EVALUATION OF PORTABLE NANOPORE STR SEQUENCING COUPLED TO A CUSTOMIZED CLOUD-ENABLED DATA ANALYSIS PLATFORM

Christopher Williams, Sarah Bailey, Seth A. Faith, Department of Molecular Biomedical Sciences and the Forensic Sciences Institute, North Carolina State University

Nanopore technologies, alongside Cloud-based bioinformatics, have shown direct application for field-based genetic typing that may have benefit for forensic crime scene analysis. Here we present a study to evaluate one nanopore technology, Oxford Nanopore Technologies' MinION, for sequence-based analysis of Short Tandem Repeats (STRs). MinION sequencing uses nanopore-sized channels to distinguish alterations in electric current as DNA sequences are shuttled through. In this study a male standard reference material, 2391c (Component B), from the National Institute of Standards and Technologies (NIST) was PCR-enriched using a prototype of the PowerSeq^{TM} AUTO/Y system (Promega Corp.). Library preparation was performed for short fragment 2d analysis with the manufacturer's Genomic DNA protocol, designed specifically for the R9/R9.4 SpotON flow cell. Data analysis was conducted using both the vendor's primary sequence analysis tool (Metrichor) and our custom-designed, Cloud-based STR sequence processing tool, *Altius,* to produce STR match results. Nearly 10,000 sequence reads passed quality filters during the runs within ONT's cloud-based Metrichor Agent. These high-quality reads were then run through *Altius* and data were compared to the published NIST SRM 2391c certificate to confirm accuracy. Over 45% of raw reads matched STR loci from the Auto/Y STR panel, with sufficient depth of coverage for 30 of the 43 amplified Autosomal STRs and 12 of the 23 amplified Y-STRs. Over the two sequencing runs, the loci with the highest level of coverage and accuracy included: CSF1PO, D2S441, D3S1358, D7S820, D16S539, PentaE, TH01 and vWA. Sequence repeat motif and composition did not show a strong correlation to sequence error and/or STR mismatch when analyzed by CLC Bioinformatics Workbench. The highest indicator of accuracy for STR genotyping was depth of coverage, which was a potential challenge in the 46-locus multiplex given the current output of the nanopore device. Lastly, the cloud-based platform allowed for rapid STR typing and database ready genotypes that could be used for searching if a database should be made available to the cloud platform. This evaluation demonstrated viable STR typing by a portable technology, the MinION, supplemented by a Cloud-based environment for remote connectivity and data processing.