MITOCHONDRIAL DNA SEQUENCING USING THE PRECISION ID NGS SYSTEM AND CONVERGE ANALYSIS SOFTWARE: A ROBUST AND SENSITIVE TILED AMPLICON ASSAY FOR FORENSIC CASEWORK APPLICATIONS

R. Hasegawa1, S. Wootton1, C. Roth1, CW. Chang1; Vijaychander, S.1, Deng. J.1, R. Lagacé1
1Thermo Fisher Scientific, South San Francisco, CA 94080

ABSTRACT

In forensic casework, mitochondrial DNA (mtDNA) is useful in the context of recalcitrant samples that fail to produce a standard STR profile. Traditional Sanger sequencing using capillary electrophoresis (CE) compiles a limitation of sequencing the mtDNA genome to the hypervariable region as sequencing of the whole mitochondrial genome (mtGenome) is both time consuming and cost-prohibitive. With the availability of massively parallel sequencing systems (MPS), the mtGenome can easily be prepared and sequenced using a tiled amplicon multiplex of 162 amplicons. Additionally, the forensic mtDNA analysis module developed on Converge™ Software and optimized specifically for the Precision ID Control Region and Whole Genome panels provides streamlined analysis for haplotype and haplogroup designations as well as robust detection of nuclear mitochondrial DNA segments (NUMTs) and point and length heteroplasmies. DNA from samples with known haplotypes were obtained through Corell and NIST. Libraries were prepared on the Ion Chef using the Precision ID mtDNA Control and Whole Genome Panel and sequenced on the Ion S5. Reads generated on the system were aligned and filtered using the CrS alignment algorithm that integrates phytophore and EMPOP information into the scoring function. Variants were called with reference to the Crs. Additionally, the closest haplogroup was calculated, and variants were evaluated based on the presence of other general metrics including frequency, strand bias, and coverage.

RESULTS

The panel showed high uniformity and efficiency for both the whole genome and control region panels (Table 3). Compared to Sanger (Table 4), variant calling with SRR 23922 samples showed little discordance, limited mostly to the 309 position. Sample normalized coverage across all the amplicons in the mtGenome (Figure 4) and also sorted coverage (Figure 5) shows amplicon uniformity across the mtGenome.

CONCLUSIONS

Control region and whole genome mtDNA sequencing can be applied in cases where DNA is limited or has been degraded. Sequencing of the panel with the Precision ID mtDNA panel on SS using the mtDNA analysis module on Converge™ showed high concordance with Sanger, primarily with exception of the 309 position, where Converge™ marked the call as “unclear.” Heteroplasmies could be detected at a rate similar to previously published MPS studies, and PhP and LHP thresholds could be lowered as a custom setting to detect lower level heteroplasmies not detected here. Sequencing of samples from 100pg of starting input down to 2pg showed remarkably little difference in coverage uniformity, dropouts, and variant calling performance. In a practical casework laboratory, the Precision ID system can be implemented in routine missing persons / disaster victim identification workflows as a robust mtDNA analysis option in cases where STRs have failed. The benefits of MPS testing relative to standard HV region sequencing using Sanger methods (e.g. increased discrimination with mtGenome sequencing, improved heteroplasmie detection and overall system sensitivity with limited-and/or degraded DNA) for mtDNA analysis offers benefits for missing persons identification.

REFERENCES


TRADEMARKS/LICENSEING

For Research, Forensics, or Paternity Use Only. Not for use in diagnostic procedures.

Table 3. Sensitivity of mtDNA panel

<table>
<thead>
<tr>
<th>Sample</th>
<th>Input (pg)</th>
<th>Uniformity</th>
<th>PhP Amplified</th>
<th>Number of PhP</th>
</tr>
</thead>
<tbody>
<tr>
<td>HG1389</td>
<td>100</td>
<td>99.96</td>
<td>91.12</td>
<td>100</td>
</tr>
<tr>
<td>HG2392</td>
<td>100</td>
<td>99.97</td>
<td>91.12</td>
<td>100</td>
</tr>
<tr>
<td>HG1389</td>
<td>10</td>
<td>99.96</td>
<td>91.12</td>
<td>100</td>
</tr>
</tbody>
</table>

*Each input level was sequenced in quadruplicate and the average of the four runs is presented in the table. Input is the gDNA equ. for the sample.

Figure 5. Performance of the panel and system at 100pg (a), 10pg (b), 5pg (c), and 2pg (d) gDNA input.

CONCLUSIONS

Control region and whole genome mtDNA sequencing can be applied in cases where DNA is limited or has been degraded. Sequencing of the panel with the Precision ID mtDNA panel on SS using the mtDNA analysis module on Converge™ showed high concordance with Sanger, primarily with exception of the 309 position, where Converge™ marked the call as “unclear.” Heteroplasmies could be detected at a rate similar to previously published MPS studies, and PhP and LHP thresholds could be lowered as a custom setting to detect lower level heteroplasmies not detected here. Sequencing of samples from 100pg of starting input down to 2pg showed remarkably little difference in coverage uniformity, dropouts, and variant calling performance. In a practical casework laboratory, the Precision ID system can be implemented in routine missing persons / disaster victim identification workflows as a robust mtDNA analysis option in cases where STRs have failed. The benefits of MPS testing relative to standard HV region sequencing using Sanger methods (e.g. increased discrimination with mtGenome sequencing, improved heteroplasmie detection and overall system sensitivity with limited-and/or degraded DNA) for mtDNA analysis offers benefits for missing persons identification.

REFERENCES


TRADEMARKS/LICENSEING

For Research, Forensics, or Paternity Use Only. Not for use in diagnostic procedures.