Concordance Assessment of the IDseek® mYSTR™ Reverse Complement PCR Kit for Y-Chromosomal STR Profiling Using the MiSeq® FGx System on a Lebanese Population

Lindsay L. Kotchey MSFS[1], Joop Theelen PhD[2], Pieter van Oers PhD[2], Brian Young PhD[3], Mirna Ghemrawi PhD[1]

[1]The Center for Forensic Science Research and Education, Willow Grove, PA, USA; [2]NimaGen B.V, Nijmegen, Netherlands; [3]NicheVision Forensics LLC, Akron, OH, USA

cfsre | NimaGen. | NicheVision Forensics, LLC

INTRODUCTION

Reverse complement PCR (RC-PCR) is a simple, one-step PCR target enrichment technology that allows for simultaneous amplification and tagging of a targeted sequence in a single tube. Compared to current massively parallel sequencing (MPS) protocols that require multi-step library preparations, including the opening of tubes and samples transfers, RC-PCR reduces labor and contamination risk by combining these steps in a single, combined reaction. The reaction kinetics of the RC-PCR results in high sensitivity and specificity because the specific primers are synthesized during the reaction, so concentrations of primers and amplicons are more in line. One kit that utilizes RC-PCR technology is the IDseek® mYSTR™ Y-Chromosomal STR Profiling kit which amplifies 27 Y-chromosomal STR markers plus Amelogenin in a single multiplex. Beyond traditional STR analysis, the mYSTR™ kit is coupled with MPS so sequence variation within the repeat and flanking regions can be detected in addition to length differences. This research aims to assess the concordance between the mYSTR™ kit with MPS and the Promega PowerPlex® Y23 and AB Yfiler™ Plus PCR amplification system with capillary electrophoresis (CE).

METHODS

Eighty-three (N = 83) unrelated male Lebanese individuals, previously typed with PowerPlex® Y23 and AB Yfiler™ Plus PCR systems, were sequenced using the IDseek® mYSTR™ Y-Chromosomal STR Profiling kit.

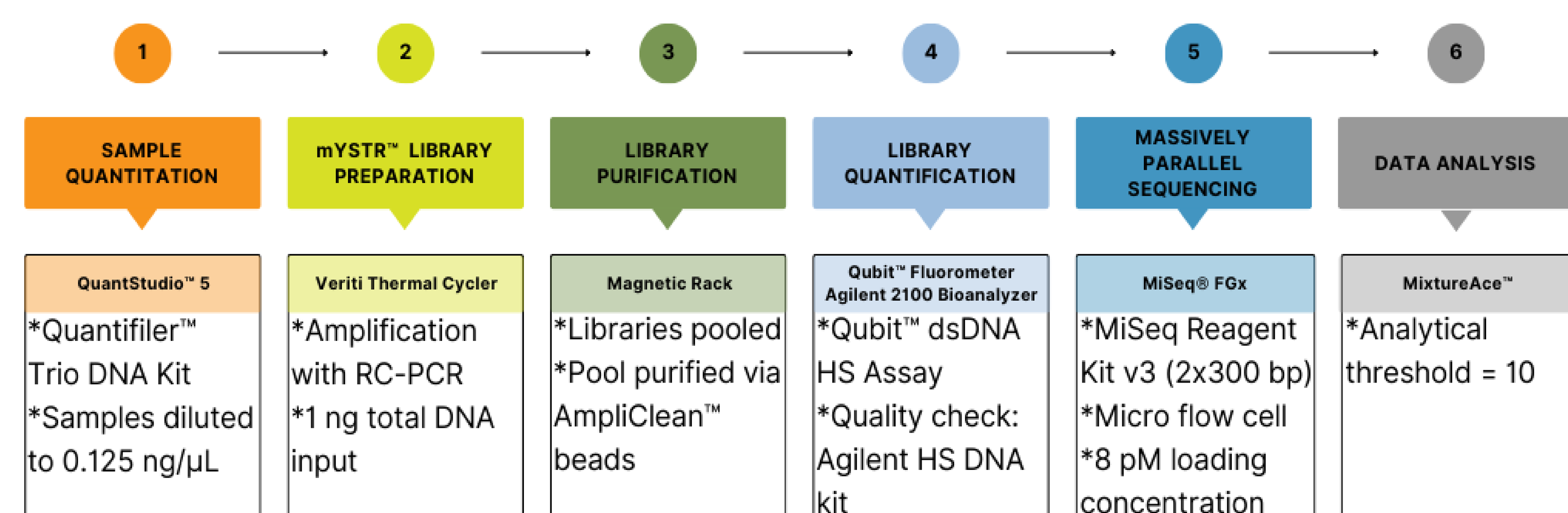
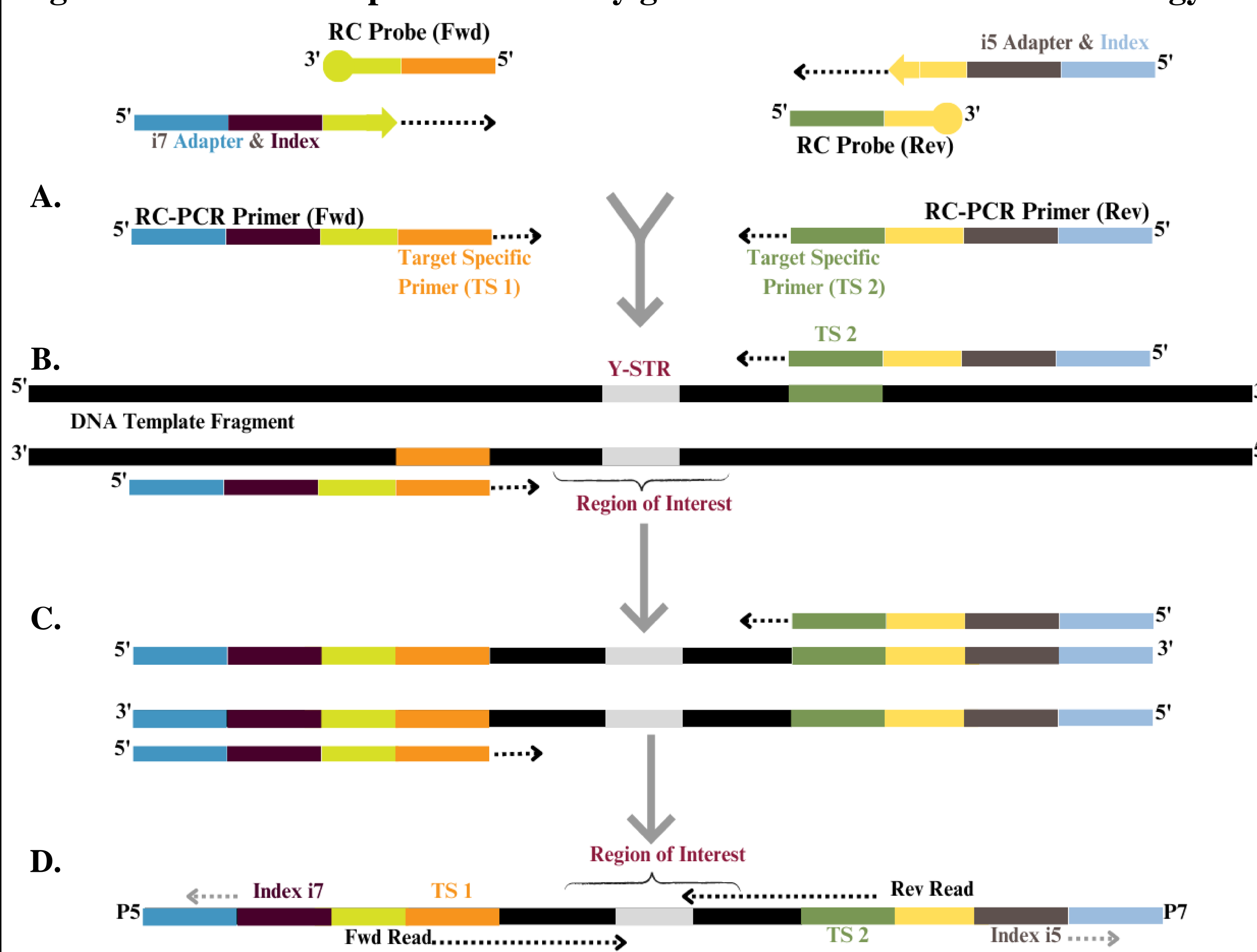


Figure 1. Schematic depiction of library generation with RC-PCR technology.



RESULTS

Table 1. MiSeq® FGx Sequencing Metrics

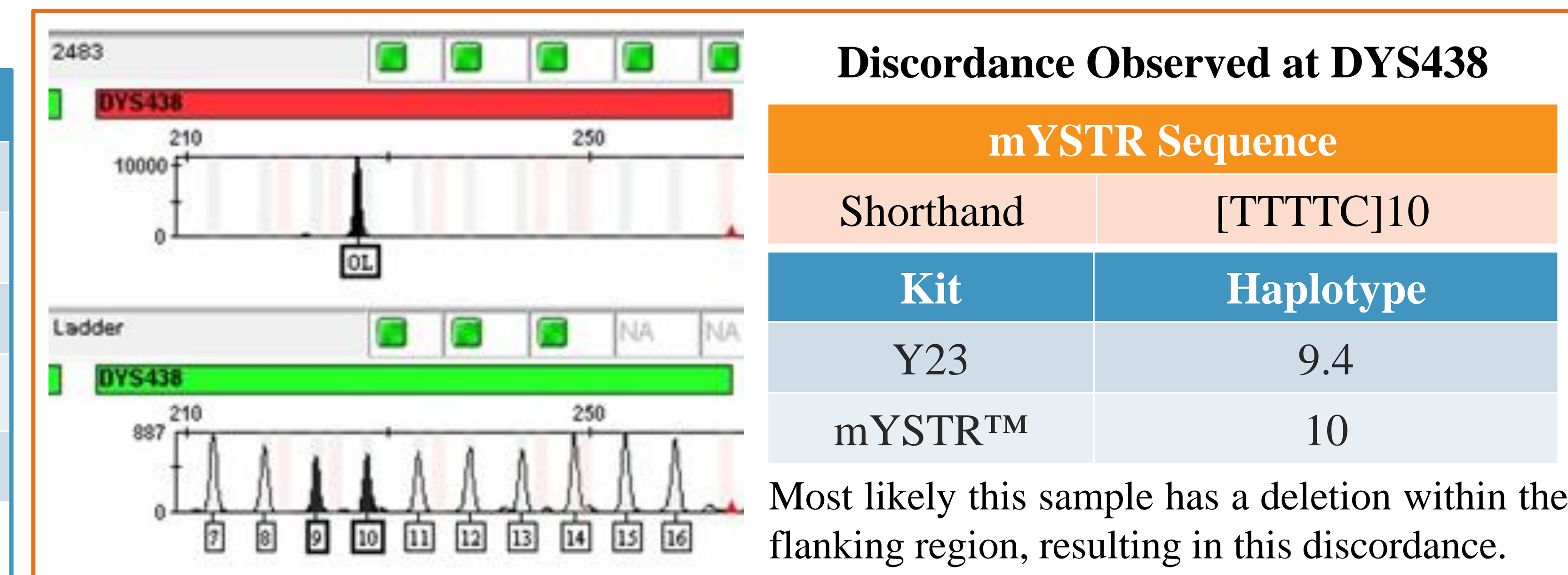
Description	Metric
Cluster Density	831 (K/mm2)
Aligned	29.43%
Error Rate	4.01%
% ≥Q30	62.7
Total Reads	4,298,916
Total Reads Passing Filter	2,960,772
% Reads Identified Passing Filter	60.5892%

The flow cell was within the desired cluster density range. PhiX was spiked in at 25% to allow for sequence variation.

Table 2. Occurrence of Dropout

LOCUS	OCCURRENCE
AMEL-X	11
AMEL-Y	10
DYS389II	6
DYS387S1a/b	1
DYS392	1

Dropout occurred a total of 29 times across all loci and samples. AMEL (X/Y) was the most dropout locus. Dropout at DYS387S1a/b and DYS392 occurred in the same sample.

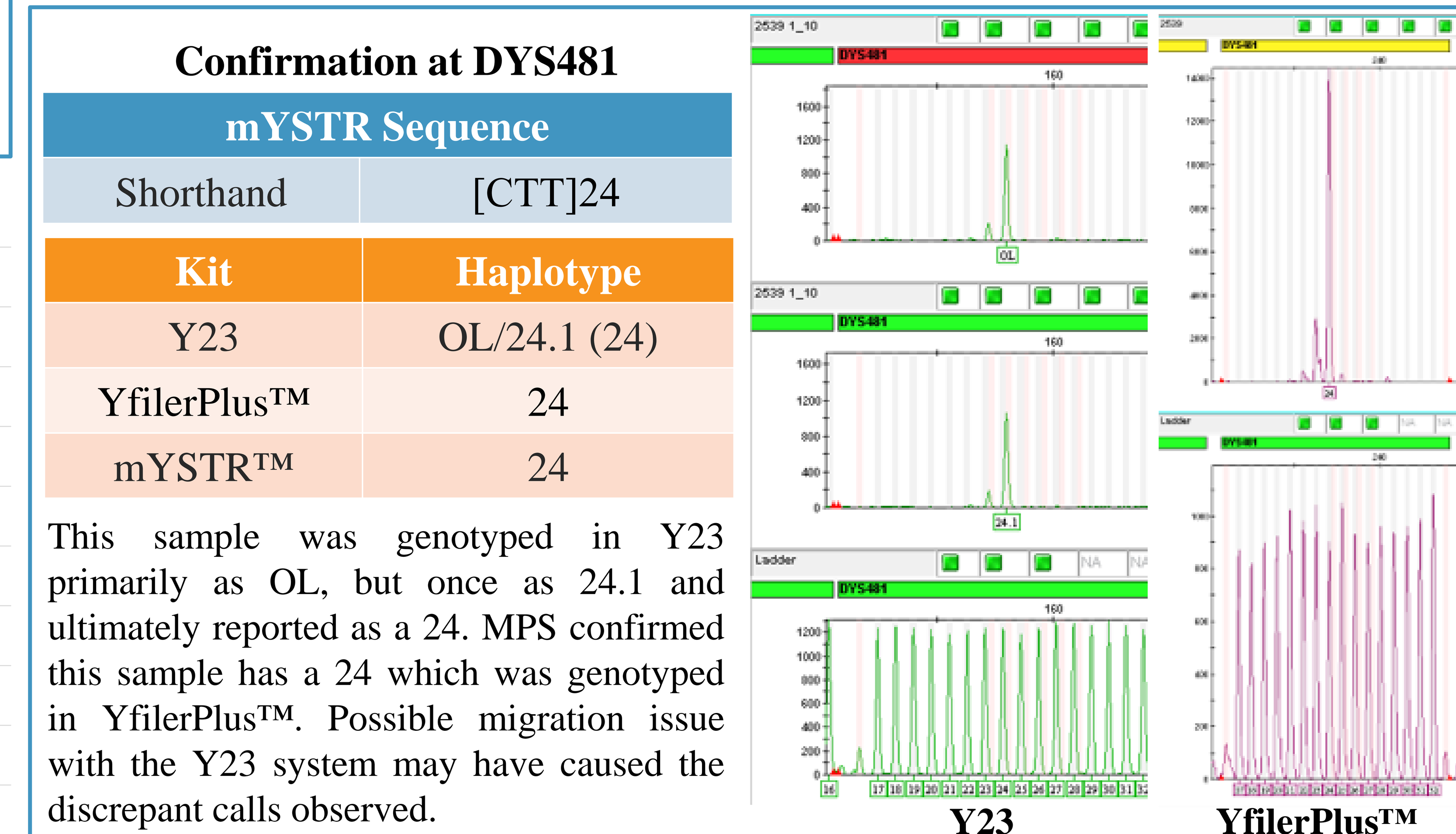


Discordance Observed at DYS438

mYSTR Sequence	
Shorthand	[TTTTC]10

Kit	Haplotype
Y23	9.4
mYSTR™	10

Most likely this sample has a deletion within the flanking region, resulting in this discordance.



Confirmation at DYS481

mYSTR Sequence	
Shorthand	[CTT]24

Kit	Haplotype
Y23	OL/24.1 (24)
YfilerPlus™	24
mYSTR™	24

This sample was genotyped in Y23 primarily as OL, but once as 24.1 and ultimately reported as a 24. MPS confirmed this sample has a 24 which was genotyped in YfilerPlus™. Possible migration issue with the Y23 system may have caused the discrepant calls observed.

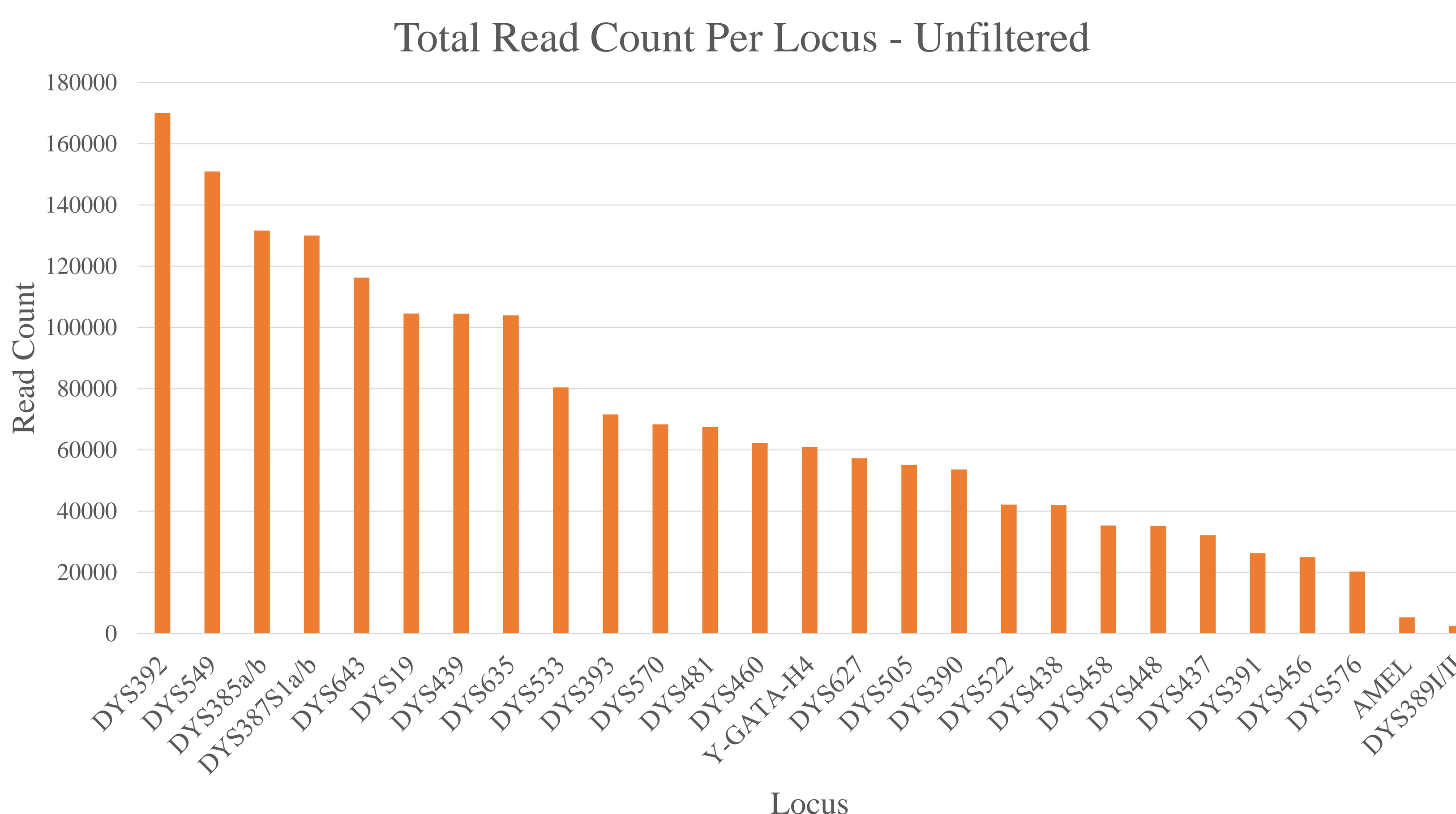


Figure 2. Comparison of total unfiltered read count per locus. Loci with the fewest reads were AMEL and DYS389I/II, correlating with the highest rate of dropout at these loci. The average total read count per sample was 20,795 reads.

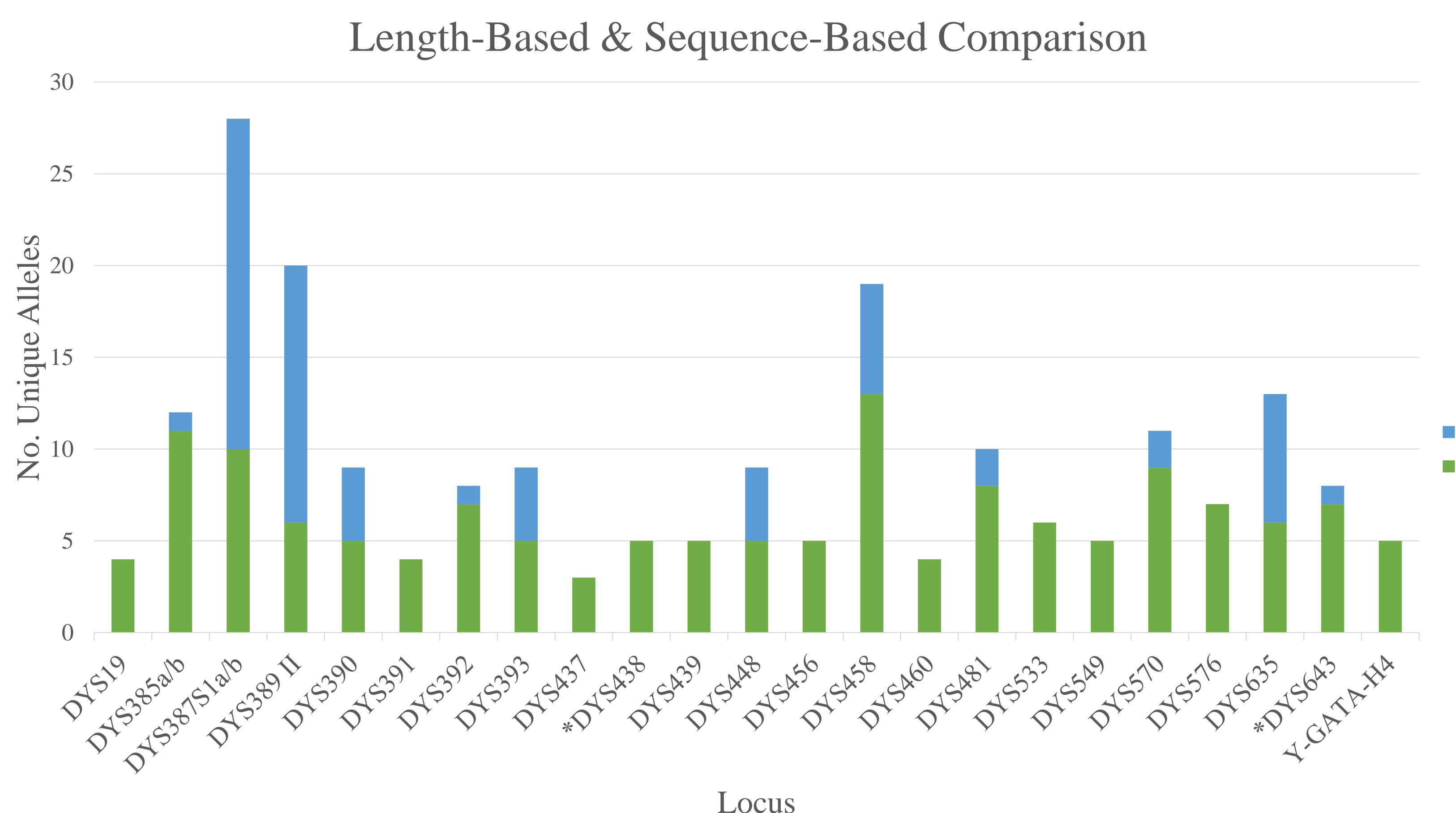


Figure 3. Comparison of alleles observed with length- or sequence-based. The highest gain in unique alleles was observed in DYS387S1a/b dividing the 10 alleles by size into 28 alleles by sequence. *DYS438 & DYS643 observed one more unique allele in the LB, however, it is believed that this is a result of an indel in the flanking regions of these samples.

CONCLUSIONS & FUTURE WORK

- Results demonstrated high concordance between length- and sequence-based data. Two discordant alleles were received from two different markers (DYS438 and DYS643) and are most likely affected by indels within the flanking regions.
- Stutter filters will be determined based on average read counts obtained.
- Flanking regions will be included in the analysis range.
- The benefits of MPS were shown with this study by highlighting the increase in diversity of STR alleles by variation in the repeat as obtained via sequencing.
- The IDseek® mYSTR™ has shown to be an effective Y-STR amplification kit with minimized hands-on laboratory time with the RC-PCR technology.

ACKNOWLEDGEMENTS

- DNA extracts were graciously provided by Dr. Susan Walsh (IUPUI) in accordance with the American University of Beirut (ALM.RM.27) and the Indiana University-Purdue University Indianapolis Institutional Review Board (study number 1409306349).
- Many thanks to Dr. Brian Young and Jeff Smith of NicheVision Forensic LLC for their technical support with the MixtureAce™ software and the mYSTR plugin.
- Reagents and materials were kindly provided by NimaGen, as well as technical guidance.
- The authors would like to thank the staff of the CFSRE for their support.

REFERENCES

Kieser, R. E. *et al.* Reverse Complement PCR: A novel one-step PCR system for typing highly degraded DNA for human identification. *Forensic Science International: Genetics* **44**, 102201 (2020).

Bus, M. M. *et al.* Reverse complement-PCR, an innovative and effective method for multiplexing forensically relevant single nucleotide polymorphism marker systems. *BioTechniques* **71**, 484–489 (2021).