Validation of a Massively Parallel Sequencing Workflow for Mitochondrial DNA Analysis at UNTHSC Center for Human Identification for Missing Persons and Traditional Casework Analyses

Jennifer D. Churchill[1], Maiko Takahashi[1], Christina Strobl[2], Dixie Peters[1], Christina Capt[1], Walther Parson[2,3], Bruce Budowle[1]

[1]Center for Human Identification, University of North Texas Health Science Center, Fort Worth, TX, USA
[2]Institute of Legal Medicine, Medical University of Innsbruck, Innsbruck, Austria
[3]Forensic Science Program, The Pennsylvania State University, PA, USA

Corresponding Author: Jennifer D. Churchill
3500 Camp Bowie Blvd, CBH-247
Fort Worth, TX 76107
1-817-735-2912
Jennifer.Churchill@unthsc.edu

Introduction

A number of studies have evaluated the applicability of massively parallel sequencing (MPS) technologies to analyze forensic biological evidence and the benefits afforded by these technologies. For mitochondrial DNA (mtDNA) sequencing, MPS technologies now make it feasible for forensic laboratories to analyze the entire mitochondrial genome (mtGenome) with increasingly streamlined and automated workflows, enhanced data analysis options, and the commercial availability of whole genome multiplex panels designed for challenged and degraded samples. Expanding analysis to the entire mtGenome offers the opportunity for an increase in discrimination power and phylogenetic resolution in mtDNA results, and the quantitative nature of MPS technologies brings new opportunities for mixture deconvolution in mtDNA analysis. As such, analysis of the mtGenome with MPS technologies can serve as the first step in transitioning from capillary electrophoresis-based to MPS-based technologies in forensic laboratories.

Materials and Methods

Library Preparation and Massively Parallel Sequencing

The mitochondrial genome was amplified with the Precision ID mtDNA Whole Genome Panel (Thermo Fisher Scientific) following the manufacturer's recommended protocols [1]. Libraries were prepared both manually with the Precision ID Library Kit (Thermo Fisher Scientific) and in an automated fashion on the Ion Chef (Thermo Fisher Scientific) following manufacturer's protocols [1]. Template preparation was completed on the Ion Chef, and sequencing was performed on the Ion S5 (Thermo Fisher Scientific) following manufacturer's protocols [1].

Data Analysis

Raw sequence data were analyzed using the Torrent Suite software. Data were aligned to an rCRS+80 reference genome to accommodate the multiplex's design [1-2]. Variant calls were generated with the Variant Caller plugin, and the variant call format (VCF) output files generated by the Variant Caller plugin were used in mitoSAVE [3] to generate haplotype calls in standard forensic nomenclature. Integrative Genomic Viewer (IGV) was used for a visual check of the aligned binary alignment map (BAM) files [4-5]. Finally, a phylogenetic check of the haplotype calls was performed in Haplogrep and Empop [6-7]. Performance metrics, including read depth, strand balance, and noise, were used to evaluate the quality of the sequencing results generated.

Results and Discussion

The Missing Persons and Forensic Units at UNTHSC's Center for Human Identification (UNTCHI) have begun the implementation process for mtDNA testing of biological evidence using a MPS workflow. Forensic analysts' experience from an extensive training program was used to develop SOPs and workflow and throughput considerations. Finally, validation studies were performed. Sensitivity and stochastic studies demonstrated the dynamic range and limit of detection with samples ranging from 300 pg to 2 pg of DNA. MPS-specific studies also addressed the limit of detection by evaluating the amount of library input and extent of sample multiplexing. Studies of reproducibility and repeatability were completed with multiple analysts and multiple instruments. The contamination assessment evaluated blanks and known haplotypes for evidence of exogenous DNA. Known and non-probative evidence samples, including family reference samples, human remains, and hairs, were sequenced and compared to previously generated Sanger sequencing results to determine performance with potentially challenging samples. Mixtures of ratios ranging from 1:2 to 1:20 were sequenced to evaluate the capability of the system to detect and resolve mixtures.

Conclusions

The continued development of massively parallel sequencing technologies supported by numerous studies throughout the genetic and forensic communities has made implementation of this technology into forensic genetic crime laboratories feasible. Data from these studies support that this MPS workflow yields reliable results for the analysis of biological evidence. Finally, our experiences and resources could assist other forensic laboratories considering implementation of massively parallel sequencing, and our procedures will be made available to others upon the completion of this study.

Acknowledgements

This work was supported in part by award no. 2016-DN-BX-K001, awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. The opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect those of the U.S. Department of Justice.

References

[1] Thermo Fisher Scientific. Precision ID Panels with Ion S5 System Application Guide. Revision B. Thermo Fisher Scientific, Waltham, MA, USA. (2017).

[2] Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat. Genet. 23 (1999) 147.

[3] King JL, Sajantila A, Budowle B. mitoSAVE: mitochondrial sequence analysis of variants in Excel. Forensic Sci. Int. Genet. 12 (2014) 122–5.

[4] Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief. Bioinform. 14 (2013) 178–92.

[5] Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP Integrative genomics viewer. Nat. Biotechnol. 29 (2011) 24–6.

[6] Kloss-Brandstatter A, Pacher D, Schonherr S, Weissensteiner H, Binna R, Specht G, Kronenberg F. HaploGrep: a fast and reliable algorithm for automatic classification of mitochondrial DNA haplogroups, Hum. Mutat. 32 (2011) 25–32.

[7] Parson W, Dur A. EMPOP-A forensic mtDNA database. Forensic Sci. Int. Genet. 1 (2007) 88-92.