The Journal of Molecular Diagnostics, Vol. 15, No. 4, July 2013

the Journal of Molecular Diagnostics
jmd.amjpathol.org

The Development of Next-Generation Sequencing Assays for the Mitochondrial Genome and 108 Nuclear Genes Associated with Mitochondrial Disorders

Shale Dames,*† Lan-Szu Chou,‡ Ye Xiao,* Tyler Wayman,* Jennifer Stocks,* Marc Singleton,§ Karen Eilbeck,† and Rong Mao*¶

From the Institute for Clinical and Experimental Pathology, ARUP Laboratories, Salt Lake City, Utah; the Departments of Biomedical Informatics,† Human Genetics,§ and Pathology,¶ University of Utah School of Medicine, Salt Lake City, Utah; and the Department of Bioagriculture,‡ National Chiayi University, Chiayi, Taiwan*

CME Accreditation Statement: This activity ("JMD 2013 CME Program in Molecular Diagnostics") has been planned and implemented in accordance with the Essential Areas and policies of the Accreditation Council for Continuing Medical Education (ACCME) through the joint sponsorship of the American Society for Clinical Pathology (ASCP) and the American Society for Investigative Pathology (ASIP). ASCP is accredited by the ACCME to provide continuing medical education for physicians.

The ASCP designates this journal-based CME activity ("JMD 2013 CME Program in Molecular Diagnostics") for a maximum of 48 AMA PRA Category 1 Credit(s)™. Physicians should only claim credit commensurate with the extent of their participation in the activity.

CME Disclosures: The authors of this article and the planning committee members and staff have no relevant financial relationships with commercial interests to disclose.

Accepted for publication March 15, 2013.

Address correspondence to Shale Dames, M.S., ARUP Labs, Molecular Genetics 289, 500 Chipeta Way, Salt Lake City, UT 84108. E-mail: shale.dames@aruplab.com.

Sanger sequencing of multigenic disorders can be technically challenging, time consuming, and prohibitively expensive. High-throughput next-generation sequencing (NGS) can provide a cost-effective method for sequencing targeted genes associated with multigenic disorders. We have developed a NGS clinical targeted gene assay for the mitochondrial genome and for 108 selected nuclear genes associated with mitochondrial disorders. Mitochondrial disorders have a reported incidence of 1 in 5000 live births, encompass a broad range of phenotypes, and are attributed to mutations in the mitochondrial and nuclear genomes. Approximately 20% of mitochondrial disorders result from mutations in mtDNA, with the remaining 80% found in nuclear genes that affect mtDNA levels or mitochondrion protein assembly. In our NGS approach, the 16,569-bp mtDNA is enriched by long-range PCR and the 108 nuclear genes (which represent 1301 amplicons and 680 kb) are enriched by RainDance emulsion PCR. Sequencing is performed on Illumina HiSeq 2000 or MiSeq platforms, and bioinformatics analysis is performed using commercial and in-house developed bioinformatics pipelines. A total of 16 validation and 13 clinical samples were examined. All previously reported variants associated with mitochondrial disorders were found in validation samples, and 5 of the 13 clinical samples were found to have mutations associated with mitochondrial disorders in either the mitochondrial genome or the 108 nuclear genes. All variants were confirmed by Sanger sequencing. *(J Mol Diagn 2013, 15: 526—534; http://dx.doi.org/10.1016/j.jmoldx.2013.03.005)*

Mitochondrial disorders genetically fall into two classes: mutations in the mitochondrial genome (mtDNA) and genes in the human nuclear genome. mtDNA is a maternally inherited, circular, 16,569-bp haploid organelle composed of 37 genes. It is estimated that up to 1500 nuclear genes may be associated with nuclear encoded mitochondrial proteins, and inheritance may be autosomal recessive, dominant, or sex-linked.[1—3] Mitochondrial disorders have an overall incidence of 1:5000, with approximately 20% of disease-causing variants found in the mtDNA and the remainder in the nuclear genome.[4,5] Common characteristics of mitochondrial disorders include ptosis, external ophthalmoplegia, proximal myopathy, exercise intolerance, cardiomyopathy, sensorineural deafness, optic atrophy, pigmentary retinopathy, diabetes mellitus, and mid- to late-term pregnancy loss.[6—9] Central nervous system findings may include fluctuating encephalopathy, seizures, dementia, migraine,

Supported by the ARUP Laboratories Institute for Clinical and Experimental Pathology.

http://dx.doi.org/10.1016/j.jmoldx.2013.03.005

stroke-like episodes, ataxia, and spasticity.[10–14] Disease onset ranges from childhood to adulthood. Because mitochondrial disorders encompass a wide range of phenotypes and a large number of genes, high-throughout next-generation sequencing (NGS) is an ideal method for variant detection.

NGS allows for a low-cost, comprehensive mitochondrial disorder panel, one that would be financially and technically difficult to perform by Sanger sequencing. Furthermore, the ability to sequence at high coverage allows for detection of low-level heteroplasmy for mtDNA mutations, which are not easily detected by Sanger sequencing. For the present study, 108 nuclear genes were chosen (Supplemental Table S1), based on known genes implicated in mitochondrial disorders at the time of assay design. The 108 sequenced nuclear genes include genes required for mtDNA integrity, fatty acid metabolism, transcription, mitochondria maintenance, and other oxidative phosphorylation–associated genes.

NGS has been used for detection of mitochondrial disorders.[15–20] Enrichment techniques include long-range PCR (LR-PCR) for the mitochondrial genome and various in-solution and chip-based capture methods for nuclear genes. Enrichment using RainDance emulsion PCR libraries (RainDance Technologies, Lexington, MA) has not previously been reported for mitochondrial disorders. RainDance emulsion PCR has been described for the enrichment of genes associated with aortopathies, glycosylation disorders, and selected genes on the X chromosome, among others.[21–23] RainDance libraries are synthesized with single primer pairs encapsulated into an emulsion droplet. Approximately 1 million droplets containing target-specific, individual PCR primer sets are used for enrichment. A PCR amplification mix with target DNA is subsequently emulsified and fused (or merged) with an individual RainDance emulsion droplet. The merged RainDance library/PCR mix is deposited into a PCR tube and amplified.[24] After amplification, the PCR fragments are concatenated and sonicated to obtain the desired length inserts for NGS libraries. Without this concatenation step, PCR fragments longer than the read lengths associated with a given NGS platform would not have internal sequence information. The main advantage of PCR enrichment methods (such as LR-PCR and RainDance) revolves around their greater specificity, compared with capture-based methods.[3,25] As with previously reported mtDNA assays, LR-PCR was chosen for enrichment based on ease, reproducibility, and cost. In the present study, we assessed the performance of LR-PCR, RainDance enrichment, and NGS for detection of mutations associated with mitochondrial disorders. We also addressed differences in bioinformatics analysis for mtDNA and RainDance-enriched genes.

Materials and Methods

mtDNA and Nuclear Gene Targeted Regions

All 37 genes of the mtDNA were sequenced. The 108 nuclear genes were included as all nuclear genes known (at the time of assay design) to be involved in mitochondrial disorders.[26,27]

mtDNA Validation Samples

Seven samples with known mtDNA mutations and two normal samples were obtained from the Coriell Cell Repositories at the Coriell Institute for Medical Research (Camden, NJ) and were used for validation of the mtDNA NGS assay. The seven mutation samples were NA10742 MT-ND4, m.11778G>A p.Arg340His; NA10744 MT-ND4, m.11778G>A p.Arg340His; NA11605, MT-ND1, m.3460G>A p.Ala52Thr; NA11906, m.8344A>G tRNA-Lys/wt; NA13411, MT-ATP6, m.8993T>G p.Leu156Arg; NA13740, MT-ATP6, m.8993T>G p.Leu156Arg; and NA13741, MT-ATP6, m.8993T>G p.Leu156Arg. The two other samples, NA12878 and NA19240, had no reported clinically associated mtDNA variants and were used as a normal control.

108 Nuclear Gene Panel Validation Samples

Three Coriell samples with no reported deleterious variants in the 108 nuclear genes examined were used as normal controls [NA11605 (which harbors a mtDNA mutation in MT-ND1), NA12878, and NA19240]. Four research samples (R459, R460, R461, and R463) were used for nuclear gene panel validation. The four research samples were shown to have no mtDNA mutations by Sanger sequencing. Two of these samples (R461 and R463) had reported mutations in genes targeted in the nuclear panel; the other two samples (R459 and R460) had no known mutations. All research samples were deidentified before enrichment and NGS sequencing.

Clinical Samples

Clinical samples were deidentified according to the ARUP Laboratories institutional review board protocol (no. 7725). A total of 13 clinical samples were processed for mitochondrial disorder assay (Table 1). Six mtDNA samples, five 108-gene panel samples, and two samples with both mtDNA and the 108-gene panel were sequenced.

Oligonucleotides

Primer sequences for mtDNA LR-PCR were as listed in Table 2. Two sets of primers were used to independently amplify the mitochondrial genome in duplicate: Mito1, Mito2, and Mito3 primers and hmt1 and hmt2 primers (Figure 1). The two sets of primers were used to avoid allele dropouts. However, because of the heteroplasmy levels associated with mtDNA, this does not ensure that all mtDNA targeted regions are amplified. Primers were synthesized by Integrated DNA Technologies (Coralville, IA) with a 5′ amino C6 modification to avoid sequence overrepresentation of amplicon ends.[28] Sequences and primers for emulsion