Genetics inMedicine ORIGINAL RESEARCH ARTICLE

Comprehensive next-generation sequence analyses of the entire mitochondrial genome reveal new insights into the molecular diagnosis of mitochondrial DNA disorders

Hong Cui, PhD¹, Fangyuan Li, MD, PhD¹, David Chen, MS¹, Guoli Wang, PhD¹, Cavatina K. Truong, BS¹, Gregory M. Enns, MD², Brett Graham, MD, PhD¹, Margherita Milone, MD, PhD³, Megan L. Landsverk, PhD¹, Jing Wang, MD¹, Wei Zhang, MD, PhD¹ and Lee-Jun C. Wong, PhD¹

Purpose: The application of massively parallel sequencing technology to the analysis of the mitochondrial genome has demonstrated great improvement in the molecular diagnosis of mitochondrial DNA-related disorders. The objective of this study was to investigate the performance characteristics and to gain new insights into the analysis of the mitochondrial genome.

Methods: The entire mitochondrial genome was analyzed as a single amplicon using a long-range PCR-based enrichment approach coupled with massively parallel sequencing. The interference of the nuclear mitochondrial DNA homologs was distinguished from the actual mitochondrial DNA sequences by comparison with the results obtained from conventional PCR-based Sanger sequencing using multiple pairs of primers.

Results: Our results demonstrated the uniform coverage of the entire mitochondrial genome. Massively parallel sequencing of the single amplicon revealed the presence of single-nucleotide

INTRODUCTION

The genetics of mitochondrial disorders is complex in that disease can be caused by mutations in either nuclear-encoded genes or maternally inherited mitochondrial DNA (mtDNA), which are present in hundreds to thousands of copies per cell.¹⁻³ Pathogenic mtDNA mutations, ranging from point mutations to large deletions, are often present in mixed proportions with wildtype mtDNA molecules within the same cell, a biological phenomenon termed heteroplasmy. The degree of mtDNA mutant heteroplasmy can vary significantly across different tissues of the same individual, and the percentage of a mutation is an important contributor to the clinical phenotype.4,5 Determination of the heteroplasmic status of any mtDNA mutation is clinically important to providing a family with informative genetic counseling regarding recurrence risk. Therefore, accurate measurement of heteroplasmy is an essential component of the molecular diagnostic scheme for mtDNA-related disorders.

The diagnosis of an individual with suspected mtDNA-related disorders begins with a thorough clinical evaluation and assessment of family history.⁶ If a disorder with matrilineal inheritance can be established, common mtDNA point mutations and large

polymorphisms and nuclear homologs of mtDNA sequences that cause the erroneous and inaccurate variant calls when PCR/Sanger sequencing approach was used. This single amplicon massively parallel sequencing strategy provides an accurate quantification of mutation heteroplasmy as well as the detection and mapping of mitochondrial DNA deletions.

Conclusion: The ability to quantitatively and qualitatively evaluate every single base of the entire mitochondrial genome is indispensible to the accurate molecular diagnosis and genetic counseling of mitochondrial DNA-related disorders. This new approach may be considered as first-line testing for comprehensive analysis of the mitochondrial genome.

Genet Med 2013:15(5):388-394

Key Words: long-range PCR; massively parallel sequencing; mitochondrial disorders; mitochondrial genome; mtDNA large deletions; nuclear mitochondrial DNA homologs (NUMTs)

deletions are typically analyzed using PCR-based targeted assays and Southern blotting methodology. If the screening for common point mutations and large deletions is negative, then the entire mitochondrial genome is usually analyzed by a conventional PCR-based Sanger sequencing approach with multiple pairs of primers. When a mutation is identified, analysis of the degree of heteroplasmy is carried out using various methods, including the commonly used allele refractory mutation system-based quantitative PCR (ARMS qPCR) for quantitative measurement.⁷ Large deletions are usually confirmed by PCR using primers encompassing the deletion break points followed by sequence analysis to map the break points, or, alternatively, by oligonucleotide array comparative genome hybridization.8 These sequential approaches are time consuming and costly. In an effort to improve efficiency and sensitivity of mtDNA mutation detection, we recently developed a comprehensive one-step long range/massively parallel sequencing (LR-PCR/MPS)-based approach that uses a single pair of PCR primers to amplify the entire mitochondrial genome. This method allows the simultaneous detection of point mutations with quantified heteroplasmy as well as large mtDNA deletions with accurately mapped break points.9

¹Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas, USA; ²Department of Genetics, Stanford University, Stanford, California, USA; ³Department of Neurology, Mayo Clinic, Rochester, Minnesota, USA. Correspondence: Wei Zhang (wzhang2@bcm.edu)

Submitted 16 August 2012; accepted 3 October 2012; advance online publication 3 January 2013. doi:10.1038/gim.2012.144

In this report, we demonstrate the clinical utility of the LR-PCR/MPS-based method, uncover unintended problems embedded in the traditional PCR-based assays, and high-light advantages of the new approach. Issues such as incorrect base calls due to primer mismatch, apparent heteroplasmy ascribed to the coamplification of nuclear homologs of mtDNA (NUMTs), and inaccurate heteroplasmy quantification are discussed. The enhanced performance of the LR-PCR/MPS approach is evidenced by both the simultaneous measurement of heteroplasmy at any nucleotide position of the entire mitochondrial genome and the detection and mapping of mtDNA deletions.

MATERIALS AND METHODS

Patients and DNA

Patients were referred to the Mitochondrial Diagnostic Laboratory at the Medical Genetics Laboratories of the Baylor College of Medicine, for the molecular evaluation of mitochondrial disorders. Total DNA was isolated from either peripheral blood lymphocytes or muscle biopsy using a commercially available DNA isolation kit (Gentra Systems, Minneapolis, MN) according to the manufacturer's protocols.

Sanger sequencing

Sanger sequencing analysis of the mtDNA was performed using 24 pairs of overlapping primers as described previously.^{10–12} Purified PCR products were sequenced using BigDye Terminator (Life Technologies, Green Island, NY) chemistry on an ABI3730XL automated DNA sequencer (Life Technologies).

LR-PCR and MPS

Mitochondrial whole-genome amplification by single-amplicon LR-PCR and construction of Illumina indexed libraries were preformed as previously described.⁹ Twelve indexed DNA libraries were pooled together at equal molar ratio and sequenced in a single lane of one flow cell on HiSeq2000 (Illumina, San Diego, CA) with 76-bp single-end reads.

Mapping of large deletion junctions

Indicated by the size of the LR-PCR products and the sequence coverage profile, single or multiple mtDNA large deletions were characterized by further alignment of the unmapped sequence reads, with modified parameters and reduced stringency, to the mitochondrial reference sequence (NC_012920) using NextGENe software (SoftGenetics, State College, PA). The deletion junctions detected at coverage greater than 5× were recorded and confirmed by targeted PCR followed by Sanger sequencing.

RESULTS

Single-amplicon, LR-PCR-based mtDNA enrichment avoids NUMTs interference

NUMTs are homologous sequences of mtDNA present in the nuclear genome.¹³ Coamplification of both mtDNA and NUMTs can potentially interfere with the accurate heteroplasmy

quantification of many mtDNA regions.14,15 To examine the abundance of NUMTs computationally, mtDNA segments generated from a sliding 175-bp window with 10-bp increments were used to search for homologous sequences in the human nuclear haploid genome. Any identified region >100 bp in size with a minimum of 80% sequence homology was plotted against the mitochondrial genome. As shown in Figure 1a, these regions occur at different frequencies, with an average copy of 9.49 \pm 7.93, indicating that each 175-bp mtDNA segment has an average of 9.5 copies of NUMTs present in the nuclear genome. Among these, multiple regions, including m.500-m.1500, m.2000-m.2500, m.4200-m.4700, m.5400m.5700, and m.12000-m.12300, have more than 20 copies. The estimated total size of NUMTs is ~1 Mb, which is 0.03% of the size of human genome. This result is consistent with the value derived from the MITOMAP database.¹⁶ Enrichment of mtDNA using oligonucleotide probe hybridization will capture NUMTs and contributes to the observed variation of sequence coverage of mtDNA (Figure 1b). The relative amount of sequence captured in different regions of the mitochondrial genome largely depends on probe design. The mtDNA sequence obtained as a by-product of exome capture without using specific mtDNA probes, although at much lower coverage, definitely contains NUMT sequences (Figure 1c). Figure 1b,c show that different regions of mtDNA are not uniformly represented in the NUMTs.

In contrast, the single-amplicon LR-PCR approach not only avoids NUMT amplification but also provides uniform mtDNA coverage (**Figure 1d**). Enrichment of the mtDNA using multiple pairs of PCR primers has been widely used for the analysis



Figure 1 Long-range PCR overcomes nuclear mitochondrial DNA homologs interference with the enrichment of mitochondrial genome. (a) Frequency of mitochondrial DNA (mtDNA) homologs presented in the nuclear genome. (b) Enrichment of mtDNA by solution capture with mtDNA-specific probes before massively parallel sequencing (MPS) results in uneven coverage of mtDNA sequence. (c) Distribution of mtDNA sequences nonspecifically captured by nuclear exome probes. (d) Long-range PCR amplification of mtDNA before MPS provides uniform coverage of the entire mitochondrial genome.

of the mitochondrial genome.^{17,18} However, multiplexed PCR does not provide uniform coverage when these amplicons are pooled and sequenced by MPS.⁹ In addition, due to the highly polymorphic feature of the mitochondrial genome, the greater the number of primer pairs used for amplification, the higher the likelihood of encountering single-nucleotide polymorphisms (SNPs) at the primer binding sites. Therefore, in the presence of highly abundant NUMTs, the multiplexed PCR-based assay will inevitably have decreased accuracy of variant calls and heteroplasmy quantifications.

To illustrate the NUMTs interference in base calls, a representative case is shown in Figure 2. Sanger sequencing of the PCR product generated by primer pair m.4013_m.4031F and m.4822_m.4804R detected m.4104A→G, m.4312C→T, m.4318C \rightarrow T, m.4456C \rightarrow T, and m.4736T \rightarrow C rare variants (Figure 2a). However, the LR-PCR/MPS results derived from the same individual did not detect these five variants, but clearly detected the m.4216T \rightarrow C and m.4232T \rightarrow C variants instead (Figure 2b). On close investigation, we found that this discrepancy between the Sanger sequencing and MPS results occurred because the individual's mtDNA contains two additional private variants, m.4017C \rightarrow T and m.4029C \rightarrow A, sitting at the forward primer used for the Sanger sequencing (Supplementary Figure S1a online). The two mismatches prevented proper mtDNA amplification, and instead allowed the amplification of NUMT sequence because as the primers have perfect homology within the NUMT sequence (**Supplementary Figure S1b** online). Thus, the five rare variants detected in the Sanger PCR amplicon are false-positive calls derived from NUMT interference.

A single SNP at a primer site may also mildly affect amplification and cause a quantitative difference in base calls. As shown in the **Supplementary Figure S2** online, the presence of a homoplasmic change (m.4020C \rightarrow T) in the mtDNA from another individual tolerated the binding of a mildly mismatched primer and allowed the coamplification of both the mtDNA and the NUMTs, leading to apparently low-level heteroplasmy of m.4318C \rightarrow T and m.4456C \rightarrow T detected by Sanger sequencing.

As illustrated in **Figure 2c**, nucleotide changes have been reported at almost every single-nucleotide position of the mtDNA with variable frequencies across different ethnic groups.¹⁹ Similarly, mtDNA pathogenic mutations are distributed throughout the mitochondrial genome (**Figure 2d**).¹⁹ The combination of highly polymorphic mtDNA and the presence of NUMTs significantly increase the risk of misdiagnosis. The results described here underscore the necessity of eliminating NUMTs interference before sequencing.

Uniform coverage of the entire mitochondrial genome by single-amplicon LR-PCR/MPS allows the detection and mapping of mtDNA deletions

Previous studies have shown that mtDNA target gene capture by oligonucleotide probe hybridization followed by MPS



Figure 2 The presence of nuclear mitochondrial DNA homologs and single-nucleotide polymorphisms (SNPs) at primer binding sites produces incorrect variant calls. (a) Targeted PCR/Sanger sequencing and (b) long-range PCR/massively parallel sequencing (MPS) rendered inconsistent results. The m.4104A \rightarrow G, m.4312C \rightarrow T, m.4318C \rightarrow T, m.4456C \rightarrow T, and m.4736T \rightarrow C homoplasmic changes were detected by Sanger sequencing but not by MPS, whereas m.4216T \rightarrow C and m.4232T \rightarrow C homoplasmic changes were solely detected by MPS. (c,d) Highly polymorphic features of mitochondrial DNA. Frequencies of MITOMAP-reported (c) SNPs and (d) mutations were plotted against mitochondrial genome.

New insights of mtDNA by MPS | CUI et al

ORIGINAL RESEARCH ARTICLE



Figure 3 Identification of single and multiple mitochondrial large deletions with mapped break points. Coverage pattern of (a) single versus (b) multiple mitochondrial large deletions (upper row) and their corresponding break points (lower row) analyzed by the long-range PCR/massively parallel sequencing method.

provides nonuniform coverage, significantly reducing the possibility of detecting mtDNA deletions.⁹ In addition, multiplexed PCR/MPS exhibits variable coverage among different PCR fragments; therefore, large mtDNA deletions cannot be detected reliably and readily. In contrast, enrichment of the mitochondrial genome with single-amplicon LR-PCR amplification provides uniform coverage of the entire mitochondrial genome (**Figure 1d**), allowing accurate detection of large deletions.

Figure 3a shows LR-PCR/MPS analysis of a DNA sample from the blood specimen of a 10-year-old boy presenting with encephalopathy, exercise intolerance, easy fatigability, and sensorineural hearing loss. In addition to 43 homoplasmic variants, the sharp decrease in read coverage from m.7638 to m.15434 indicates a heteroplasmic large deletion of 7,797 bp. The degree of deletion heteroplasmy was estimated by comparing the coverage of deleted versus nondeleted regions. The deletion break point at a single-base resolution is clearly revealed by our LR-PCR/MPS method. Conventional PCR using primers flanking the deletion followed by Sanger sequencing of the deletion junction confirmed the MPS result with exactly the same break points.²⁰

Another example is the coverage profile of a muscle sample from a 70-year-old man with myopathy shown in **Figure 3b**. Instead of sharp deletion junctions, an arch-shaped coverage pattern was observed, suggesting multiple mtDNA deletions. Realignment of the unmapped sequences to the reference sequences with less stringent parameters revealed multiple deletion junctions. A total of 48 junction sequences were identified. The exact deletion break points were confirmed by targeted PCR followed by Sanger sequencing.

Low-level mutation heteroplasmy and its clinical and genetic significance

In general, Sanger sequencing does not detect heteroplasmy <15%.²¹ Although specific primers or probes can be designed for the quantification of target positions,^{7,22} it is laborious to validate the method for every novel mtDNA point mutation. In addition, primers may contain modifications and probes that are specific for either the wild-type or mutant allele,^{23–25} thus, a difference in PCR amplification efficiency for these two



Figure 4 Low-level heteroplasmy detected by massively parallel sequencing (MPS) method in pedigree studies. MPS method was used to quantify the levels of heteroplasmy for the mutations among different families: (a) family 1, m.3243A \rightarrow G; (b) family 2, m.5001dupA; (c) family 3, m.4296G \rightarrow A; (d) family 4, m.7222A \rightarrow G. Additional results of allele refractory mutation system–based quantitative PCR are shown in (b) following the star sign inside parentheses.

alleles is expected, leading to inherent inaccuracies with the measurement of the degree of heteroplasmy. Our group has recently demonstrated that, for mtDNA analysis at $20,000 \times$ coverage, the experimental error rate of the MPS using the Illumina HiSeq2000 platform was $0.326 \pm 0.335\%$, with a limit of detection of 1.33%.⁹ By applying LR-PCR/MPS-based analysis, the level of minor allele heteroplasmy can be more accurately detected with reproducibility (**Supplementary Table S1** online). The results from the following cases illustrate the power of LR-PCR/MPS in providing information necessary for more accurate risk assessment and genetic counseling.

For family 1, the proband had a history of hearing loss and vision problems (**Figure 4a**). Screening for mtDNA common point mutations by allele-specific oligonucleotide hybridization detected the m.3243A \rightarrow G mutation in the tRNA^{Leu(UUR)} gene.²⁶ This mutation was too low to be detected by Sanger,

but was measured as 7% heteroplasmy by LR-PCR/MPS (**Supplementary Figure S3** online). Her daughter inherited this mutation at 9.3% heteroplasmy in her blood sample, as determined by MPS. This case illustrates that Sanger sequencing would have missed the detection of this low heteroplasmic mutation, whereas it is readily detected by MPS. Although these levels of heteroplasmy are low, it is well established that there is selection against the m.3243A \rightarrow G mutation in the rapidly dividing blood cells.^{27,28} Thus, these findings prompt examination of mutation heteroplasmy levels in other tissues.

The 2-year-old proband in family 2 presented with developmental delay, seizures, cardiomyopathy, and lactic acidemia. MPS detected an m.5001dupA (p.1178NfsX22, ND2) mutation at 99.5% and 16% heteroplasmy in the muscle and blood, respectively. Because this mutation was not detected in a blood specimen from the proband's asymptomatic mother (**Figure 4b**), it probably occurred *de novo* in the proband; however, germline mosaicism cannot be excluded. The determination of whether an mtDNA mutation is inherited or sporadic is clearly dependent on the sensitivity of the detection method used. Accurate detection of low-level heteroplasmy therefore plays an important role in the molecular diagnosis and counseling of mtDNA-related disorders.

In family 3, the proband is a 4-year-old child with developmental delay, hypotonia, encephalopathy, peripheral neuropathy, and elevated lactate. This patient was found to harbor a nearly homoplasmic novel mutation, m.4296G \rightarrow A, in the tRNA^{Ile} gene by Sanger sequencing (Supplementary Figure S4a online).²⁹ Her asymptomatic mother also had the same mutation detected in blood by Sanger sequencing, but the signal on the electrophoretic chromatogram was only a fraction of the wild-type peak, suggesting low-level heteroplasmy (Supplementary Figure S4a online). To obtain more quantitative results, ARMS qPCR was performed on blood samples from both individuals. The ARMS qPCR analyses showed 78% and 4% heteroplasmy in the affected child and her mother, respectively (Figure 4c). These results were not consistent with the detection limit of Sanger sequencing, in which homoplasmic changes above 85% are expected to be quantified and heteroplasmic changes under 15% cannot be clearly identified. Subsequently, LR-PCR/MPS analyses revealed 97.9% and 27.8% heteroplasmy in the proband and the mother (Supplementary Figure S4b online), respectively, which are consistent with Sanger sequencing results. Thus, LR-PCR/MPS outperformed ARMS qPCR in this instance. These results illustrate the limitations of current PCR-based methods for the quantification of mtDNA heteroplasmy.

In family 4 (Figure 4d), the proband is a 65-year-old woman with peripheral neuropathy, muscle weakness, ptosis, abnormal muscle histological findings, and abnormal electromyography.³⁰ A heteroplasmic mutation, m.7222A \rightarrow G (p.Y440C in COI protein), was detected in this patient's muscle specimen by Sanger sequencing. LR-PCR/MPS analyses indicated 37% mutation heteroplasmy in her muscle sample, but did not detect the mutation in her blood or in blood samples from her sister and

daughter. The collective results thus suggest that the mutation most likely arose sporadically in the proband's muscle tissue. Therefore, the risk of passing this likely somatic mutation to her children is very low.

These four examples illustrated the power of MPS-based analyses in detecting low-level heteroplasmy and assessing whether a mutation is inherited, *de novo*, or somatic.

DISCUSSION

The interference of NUMTs: an old problem with a new solution

mtDNA exists as a multicopy molecule within cells and is highly polymorphic. These unique characteristics challenge both the traditional stepwise analytical methods and various newly developed MPS-based approaches for the detection of mtDNA mutations at all levels. Oligonucleotide probehybridization has been used for enrichment of mtDNA in various MPS-based studies.^{17,31} However, our computational in silico search has identified more than 1,000 NUMT sites with >80% mtDNA sequence homology, similar to the findings in a recent report.³² The scope of NUMT signal is obvious by mapping capture-enriched sequencing data from a mtDNA-depleted p zero cell line. It is estimated that 0.1% of sequence reads from the total genomic DNA mapped to the mtDNA reference sequence. The presence of NUMTs resulted in many false-positive variant calls and numerous low heteroplasmic calls. Even after the removal of NUMTs using in-house developed software that distinguishes low-level heteroplasmy from sequencing error,³³ the false-negative and false-positive rates were still significantly above levels acceptable for use in clinical diagnostic laboratories.32,34 Although reads generated from NUMT regions can be partially deselected by a stringent alignment algorithm, regions that are nearly identical to mtDNA, such as fragments in the regions of 9p24.3 and 12q11, are practically impossible to remove. Therefore, regardless of stringency, exome capture will inevitably cocapture NUMTs, which will then confound mutational analyses and lead to errant clinical diagnosis.

Another strategy of mtDNA enrichment using multiple amplicons not only fails to address the problem of contaminating NUMTs but also introduces additional problems. These problems include the increased probability of one or more SNPs sitting within primer binding sites, which results in reduced amplification efficiency, masked SNP data, and possibly preferential amplification of NUMTs. For instance, the m.4456C \rightarrow T variant (Figure 2a) has also been reported as a polymorphic change in MITOMAP.35 This variant was identified within one of the 36 short PCR amplicons tiled to cover the entire mtDNA. Close examination revealed that the mtDNA primers used for the amplification of the fragment encompassing the m.4456C \rightarrow T variant were completely homologous to a NUMT sequence on chromosome 1. Thus, further investigation is warranted to confirm that this variant is an authentic mtDNA variant and not part of the NUMT.

Our approach of using one pair of primers to amplify the entire mitochondrial genome can effectively resolve the problems associated with NUMTs. The LR-PCR primers were carefully designed to avoid any reported SNPs. Moreover, the primer sites are also sequenced by Sanger sequencing for each sample to ensure that no SNPs are sitting at the primer sites. When this happens, alternative LR-PCR primers that are clear of SNPs are used to repeat LR-PCR/MPS. Our current and previous studies demonstrated that the MPS-based method offers several advantages over conventional Sanger sequencing.⁹ This is mainly achieved by the unbiased enrichment of the entire mitochondrial genome and avoiding the interference of NUMTs and SNPs that are embedded in other enrichment approaches. Adherence to these measures may abrogate current limitations of MPSbased mtDNA analysis in forensic studies.³⁶

In addition to forensic applications, heteroplasmic mtDNA variants have been considered useful genetic markers for genetic disease diagnosis, cancer prognosis, and other research studies.^{17,18,37,38} Knowing the limitation of the MPS technologies and the interference of NUMTs and SNPs, evaluation and interpretation of mtDNA sequence results should be carried out with caution. When paraffin-fixed tumor tissues are used and target enrichment is limited to the use of multiple pairs of primers for PCR enrichment,¹⁸ interference of NUMTs and SNPs is significant, and cautious interpretation of results derived from these approaches is indicated.

Detection and mapping of mtDNA deletions

The nonuniform coverage profile produced by capture enrichment (Figures 1b,c) or multiplex PCR amplification makes the detection of heteroplasmic mtDNA deletions unreliable, if not impossible. Enrichment of the entire mitochondrial genome by single-amplicon LR-PCR allows accurate detection of mtDNA large deletions with unequivocally mapped break points (Figure 3). Southern blot analysis is not sensitive enough to reliably detect low heteroplasmic mtDNA deletions. By applying LR-PCR/MPS, the detection sensitivity of mtDNA multiple deletions is enhanced, leading to an increased clinical diagnostic yield. mtDNA multiple deletions are associated with aging and oxidative damage, or more importantly, secondary to primary mutations in nuclear genes responsible for mtDNA biosynthesis and integrity maintenance, such as POLG, TWINKLE, and OPA1.16,39,40 Fifteen DNA samples from muscle showing mtDNA multiple deletions on MPS-based analysis were subjected to sequence analysis of a panel of nuclear genes responsible for the maintenance of mtDNA integrity. POLG mutations were identified in three patients and RRM2B and OPA1 mutations in one patient each (unpublished data). All 15 patients with mtDNA multiple deletions were adults. Among them, nine were more than 50 years old, suggesting that the accumulation of mtDNA multiple deletions may be secondary to both nuclear gene defects and aging. Clearly, MPS-based analysis of the entire mitochondrial genome is useful for the diagnosis of mtDNA multipledeletion-related mitochondrial disorders that are caused by the nuclear-encoded genes.

Quantification of mtDNA heteroplasmy

The degree of mutant heteroplasmy, its tissue distribution, and threshold significantly contribute to disease phenotype and severity. Before the availability of MPS technologies, mtDNA heteroplasmy was typically measured by various PCR-based methods, including ARMS qPCR and pyrosequencing^{7,22} for a limited number of common point mutations. If a variant is novel and target quantification is not available, Sanger sequencing of the target region can be used to detect the high levels of heteroplasmy. However, due to the high frequencies of mtDNA SNPs distributed along the entire mitochondrial genome and assay limitations, accurate quantification of heteroplasmy by these methods can be problematic.

LR-PCR/MPS-based analysis not only provides accurate quantification of nucleotide heteroplasmy, but it is also sensitive enough to detect heteroplasmy as low as 1.5% at every singlenucleotide position of the entire mitochondrial genome without the interference of NUMTs and SNPs. As illustrated in Figure 4, as compared with Sanger sequencing and/or PCR-based quantification methods, heteroplasmy measurements via LR-PCR/MPS for various tissues of the proband and matrilineal relatives will better distinguish between inherited and de novo mutations, and therefore improve diagnosis and recurrence risk estimates. The presence of nonspecifically enriched NUMTs either by capture or short PCR amplification may result in false-positive assignment of carrier status. In general, the consequence of NUMT interference is less severe than that of the SNPs due to the excessive copy number of mtDNA relative to NUMTs. However, depending on the position and the number of rare mtDNA variants residing within PCR primer binding sites, suboptimal annealing conditions may cause preferential amplification of the NUMTs to various extents, potentially leading to complete dropout of authentic mtDNA variants. By comparing the data generated from traditional Sanger sequencing and LR-PCR/MPS described in this report, the false-positive calls by Sanger sequencing are mostly due to private SNPs residing within the PCR primer binding sites and previously unrecognized skewed amplification of the NUMTs. These results clearly demonstrate that LR-PCR/MPS has better analytic performance than the traditional Sanger sequencing or PCR-based targeted quantification assay.

Conclusion

Sanger sequencing has long been considered as the gold standard for molecular diagnosis of mtDNA-based disorders. Our results demonstrated that single-amplicon LR-PCR amplification of the entire mitochondrial genome not only eliminates potential NUMTs and SNP interference but also renders a uniform coverage of the entire 16.6-kb mitochondrial genome, which is essential for reliable and accurate detection of single and multiple large mtDNA deletions. Coupled with MPS, the accurate detection of low-level heteroplasmy at <15% also becomes possible. With the higher sensitivity, specificity, and accuracy in the detection of a wider spectrum of mutation types, we propose that singleamplicon LR-PCR/MPS be adopted as the new standard for the comprehensive analysis of the mitochondrial genome.

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper and available at the http://www.nature.com/gim

ACKNOWLEDGMENT

We thank Eric S. Schmitt for valuable discussion and input on the manuscript.

DISCLOSURE

Many of the authors are faculty members in the Department of Molecular and Human Genetics at BCM. The Medical Genetics Laboratories of the department offer extensive, fee-based genetic tests including the use of massively parallel sequencing for molecular analyses. The authors (W.Z., H.C., L.-J.W.) have one pending patent application on the comprehensive analysis of the mitochondrial genome by next-generation sequencing.

REFERENCES

- 1. Gahan ME, Miller F, Lewin SR, et al. Quantification of mitochondrial DNA in peripheral blood mononuclear cells and subcutaneous fat using real-time polymerase chain reaction. *J Clin Virol* 2001;22:241–247.
- Barthélémy C, Ogier de Baulny H, Diaz J, et al. Late-onset mitochondrial DNA depletion: DNA copy number, multiple deletions, and compensation. *Ann Neurol* 2001;49:607–617.
- 3. Cavelier L, Jazin EE, Eriksson I, et al. Decreased cytochrome-c oxidase activity and lack of age-related accumulation of mitochondrial DNA deletions in the brains of schizophrenics. *Genomics* 1995;29:217–224.
- Bai RK, Wong LJ. Simultaneous detection and quantification of mitochondrial DNA deletion(s), depletion, and over-replication in patients with mitochondrial disease. J Mol Diagn 2005;7:613–622.
- Morten KJ, Ashley N, Wijburg F, et al. Liver mtDNA content increases during development: a comparison of methods and the importance of age- and tissue-specific controls for the diagnosis of mtDNA depletion. *Mitochondrion* 2007;7:386–395.
- Wong LJ, Scaglia F, Graham BH, Craigen WJ. Current molecular diagnostic algorithm for mitochondrial disorders. *Mol Genet Metab* 2010;100: 111–117.
- Wang J, Venegas V, Li F, Wong L-J. Analysis of Mitochondrial DNA Point Mutation Heteroplasmy by ARMS Quantitative PCR. John Wiley & Sons: Hoboken, NJ, 2011.
- Chinault AC, Shaw CA, Brundage EK, Tang LY, Wong LJ. Application of dualgenome oligonucleotide array-based comparative genomic hybridization to the molecular diagnosis of mitochondrial DNA deletion and depletion syndromes. *Genet Med* 2009;11:518–526.
- Zhang W, Cui H, Wong LJ. Comprehensive 1-step molecular analyses of mitochondrial genome by massively parallel sequencing. *Clin Chem* 2012; e-pub ahead of print 9 July 2012.
- Brautbar A, Wang J, Abdenur JE, et al. The mitochondrial 13513G→A mutation is associated with Leigh disease phenotypes independent of complex I deficiency in muscle. *Mol Genet Metab* 2008;94:485–490.
- Wang J, Brautbar A, Chan AK, et al. Two mtDNA mutations 14487T→C (M63V, ND6) and 12297T→C (tRNA Leu) in a Leigh syndrome family. *Mol Genet Metab* 2009;96:59–65.
- 12. Ware SM, El-Hassan N, Kahler SG, et al. Infantile cardiomyopathy caused by a mutation in the overlapping region of mitochondrial ATPase 6 and 8 genes. *J Med Genet* 2009;46:308–314.
- Tsuzuki T, Nomiyama H, Setoyama C, Maeda S, Shimada K. Presence of mitochondrial-DNA-like sequences in the human nuclear DNA. *Gene* 1983;25:223–229.
- Parfait B, Rustin P, Munnich A, Rötig A. Co-amplification of nuclear pseudogenes and assessment of heteroplasmy of mitochondrial DNA mutations. *Biochem Biophys Res Commun* 1998;247:57–59.
- Hirano M, Shtilbans A, Mayeux R, et al. Apparent mtDNA heteroplasmy in Alzheimer's disease patients and in normals due to PCR amplification of nucleus-embedded mtDNA pseudogenes. *Proc Natl Acad Sci USA* 1997;94:14894–14899.

- Milone M, Younge BR, Wang J, Zhang S, Wong LJ. Mitochondrial disorder with OPA1 mutation lacking optic atrophy. *Mitochondrion* 2009;9:279–281.
- Calvo SE, Compton AG, Hershman SG, et al. Molecular diagnosis of infantile mitochondrial disease with targeted next-generation sequencing. *Sci Transl Med* 2012;4:118ra10.
- He Y, Wu J, Dressman DC, et al. Heteroplasmic mitochondrial DNA mutations in normal and tumour cells. *Nature* 2010;464:610–614.
- 19. MITOMAP. A Human Mitochondrial Genome Database. http://www. mitomap.org. 2011.
- Wong LJ, Perng CL, Hsu CH, et al. Compensatory amplification of mtDNA in a patient with a novel deletion/duplication and high mutant load. J Med Genet 2003;40:e125.
- Rohlin A, Wernersson J, Engwall Y, Wiklund L, Björk J, Nordling M. Parallel sequencing used in detection of mosaic mutations: comparison with four diagnostic DNA screening techniques. *Hum Mutat* 2009;30:1012–1020.
- Venegas V, Halberg MC. Quantification of mtDNA mutation heteroplasmy (ARMS qPCR). *Methods Mol Biol* 2012;837:313–326.
- Alston CL, Lowe J, Turnbull DM, Maddison P, Taylor RW. A novel mitochondrial tRNAGlu (MTTE) gene mutation causing chronic progressive external ophthalmoplegia at low levels of heteroplasmy in muscle. *J Neurol Sci* 2010;298:140–144.
- White HE, Durston VJ, Seller A, Fratter C, Harvey JF, Cross NC. Accurate detection and quantitation of heteroplasmic mitochondrial point mutations by pyrosequencing. *Genet Test* 2005;9:190–199.
- 25. Shanske S, Wong LJ. Molecular analysis for mitochondrial DNA disorders. *Mitochondrion* 2004;4:403–415.
- Tang S, Halberg MC, Floyd KC, Wang J. Analysis of common mitochondrial DNA mutations by allele-specific oligonucleotide and Southern blot hybridization. *Methods Mol Biol* 2012;837:259–279.
- Rahman S, Poulton J, Marchington D, Suomalainen A. Decrease of 3243 A→G mtDNA mutation from blood in MELAS syndrome: a longitudinal study. *Am J Hum Genet* 2001;68:238–240.
- Shanske S, Pancrudo J, Kaufmann P, et al. Varying loads of the mitochondrial DNA A3243G mutation in different tissues: implications for diagnosis. *Am J Med Genet A* 2004;130A:134–137.
- 29. Cox R, Platt J, Chen LC, Tang S, Wong LJ, Enns GM. Leigh syndrome caused by a novel m.4296G→A mutation in mitochondrial tRNA isoleucine. *Mitochondrion* 2012;12:258–261.
- Massie R, Wang J, Chen LC, et al. Mitochondrial myopathy due to novel missense mutation in the cytochrome c oxidase 1 gene. J Neurol Sci 2012;319:158–163.
- 31. Tucker EJ, Hershman SG, Köhrer C, et al. Mutations in MTFMT underlie a human disorder of formylation causing impaired mitochondrial translation. *Cell Metab* 2011;14:428–434.
- Li M, Schroeder R, Ko A, Stoneking M. Fidelity of capture-enrichment for mtDNA genome sequencing: influence of NUMTs. *Nucleic Acids Res* 2012;40:e137.
- 33. Li M, Stoneking M. A new approach for detecting low-level mutations in next-generation sequence data. *Genome Biol* 2012;13:R34.
- Li M, Schönberg A, Schaefer M, Schroeder R, Nasidze I, Stoneking M. Detecting heteroplasmy from high-throughput sequencing of complete human mitochondrial DNA genomes. *Am J Hum Genet* 2010;87:237–249.
- 35. Dobrowolski SF, Hendrickx AT, van den Bosch BJ, et al. Identifying sequence variants in the human mitochondrial genome using high-resolution melt (HRM) profiling. *Hum Mutat* 2009;30:891–898.
- Bandelt HJ, Salas A. Current next generation sequencing technology may not meet forensic standards. *Forensic Sci Int Genet* 2012;6: 143–145.
- Zaragoza MV, Fass J, Diegoli M, Lin D, Arbustini E. Mitochondrial DNA variant discovery and evaluation in human Cardiomyopathies through nextgeneration sequencing. *PLoS ONE* 2010;5:e12295.
- Guo Y, Cai Q, Samuels DC, et al. The use of next generation sequencing technology to study the effect of radiation therapy on mitochondrial DNA mutation. *Mutat Res* 2012;744:154–160.
- Milone M, Brunetti-Pierri N, Tang LY, et al. Sensory ataxic neuropathy with ophthalmoparesis caused by POLG mutations. *Neuromuscul Disord* 2008;18:626–632.
- Van Hove JL, Cunningham V, Rice C, et al. Finding twinkle in the eyes of a 71-year-old lady: a case report and review of the genotypic and phenotypic spectrum of TWINKLE-related dominant disease. *Am J Med Genet A* 2009;149A:861–867.