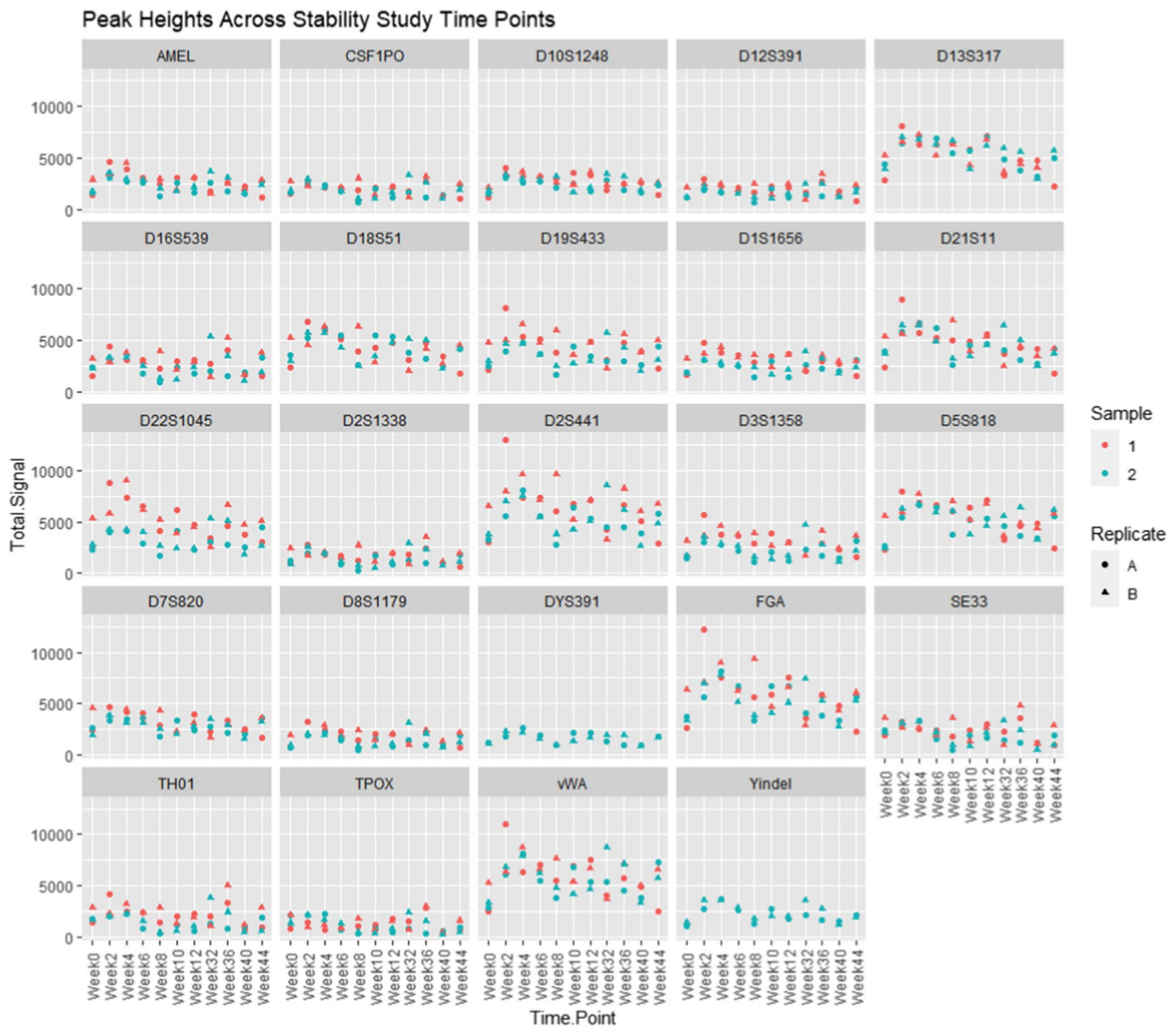
Fig. 17 Scatter plot of Total Signal values for each swab at each evaluated time point grouped by locus

Concordance Study, this allele call at the DYS391 locus was likely a pull-up artifact. Manual review flagged and removed this allele call as no additional Y-alleles were called in this profile and the 6 allele called at the DYS391 locus was off-balance from the rest of the profile.

The Total Signal and PHR of each locus genotyped in the Swab Reprocessing Study is illustrated in Figs. 13 and 14. When grouping the swabs by their experimental category, performance metrics averages were calculated for each of the original runs (Rerun 0) and reruns completed in each category. In Category 1, the average Total Signal was 3,3812 (±1,754) RFUs, 2,460 (±1,111) RFUs, and 2,367 (±1,306) RFUs for rerun 0, rerun 1, and rerun 2, respectively. In Category 2, the average Total Signal was 3,496

(±2,378) RFUs, 2,706 (±1,136) RFUs, and 2,970 (±1,626) RFUs for rerun 0, rerun 1, and rerun 2, respectively. In Category 3, the average Total Signal was 3,214 (±2,179) RFUs and 5,809 (±3,254) RFUs for rerun 0 and rerun 1, respectively. In Category 4, the average Total Signal was 1,253 (±607) RFUs and 1,065 (±455) RFUs for rerun 0 and rerun 1, respectively. A decrease was seen in the average Total Signal for reruns 1 and 2 in Category 1 and 2, suggesting that reprocessing a swab had an effect on Total Signal. This observation is not unexpected as less DNA would be left on the swab with each subsequent run on the RapidHIT ID instrument. However, this trend in Total Signal was not seen in Categories 3 and 4. The extraction, quantification, and normalization steps in the 3500 Genetic Analyzer workflow

Fig. 18 Scatter plot of PHR values for each swab at each evaluated time point grouped by locus

for the reruns in Category 3 and 4 could explain the difference in the Total Signal trend seen in Category 1 and 2 reruns. In Category 1, the average PHR was 0.89 ($\pm$0.08), 0.88 ($\pm$0.09), and 0.89 ($\pm$0.09) for rerun 0, rerun 1, and rerun 2, respectively. In Category 2, the average PHR was 0.88 ($\pm$0.09), 0.91 ($\pm$0.07) and 0.87 ($\pm$0.10) for rerun 0, rerun 1, and rerun 2, respectively. In Category 3, the average PHR was 0.87 ($\pm$0.10) and 0.91 ($\pm$0.08) for rerun 0 and rerun 1, respectively. In Category 4, the average PHR was 0.82 ($\pm$0.11) and 0.83 ($\pm$0.12) for rerun 0 and rerun 1, respectively. The comparable average PHR values across the runs in each category suggests reprocessing a swab had minimal effect on the PHR performance metric.

Stability study

The purpose of the Stability study was to assess the effect of aged swabs, maintained under ambient conditions for different lengths of time, on the generation of STR profiles from reference buccal swabs with the RapidHIT ID System. A total of 44 swabs were evaluated in this study. Due to the COVID-19 pandemic, this experiment was stopped and re-started. However, this impact offered the opportunity to add additional, longer-term time points into the study (i.e., the 32-, 36-, 40-, and 44-week time points). Each of the 44 swabs generated full profiles with zero quality flags (producing a "green check"), and each
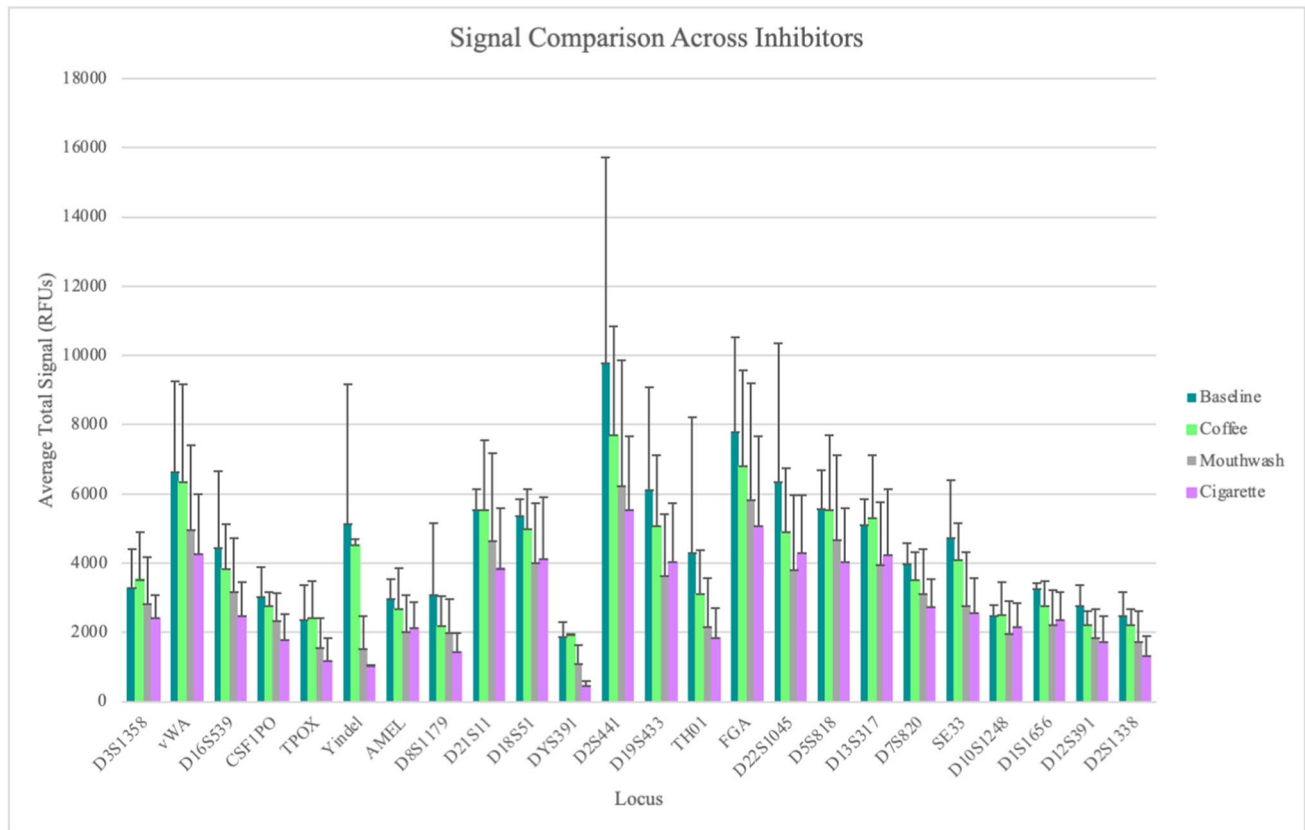
Fig. 19 The average total signal for each locus included in the GlobalFiler Express multiplex grouped by inhibitor (or lack thereof) is pictured. Vertical black bars represent one standard deviation

of these aged profiles was concordant with its associated baseline swab (zero-week time point). Peak heights and PHRs are illustrated in Figs. 15 and 16 with no visible trend or effect on either metric as storage time increased. Figures 17 and 18 illustrate the relatively consistent peak heights and PHRs across each evaluated time point on a per-marker basis. One trend worth noting is the increase in peak height for many of the loci from the Week0 to Week2 time point, suggesting a potential increase in performance if the swabs are given some time to dry versus loading into the sample cartridge immediately after collection.

Inhibition study

The purpose of the Inhibition study was to assess the effect of inhibitors, often found in the mouth, on the generation of STR profiles from reference buccal swabs with the RapidHIT ID System. A total of 18 swabs and three inhibitors, coffee, mouthwash, and cigarettes, were evaluated in this study. Each of the 18 swabs generated full profiles with zero quality flags (producing a "green check"), and each of these inhibited profiles was concordant with its associated baseline swab run as part of

the Concordance study. To further evaluate the inhibitor's potential effect on generation of an STR profile, the inhibited swabs were grouped by inhibitor, and the Total Signal from each locus included in the GlobalFiler Express multiplex was averaged across the profiles generated from each swab containing an inhibitor (and associated baseline). The average Total Signals for the three inhibitors are shown in Fig. 19. The average Total Signal ranged from 1,857 RFUs to 9,773 RFUs for the baseline swabs, from 1,954 RFUs to 7,709 RFUs for the coffee inhibitor swabs, from 1,081 RFUs to 6,230 for the mouthwash inhibitor samples, and from 455 RFUs to 5,527 RFUs for the cigarette inhibitor samples. Cigarette inhibitor samples had the lowest range of total signal values from each of the evaluated loci, consistent with the results seen in the inhibition studies of Wiley et al. [14]. The DYS391 and D2S441 loci generated the minimum and maximum RFUs, respectively, consistent with Concordance Study performance metric results, for each inhibitor's (and baseline) Total Signal ranges indicating amplification efficiency played a part in the variability seen in each inhibition category and was consistent across each category. Overall, the above results suggest
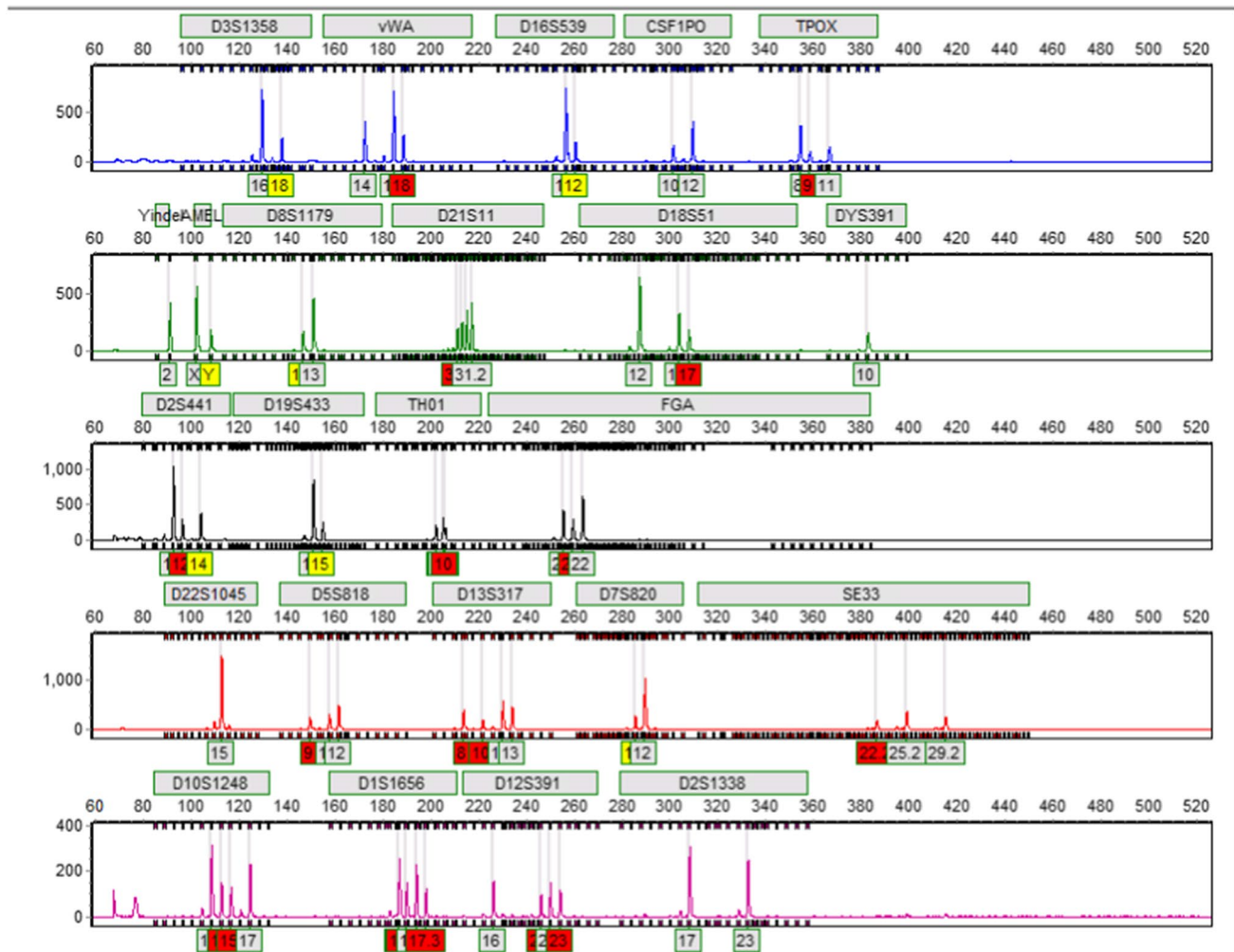
Fig. 20 Electropherogram of the profile for Mixture 1–4 (Replicate A). This profile received both heterozygote imbalance (IMB) and ploidy (PL) quality flags as all 63 possible alleles from these two contributors are represented

that the evaluated inhibitors did not impact the RapidHIT ID System's ability to produce complete and concordant profiles, despite the varying impact each inhibitor had on performance metrics.

Mixture study

The purpose of the Mixture study was to assess the RapidHIT ID System's ability to flag mixtures. A total of 21 swabs were evaluated in this study, and mixture ratios were based on the volume of saliva (not the amount of input DNA) from each contributor added to the swab. Each of the single-source swabs genotyped in this study produced complete profiles concordant with each replicate swab included in the Mixture study and concordant with the replicate swab included in the Concordance study. The RapidHIT ID System detected and flagged

(producing a "yellow check") a second contributor in 12 of the 15 mixed samples (80%) genotyped in this study (example shown in Fig. 20). Each of the 9–1 mixture ratio replicates produced a non-flagged (or "green check") profile (example shown in Fig. 21). To further evaluate these results, the number of unique alleles called for each contributor was calculated (Fig. 22). For the single source samples, only alleles from one contributor were called, aligning with the non-flagged ("green check") profiles that were generated for these samples. For the majority of the mixed samples (80%), unique alleles from both contributors were called and subsequently flagged. For Replicate B of the 9–1 mixture, two alleles from the minor contributor were called but not flagged by the RapidHIT ID System as a mixed profile. Further review of this profile (Fig. 21) showed that this mixture was the result of a male and female contributor,
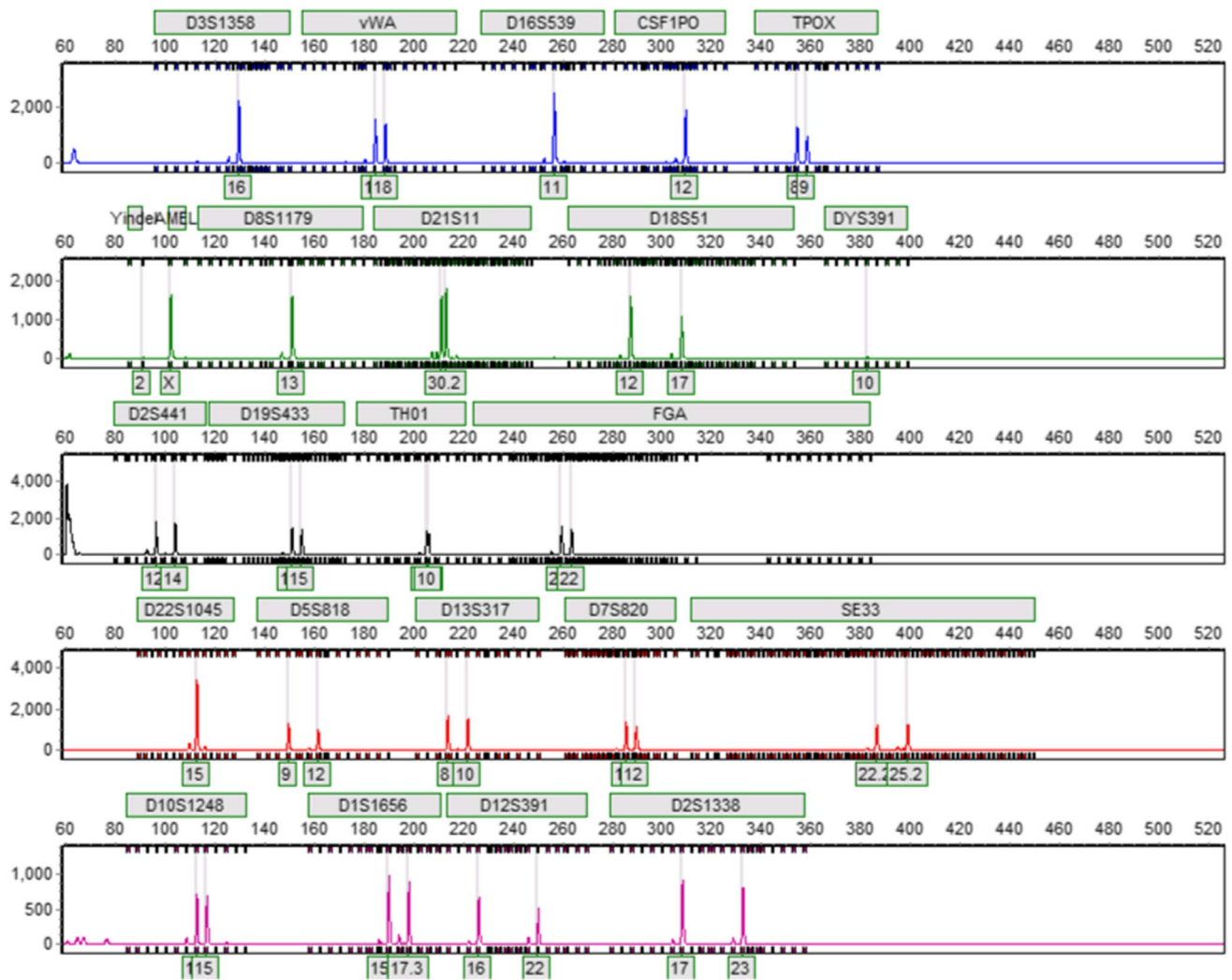
Fig. 21 Electropherogram of the profile for Mixture 9–1 (Replicate B). This profile illustrates alleles from both contributors but did not receive any quality flags during automated analysis

and the two non-flagged alleles from the minor contributor were Y alleles that were not present in the female contributor. Thus, the RapidHIT ID System did not flag one (6.7%) of the mixed samples with alleles called from both contributors. The two Y-alleles from the minor contributor were identified upon secondary manual review. Additionally, the Y-alleles are not included in the auto-generated CMF file from the GeneMarker HID software, eliminating the possibility of including alleles from a male minor contributor with a female major contributor's profile when conducting database searches [53].

Precision study

Accurate sizing is needed for generation of reliable genotyping calls. Thus, the purpose of the Precision Study was to assess ladder selection and the difference in sizes

between alleles of the "unknown" sample and selected ladder. A total of 253 samples were processed in this validation study, where 247 (98%) of these samples were processed on the same RapidHIT ID instrument, offering an opportunity to evaluate ladder selection trends. Run number and sample are likely to impact ladder selection and are included as variables in Fig. 23 to visualize any trends or potential effects with the ladder selected. Progressively aging polymer potentially could affect migration of the amplified DNA fragments and, subsequently, ladder selection. However, no visible trend was observed in ladder selection across run number (i.e., age of the primary cartridge). In fact, three ladders (ladder 49, ladder 54, and the native ladder) were chosen in 193 (78%) of the RapidHIT ID System runs on the same instrument. Due to the design of some of the studies included in this validation (e.g., Stability Study and Reproducibility and Repeatability Study), multiple swabs
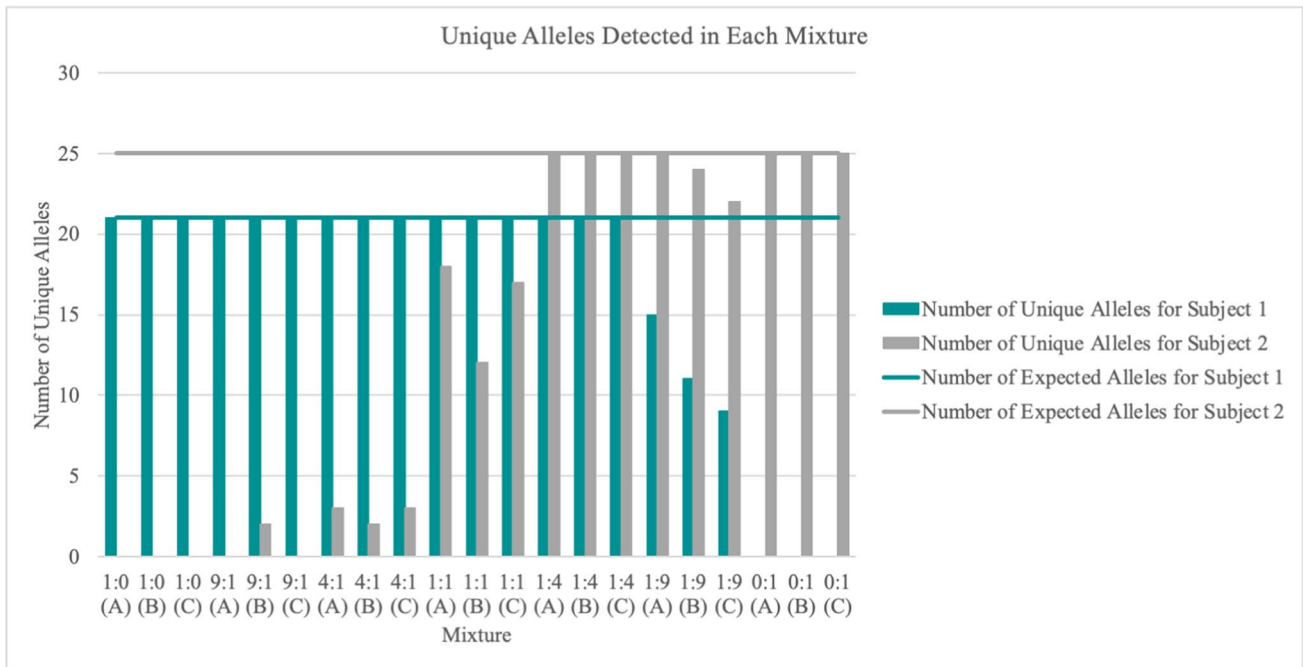
Fig. 22 The number of unique alleles from each contributor included in each single-source and mixed profile is illustrated. Mixture ratios were based on the volume of saliva from each contributor that was added to the swab. The horizontal lines represent the total number of possible unique alleles for each contributor

were collected from the same individual and evaluated. This repetitive swab collection potentially impacted ladder selection as well. While ladder 49 was selected in 60 (24%) of the RapidHIT ID instrument 1 runs, 67% of the ladder 49 runs included swabs collected from the same individual. Ladder 54 was also selected in 60 (24%) of the RapidHIT ID instrument 1 runs, where 32% of the ladder 54 runs included swabs collected from the same individual. The native ladder was the most selected ladder and was used in 73 (30%) of the RapidHIT ID instrument 1 runs, with 25% of these runs including swabs collected from the same individual.

Difference in sizes between alleles of the unknown sample and selected ladder (regardless of the ladder chosen) was also used as a performance metric for precision. This metric is included in the output file from the GeneMarker HID software. For the precision study, allele calls removed during manual secondary analysis (i.e., pull-up artifacts and stutter peaks) were considered outliers and removed from averages and statistical calculations, leaving a total of 9,439 allele calls for evaluation. The average difference in sizes between alleles of the unknown sample and selected ladder was 0.048 ($\pm$ 0.060) base pairs (Fig. 24). A total of 9,437 (99.98%) of the allele calls had a difference in allele size of 0.4 base pairs or less. The remaining two allele calls, with differences of 0.6 and 0.77 base pairs, were the previously discussed SE33 allele calls labeled with OB quality flags in the

Concordance Study (Table 2). A total of 8,999 (95%) of the allele calls had a difference in allele size of 0.1 base pairs or less. A per locus representation of Fig. 24 is provided with (Supplementary Fig. 5) and without (Supplementary Fig. 6) outliers. On a per locus basis (without outliers), the average difference in sizes between alleles of the unknown sample and selected ladder ranged from 0.01 ($\pm$ 0.04) base pairs at the D22S1045 locus to 0.14 ($\pm$ 0.06) base pairs at the D8S1179 locus.

Success rate for runs on RapidHIT ID system

First-pass success rate of the RapidHIT ID System runs inform on the reliability and efficacy of the system for use by non-traditional users (e.g., police and border patrol agents). Thus, the first-pass success rate for the entire validation study was calculated. First-pass success was defined as generation of a full profile with zero quality flags during primary analysis by the RapidHIT ID expert system (i.e., without the use of manual secondary review by a qualified/trained scientist). A total of 253 swabs were processed as part of this complete validation study. However, 25 of these swabs (i.e., the mixtures and unused swabs processed as part of the Contamination Study) were expected to generate a "yellow check" after evaluation and, thus, were not included in calculating the study's first-pass success rate. For the remaining 228 swabs, a total of 209 of the swabs
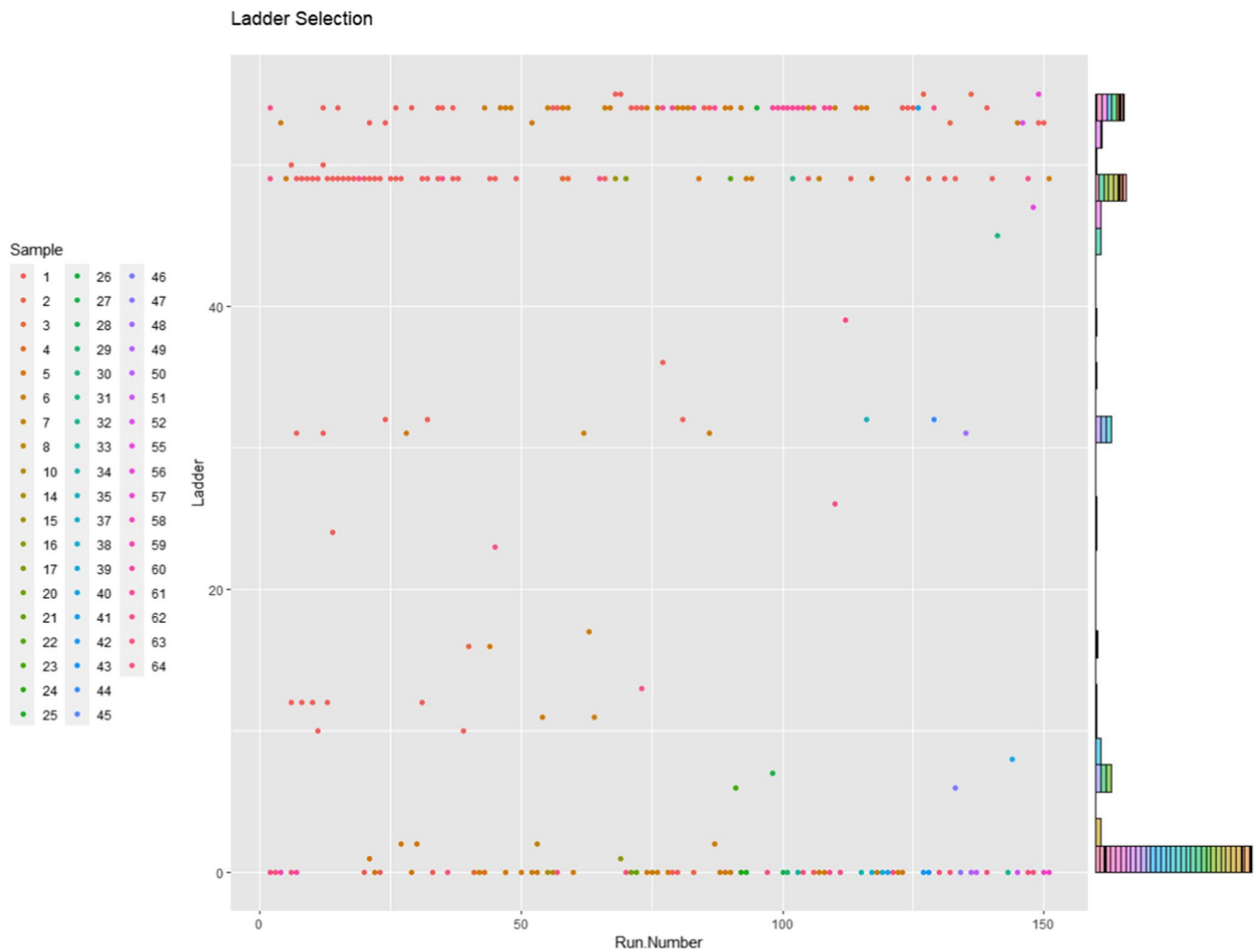
Fig. 23 Scatterplot detailing ladder selection across primary cartridge usage. Histogram details which sample was evaluated for each RapidHIT ID run included in scatterplot

generated a full profile with zero quality flags (or "green check") for a first-pass success rate of 92%. This first-pass success rate includes all the inhibitor, aged, and low-quantity (i.e., low number of swipes) swabs. The first-pass success rate of 92% is likely an underestimate of what laboratories or agencies may expect operationally when analyzing reference buccal swabs.

Each of the 19 swabs that did not generate a full profile with zero quality flags during primary analysis was discussed above. Three of these 19 swabs did not produce a usable profile due to a "size call failed" error (Concordance and Reproducibility and Repeatability Studies). These swabs were immediately re-run in a new sample cartridge, producing a full, quality-flag free profile. The remaining 16 swabs generated profiles with quality flags. Five (each from the Sensitivity Study) of these 16 swabs produced partial profiles. Within these 16 profiles with quality flags, there were 24 flagged allele calls (some received more than one quality flag). These quality flags included: 7 IHO quality flags; 12

IMB quality flags; 3 PL quality flags; 10 OB quality flags; 1 OL quality flag.

Modified Rapid DNA Analysis, or secondary manual review of the STR profiles, allowed for identification of stutter peaks or pull-up artifacts that had been called by the expert system or given a quality flag. Manual review of these "yellow check" profiles resulted in an additional eight profiles with zero quality flags, raising the pass rate to 217/228 or 95%.

Conclusions

The study herein details a series of validation studies for the Applied Biosystems RapidHIT ID Instrument, ACE GlobalFiler Express Sample Cartridge, and RapidLINK software workflow. These validation studies assess the workflow's concordance with traditionally used methodologies, sensitivity, contamination, reproducibility and
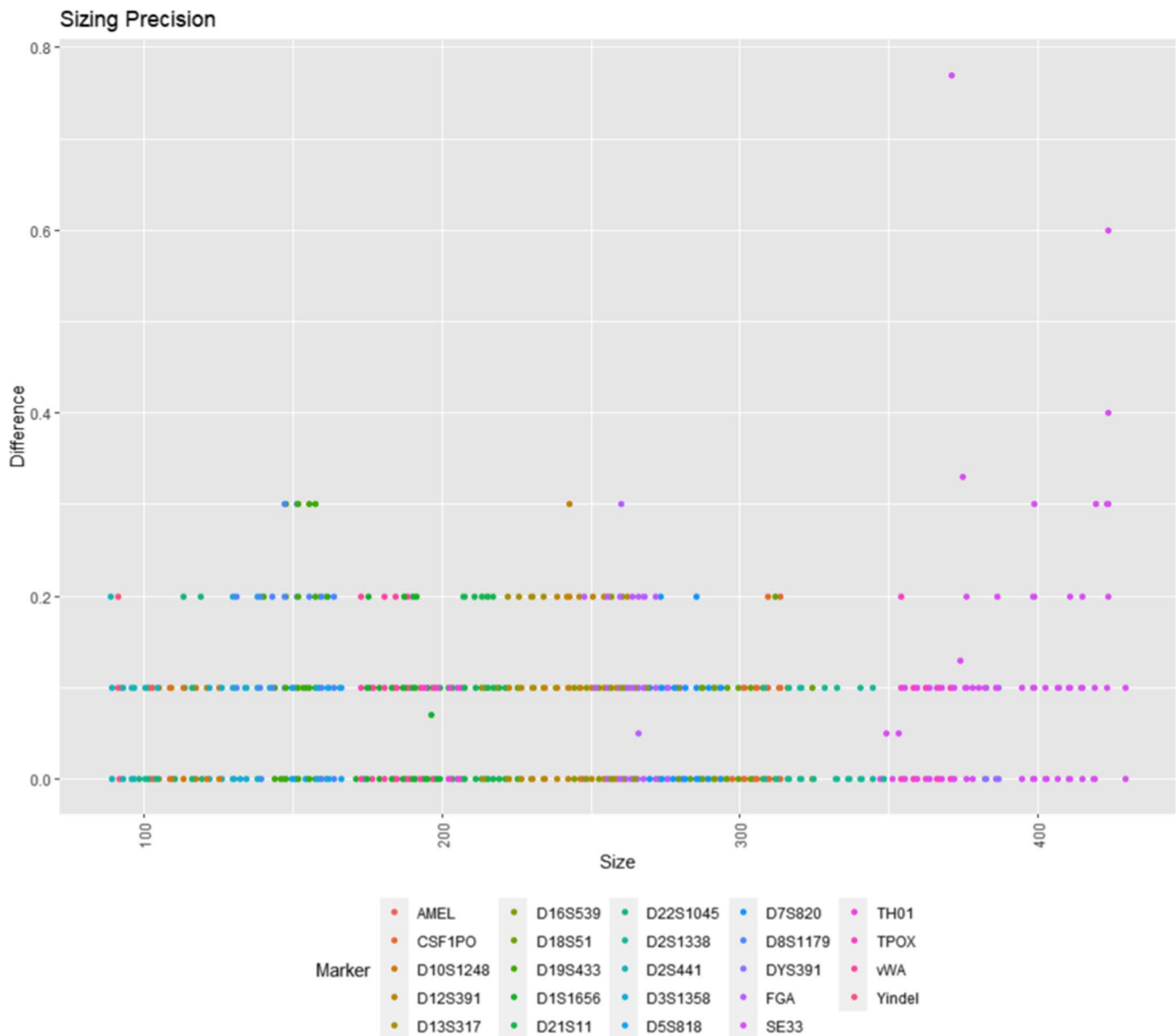
Fig. 24 Scatterplot illustrating difference in sizes between alleles of the unknown sample and selected ladder. Precision study outliers were not included in graph

repeatability, swab stability, effect of common inhibitors, mixture detection, swab re-analysis, and precision. Overall, the data produced in this study illustrate that this workflow generates reliable and reproducible results when evaluating reference buccal swabs.

The data evaluated herein produced a first-pass success rate of 92% with results that did not appear to be affected by common inhibitors or swab age. Collected swabs were able to be analyzed successfully more than once on the RapidHIT ID System or subsequently on the 3500 Genetic Analyzer, and no evidence of run-to-run contamination was detected. These results highlight a flexible and reliable workflow suitable for operators with widely varying backgrounds. However, results from this study also

highlight the benefit of increasing pass rate success if a trained forensic analyst(s) is included in the RapidHIT ID System workflow.

Declarations

Ethics approval The samples were anonymized and collected in accordance with methods approved by the Institutional Review Board for the University of North Texas Health Science Center in Fort Worth, TX, USA.

Conflict of interest The authors declare no competing interests.

References

1. Micka KA, Sprecher CJ, Lins AM, Comey CT, Koons BW, Crouse C, Endean D, Pirelli K, Lee SB, Duda N (1996) Validation of multiplex polymorphic STR amplification sets developed for personal identification applications. J Forensic Sci 41:582–590

2. Moretti T, Baumstark A, Defenbaugh D, Keys K, Smerick J, Budowle B (2001) Validation of short tandem repeats (STRs) for forensic usage: performance testing of fluorescent multiplex STR systems and analysis of authentic and simulated forensic samples. J Forensic Sci 46:647–660

3. Collins PJ, Hennessy LK, Leibelt CS, Roby RK, Reeder DJ, Foxall PA (2004) Developmental validation of a single-tube Amplification of the 13 CODIS STR Loci, D2S1338, D19S433, and amelogenin: the AmpFSTR1 Identifiler1 PCR Amplification Kit. J Forensic Sci 49:1265–1277

4. Ensenberger MG, Thompson J, Hill B, Homick K, Kearney V, Mayntz-Press KA, Mazur P, McGuckian A, Myers J, Raley K, Raley SG, Rothove R, Wilson J, Wieczorek D, Fulmer PM, Storts DR, Krenke BE (2010) Developmental validation of the PowerPlex 16 HS System: an improved 16-locus fluorescent STR multiplex. Forensic Sci Int Genet 4:257–264

5. Kraemer M, Prochnow A, Bussmann M, Scherer M, Peist R, Steffen C (2017) Developmental validation of QIAGEN Investigator 24plex QS Kit and Investigator 24plex GO! Kit: Two 6-dye multiplex assays for the extended CODIS core loci. Forensic Sci Int Genet 29:9–20

6. Ludeman MJ, Zhong C, Mulero JJ, Lagace RE, Hennessy LK, Short ML, Wang DY (2018) Developmental validation of GlobalFiler PCR amplification kit: a 6-dye multiplex assay designed for amplification of casework samples. Int J Legal Med 132:1555–1573

7. Stangegaard M, Frøslev TG, Frank-Hansen R, Hansen AJ, Morling N (2011) Automated extraction of DNA from blood and PCR setup using a Tecan Freedom EVO liquid handler for forensic genetic STR typing of reference samples. J Lab Autom 16:134–140

8. Anslinger K, Bayer B, Rolf B, Keil W, Eisenmenger W (2005) Application of the BioRobot EZ1 in a forensic laboratory. Leg Med 7:164–168

9. Park SJ, Kim JY, Yang YG, Lee SH (2008) Direct STR amplification from whole blood and blood or saliva spotted FTA without DNA purification. J Forensic Sci 53:335–341

10. Flores S, Sun J, King J, Budowle B (2014) Internal validation of the GlobalFiler Express PCR Amplification Kit for the direct amplification of reference DNA samples on a high-throughput automated workflow. Forensic Sci Int Genet 10:33–39

11. Boelens D, Mariot RF, Ghemrawi M, Kloosterman AD, McCord BR (2021) The development of miniSTRs as a method for high-speed direct PCR. Electrophoresis 42:12–13

12. Taylor M (2020) FBI Pilots Rapid DNA at Booking Stations, ID's Challenges for Crime Scenes. Forensic Magazine https://www.forensicmag.com/568259-FBI-Pilots-Rapid-DNA-at-Booking-Stations-ID-s-Challenges-for-Crime-Scenes/. Accessed 19 Jul 2021

13. LaRue BL, Moore A, King JL, Marshall PL, Budowle B (2014) An evaluation of the RapidHIT system for reliably genotyping reference samples. Forensic Sci Int Genet 13:104–111

14. Wiley R, Sage K, LaRue B, Budowle B (2017) Internal validation of the RapidHIT ID system. Forensic Sci Int Genet 31:180–188

15. Holland M, Wendt F (2015) Evaluation of the RapidHIT 200, an automated human identification system for STR analysis of single source samples. Forensic Sci Int Genet 14:76–85

16. Merchant N (2020) US to start collecting DNA from people detained at border. The Associated Press https://apnews.com/article/immigration-tx-state-wire-donald-trump-us-news-8e7d4e3d6e2ef24dcc4fd9e79552a3d0. Accessed 19 Jul 2021

17. Marino M (2021) Investigators using 'rapid DNA testing' to identify victims of Surfside tragedy. Tampa News Channel 8 (Nexstar Media Inc) https://www.wfla.com/news/pinellas-county/investigators-using-rapid-dna-testing-to-identify-victims-of-surfside-tragedy/. Accessed 19 Jul 2021

18. Gin K, Tovar J, Bartelink EJ, Kendell A, Milligan C, Willey P, Wood J, Tan E, Turingan RS, Selden RF (2020) The 2018 California Wildfires: Integration of Rapid DNA to Dramatically Accelerate Victim Identification. J Forensic Sci 65:791–799

19. FBI. Rapid DNA General Information. https://www.fbi.gov/services/laboratory/biometric-analysis/codis/rapid-dna. Accessed 19 Jul 2021

20. Verheij S, Clarisse L, van den Berge M, Sijen T (2013) RapidHIT 200, a promising system for rapid DNA analysis. Forensic Sci Int Genet: Supp Series 4:e254–e255

21. Thong Z, Phua YH, Loo ES, Goh SK, Ang J, Looi WF, Syn CKC (2015) Evaluation of the RapidHIT 200 System: A comparative study of its performance with Maxwell DNA IQ/Identifiler Plus/ABI 3500xl workflow. Forensic Sci Int Genet 19:22–27

22. Shackleton D, Gray N, Ives L, Malson S, Vanhinsbergh D (2019) Development of RapidHIT ID using NGMSElect Express chemistry for the processing of reference samples within the UK criminal justice system. Forensic Sci Int 295:179–188

23. Shackleton D, Pagram J, Ives L, Vanhinsbergh D (2018) Development and validation of the RapidHIT 200 utilising NGMSElect Express for the processing of buccal swabs. Forensic Sci Int 289:244–252

24. Romsos EL, French JL, Smith M, Figarelli V, Harran F, Vandegrift G, Moreno LI, Callaghan TF, Brocato J, Vaidyanathan J, Pedroso JC, Amy A, Stoiloff S, Morillo VH, Czetyrko K, Johnson ED, de Tagyos J, Murray A, Vallone PM (2020) Results of the 2018 Rapid DNA Maturity Assessment. J Forensic Sci 65:953–959

25. Ragazzo M, Melchiorri S, Manzo L, Errichiello V, Puleri G, Nicastro F, Giardina E (2020) Comparative Analysis of ANDE 6C Rapid DNA Analysis System and Traditional Methods. Genes 11:582

26. Murakami C, Irie W, Sasaki C, Nakamaru N, Sakamoto M, Nagato J, Satoh F (2020) Individual identification using the RapidHIT ID system for forensic samples. Leg Med 47:101776

27. Moreno LI, Brown AL, Callaghan TF (2017) Internal validation of the DNAscan/ANDE Rapid DNA Analysis platform and its associated PowerPlex 16 high content DNA biochip cassette for use as an expert system with reference buccal swabs. Forensic Sci Int Genet 29:100–108

28. Manzella AM, Carte KM, King SL, Moreno LI (2021) Assessment of the ANDE 6C Rapid DNA system and investigative biochip for the processing of calcified and muscle tissue. Forensic Sci Int Genet 53:102526

29. Manzella AM, Moreno LI (2020) Assessing the impact of using conventional swabs on the ANDE 6C arrestee biochip. Forensic Sci Int Genet 48:102358

30. Kitayama T, Fukagawa T, Watahiki H, Mita Y, Fujii K, Unuma K, Sakurada K, Uemura K, Sekiguchi K, Mizuno N (2020) Evaluation of Rapid DNA system for buccal swab and disaster victim identification samples. Leg Med 46:101713

31. Hinton N, Springstead T, Wright K, McNamara KP (2021) Evaluation of rapid DNA using ANDE™ in a technical exploitation Level 2 laboratory workflow. J Forensic Sci 66(5):1879–1888. https://doi.org/10.1111/1556-4029.14728

32. Date-Chong M, Hudlow WR, Buoncristiani MR (2016) Evaluation of the RapidHIT 200 and RapidHIT GlobalFiler Express kit for fully automated STR genotyping. Forensic Sci Int Genet 23:1–8

33. Chen A, Yang Y, Yang Q, Tao R, Chen C, Zhang S, Li C (2021) Analytical validation of an RI sample cartridge with the RapidHIT ID system. Int J Legal Med 135:1257–1265

34. Amick GD, Swiger RR (2019) Internal Validation of RapidHIT ID ACE Sample Cartridge and Assessment of the EXT Sample Cartridge. J Forensic Sci 64:857–868

35. Turingan RS, Brown J, Kaplun L, Smith J, Watson J, Boyd DA, Steadman DW, Selden RF (2020) Identification of human remains using Rapid DNA analysis. Int J Legal Med 134:863–872

36. Turingan RS, Tan E, Jiang H, Brown J, Estari Y, Krautz-Peterson G, Selden RF (2020) Developmental Validation of the ANDE 6C System for Rapid DNA Analysis of Forensic Casework and DVI Samples. J Forensic Sci 65:1056–1071

37. Turingan RS, Vasantgadkar S, Palombo L, Hogan C, Jiang H, Tan E, Selden RF (2016) Rapid DNA analysis for automated processing and interpretation of low DNA content samples. Investig Genet 7:2

38. Tan E, Turingan RS, Hogan C, Vasantgadkar S, Palombo L, Schumm JW, Selden RF (2013) Fully integrated, fully automated generation of short tandem repeat profiles. Investig Genet 4:16

39. Manna AD, Nye JV, Carney C, Hammons JS, Mann M, Shamali FA, Vallone PM, Romsos EL, Marne BA, Tan E, Turingan RS, Hogan C, Selden RF, French JL (2016) Developmental validation of the DNAscan Rpaid DNA Analysis instrument and expert system for reference sample processing. Forensic Sci Int Genet 25:145–156

40. Grover R, Jiang H, Turingan RS, French JL, Tan E, Selden RF (2017) FlexPlex27 – highly multiplexed rapid DNA identification for law enforcement, kinship, and military applications. Int J Legal Med 131:1489–1501

41. Carney C, Whitney S, Vaidyanathan J, Persick R, Noel F, Vallone PM, Romsos EL, Tan E, Grover R, Turingan RS, French JL, Selden RF (2019) Developmental validation of the ANDE rapid DNA system with FlexPlex assay for arrestee and reference buccal swab processing and database searching. Forensic Sci Int Genet 40:120–130

42. Salceda S, Barican A, Buscaino J, Goldman B, Klevenberg J, Kuhn M, Lehto D, Lin F, Nguyen P, Park C, Pearson F, Pittaro R, Salodkar S, Schueren R, Smith C, Troup C, Tsou D, Vangbo M, Wunderle J, King D (2017) Validation of a rapid DNA process with the RapidHIT ID system using GlobalFiler Express chemistry, a platform optimized for decentralized testing environments. Forensic Sci Int Genet 28:21–34

43. Jovanovish S, Bogdan G, Belcinski R, Buscaino J, Burgi D, Butts ELR, Chear K, Ciopyk B, Eberhart D, El-Sissi O, Franklin H, Gangano S, Gass J, Harris D, Hennessy L, Kindwall A, King D, Klevenberg J, Li Y, Mehendale N, McIntosh R, Nielsen B, Park C, Pearson F, Schueren R, Stainton N, Troup C, Vallone PM, Vangbo M, Woudenberg T, Wyrick D, Williams S (2015) Developmental validation of a fully integrated sample-to-profile rapid human identification system for processing single-source reference buccal samples. Forensic Sci Int Genet 16:181–194

44. Hennessy LK, Mehendale N, Chear K, Jovanovich S, Williams S, Park C, Gangano S (2014) Developmental validation of the GlobalFiler express kit, a 24-marker STR assay, on the RapidHIT System. Forensic Sci Int Genet 13:247–258

45. Gangano S, Elliott K, Anoruo K, Gass J, Buscaino J, Jovanovich S, Harris D (2013) DNA investigative lead development from blood and saliva samples in less than two hours using the RapidHIT Human DNA Identification System. Forensic Sci Int Genet Supp Series 4:e43–e44

46. Buscaino J, Barican A, Farrales L, Goldman B, Klevenberg J, Kuhn M, Lin F, Nguyen P, Salceda S, Schueren R, Smith C, Troup C, Tsou D, Vangbo M, King D (2018) Evaluation of a rapid DNA process with the RapidHIT ID system using a specialized cartridge for extracted and quantified human DNA. Forensic Sci Int Genet 34:116–127

47. Thermo Fisher Scientific (2019) RapidINTEL Sample Cartridge for blood and saliva samples USER BULLETIN. MAN0018979:Revision A. https://www.thermofisher.com/document-connect/document-connect.html?url=https%3A%2F%2Fassets.thermofisher.com%2FTFS-Assets%2FLSG%2Fmanuals%2FMAN0018979_RapidINTEL_RHIT_v1_1_3_Validation_UB.pdf&title=VXNlciBCdWxsZXRpbjogUmFwaWRJTlRFTCBTYW1wbGUgQ2FydHJpZGdlIGZvciBibG9vZCBhbmQgc2FsaXZhIHNhbXBsZXM=. Accessed 19 Jul 2021

48. Thermo Fisher Scientific (2021) Technical Note: Applied Biosystems RapidHIT ID System v1.3 with RapidHIT Ace GlobalFiler Express Sample Cartridges Validation. Revision A. https://www.thermofisher.com/document-connect/document-connect.html?url=https%3A%2F%2Fassets.thermofisher.com%2FTFS-Assets%2FGSD%2FTechnical-Notes%2Frapidhit-id-validation-technical-note.pdf&title=VGVjaG5pY2FsIE5vdGU6IEFwcGxpZWQgQmlvc3lzdGVtcyBSYXBpZEhJVCBJRCBTeXN0ZW0gdjEuMyB3aXRoIFJhcGlkSElUIEFjZSBHbG9iYWxGaWxlclRNIEV4cHJlc3MgU2FtcGxlIENhcnRyaWRnZXMgVmFsaWRhdGlvbg==. Accessed 19 Jul 2021

49. Scientific Working Group on DNA Analysis Methods (2016) SWGDAM Validation Guidelines for DNA Analysis Methods. https://1ecb9588-ea6f-4feb-971a-73265dbf079c.filesusr.com/ugd/4344b0_813b241e8944497e99b9c45b163b76bd.pdf. Accessed 6/10/2021

50. Thermo Fisher Scientific (2012) PrepFiler ® and PrepFiler ® BTA Forensic DNA Extraction Kits User Manual. 4468126

51. Thermo Fisher Scientific (2017) Quantifiler TM HP and Trio DNA Quantification Kits User Guide. 4485354: G

52. Thermo Fisher Scientific (2014) GlobalFiler Express PCR Amplification Kit User Guide. 4477672: G

53. Thermo Fisher Scientific (2021) RapidHIT ID System v1.3.1 USER GUIDE. MAN0018938:A. https://www.thermofisher.com/document-connect/document-connect.html?url=https%3A%2F%2Fassets.thermofisher.com%2FTFS-Assets%2FLSG%2Fmanuals%2FMAN0018938_RapidHIT_ID_System_v1_3_1_UG.pdf&title=VXNlciBHdWlkZTogUmFwaWRISVQgSUQgU3lzdGVtIHYxLjMuMQ==. Accessed 19 Jul 2021

54. Thermo Fisher Scientific (2018) RapidLINK Software v1.0 USER GUIDE. MAN0018038:A. https://www.thermofisher.com/document-connect/document-connect.html?url=https%3A%2F%2Fassets.thermofisher.com%2FTFS-Assets%2FLSG%2Fmanuals%2FMAN0018038_RapidLinkSW1_UG.pdf&title=VXNlciBHdWlkZTogUmFwaWRMSU5LIFNvZnR3YXJlIHYxLjA=. Accessed 19 Jul 2021