Rapid identification of a novel complex I MT-ND2 m.10134C>A mutation in a Leigh syndrome patient

David K. Miller  
*University of Queensland*

Minal J. Menezes  
*Children's Hospital at Westmead, Sydney, New South Wales, Australia*

Cas Simons  
*University of Queensland*

Lisa G. Riley  
*Children's Hospital at Westmead, Sydney, New South Wales, Australia*

Sandra T. Cooper  
*Children's Hospital at Westmead, Sydney, New South Wales, Australia*

*See next page for additional authors*

Follow this and additional works at: [http://hsrc.himmelfarb.gwu.edu/smhs_intsysbio_facpubs](http://hsrc.himmelfarb.gwu.edu/smhs_intsysbio_facpubs)

Part of the [Systems Biology Commons](http://hsrc.himmelfarb.gwu.edu/smhs_intsysbio_facpubs)

**Recommended Citation**

Rapid Identification of a Novel Complex I MT-ND3 m.10134C>A Mutation in a Leigh Syndrome Patient

David K. Miller1,*, Minal J. Menezes2,3*, Cas Simons4*, Lisa G. Riley2, Sandra T. Cooper3,5, Sean M. Grimmond1, David R. Thorburn6,7, John Christodoulou2,3,8, Ryan J. Taft4,9,*

Abstract
Leigh syndrome (LS) is a rare progressive multi-system neurodegenerative disorder, the genetics of which is frequently difficult to resolve. Rapid determination of the genetic etiology of LS in a 5-year-old girl facilitated inclusion in Edison Pharmacetical’s phase 2B clinical trial of EPI-743. SNP-arrays and high-coverge whole exome sequencing were performed on the proband, both parents and three unaffected siblings. Subsequent multi-tissue targeted high-depth mitochondrial sequencing was performed using custom long-range PCR amplicons. Tissue-specific mutant load was also assessed by qPCR. Complex I was interrogated by spectrophotometric enzyme assays and Western Blot. No putatively causal mutations were identified in nuclear-encoded genes. Analysis of low-coverage off-target mitochondrial reads revealed a previously unreported mitochondrial mutation in the proband in MT-ND3 (m.10134C>A, p.Q26K), a Complex I mitochondrial gene previously associated with LS. Targeted investigations demonstrated that this mutation was 1% heteroplasmic in the mother’s blood and homoplasmic in the proband’s blood, fibroblasts, liver, and muscle. Enzyme assays revealed decreased Complex I activity. The identification of this novel LS MT-ND3 variant, the genomics of which was accomplished in less than 3.5 weeks, indicates that rapid genomic approaches may prove useful in time-sensitive cases with an unresolved genetic diagnosis.

Introduction
Leigh syndrome (LS) is a rare progressive neurodegenerative disorder that is characterised by early onset (typically infancy or early childhood) and a combination of signs including muscular hypotonia or spasticity, dystonia, nystagmus, psychomotor retardation and occasionally epilepsy [1]. LS patients may show rapid loss of previously acquired skills, failure to thrive and feeding difficulties, with survival generally measured in years [1,2]. Bilateral symmetrical lesions in the central nervous system, particularly in the brain stem, thalamus, and spinal cord are diagnostic, as are findings of lactic acidosis and a Magnetic Resonance Spectroscopy (MRS) lactate peak [1].

LS can be caused by a heterogeneous array of mitochondrial or nuclear genetic mutations that decrease aerobic energy production. Although the most prevalent LS mutations are associated with Complex IV (cytochrome C oxidase complex; e.g. COX15, SURF1) [OMIM 603646, OMIM 185620], Complex I subunit mutations are also common. Damaging variants have been identified in the mitochondrial genome encoded genes MT-ND1, MT-ND2, MT-ND3, MT-ND4, MT-ND5, MT-ND6 [3,4], and ten nuclear encoded Complex I NADH-ubiquinone oxidoreductase (NDU) proteins [5]. LS is the most common clinical presentation of a Complex I deficiency [6,7]. A clinical diagnosis of LS can be reached without knowledge of the causal mutation, the identification of which, until recently, has...
often been impractical due to the technical limitations and cost associated with Sanger sequencing multiple nuclear and mitochondrial candidate genes. Here we present the case of a child who required a definitive genetic diagnosis for inclusion in an ongoing Edison Pharmaceutical Phase 2B clinical trial of EPI-743, a para-benzoquinone that facilitates repletion of reduced intracellular glutathione and shows promise for LS patients. All ten LS children participating in a recent Phase 2A EPI-743 study exhibited reversal of disease progression regardless of genetic determinant or disease severity [8]. To identify the causal genetic variant responsible for LS in the proband we employed tiered set genomic investigations across all six members of the immediate family, including development of a targeted Nexxera-based mitochondrial DNA sequencing assay, which led to the discovery of a novel LS MT-ND3 mutation.

Patients and Methods

Case Report

The proband, a girl, is the youngest of four healthy children born to unrelated and clinically unremarkable Australian parents (Figure 1A). There were no antenatal or postnatal issues of note, and early growth and development were normal. First concerns were raised when the proband was approximately 4 years of age, when it became apparent that she had lost her ability to jump and had developed an abnormal gait. It was also noted at this time that her speech development was delayed. The proband’s parents stated that she was slower in acquiring new skills than her siblings from an early age. Following an MRI/MRS of the brain under general anaesthesia at 4.5 years of age her speech became more difficult to understand, her mobility and balance worsened, and her behaviour deteriorated. Formal cardiology, audiology and ophthalmology assessments were normal.

The patient’s score on the Newcastle Paediatric Mitochondrial Disease Scale for sections I-III was 17 (range 0–79). Diagnostic investigations were consistent with LS. The brain MRI revealed extensive signal abnormality involving putamen, and to a lesser extent, the globus pallidus bilaterally and the cerebral peduncles (Figure 1B). MRS also revealed elevated lactate in the basal ganglia (data not shown), although blood and CSF lactate levels were normal. Liver function enzymes were all within the normal range. The muscle biopsy showed fat accumulation and slightly increased subsarcolemmal mitochondrial aggregates. However there were no ragged red fibres and COX/SDH immunohistochemical staining was normal. Histological studies on the liver showed occasional moderately enlarged mitochondria with paracrystalline and crystalline inclusions.

Genomics and Bioinformatics

High-density SNP Arrays & Exome Sequencing. This research was approved by the Human Research Ethics Committees of the Children’s Hospital at Westmead and the University of Queensland, and written consent was obtained from the parents on behalf of themselves and their children (all children aged less than 10 years), as approved by the Human Research Ethics Committees. DNA was extracted from peripheral blood samples from the patient, both parents and three unaffected siblings. Copy number changes (e.g. large indels) and genomic rearrangements were assessed using the Illumina HumanOmni2.5-8 Multi-Use SNP arrays. Exome sequencing libraries were prepared with the TruSeq DNA sample prep kit, and exome capture was performed using the Nimblegen v3 Human EZ-Exome kits. Libraries from all six individuals were generated within 96 hours of sample receipt. Rapid sequencing (including on-instrument cluster generation) was performed on the Illumina HiSeq 2500 in less than twenty-seven hours.

Targeted Mitochondrial Sequencing. Three primers pairs [9] were used to generate overlapping PCR products (each ~6,000 bp) that encompassed the entire mitochondrial genome. PCR products from each member of the family were used to create libraries compatible with MiSeq and Ion Torrent PGM sequencing. MiSeq libraries were prepared from 1 ng of PCR product using the Nextera XT library preparation kits, and sequencing was performed using the MiSeq 300 cycle v2 kits as per the manufacturer’s instructions including on-instrument cluster generation. Ion Torrent libraries were prepared using 100 ng of each product in combination with the Ion Xpress Plus Library Kit for AB Library Builder. Library Fragments were clonally amplified onto Ion Sphere particles by emulsion PCR using the Ion One Touch Template Preparation kits, and sequenced on Ion Torrent S18 chips. In both cases libraries were pooled prior to sequencing and results obtained within 24 hours of the initial PCR reactions.

Informatics Pipeline. SNP arrays were analysed with Illumina BeadStudio, PLINK and a set of custom scripts. Exome and targeted mitochondrial sequencing reads were aligned to the reference human genome (hg19) using the Burrows-Wheeler alignment (BWA) and down-stream processing of sequence data was done with Picard v1.8, SAMtools v0.1.18 and the Genome Analysis Toolkit (GATK) version 2.2.8. Variants (SNP and Indels) were identified using GATK, following version four of the GATK Best Practice Variant Detection guide. Variants were annotated using Annovar with UCSC Known Genes models, and known polymorphisms were identified using dbSNP135, 1000 genomes and NHLBI exome project with minor allele frequencies recorded from each dataset. Subsequent analysis and identification of candidate variants was performed using an in-house workflow incorporating the annotated variant data and pedigree information.

Spectrophotometric assays of the respiratory chain enzyme activity

Spectrophotometric determinations of enzyme activity for all respiratory chain enzymes from skeletal muscle and liver biopsies were performed as described previously [10].

qPCR quantification of mutant load

PCR primers specific to the MT-ND3 (m.10134C>A) mutation, and to wild-type controls were designed and are available upon request. Quantitative PCR was performed using SYBR green dye and 1 Unit of Immolase DNA Polymerase (Bioline, Alexandria, NSW, Australia), with a final concentration of 1.5 mM MgCl₂, 500 pM of each primer, 2% DMSO and 1 M betaine. PCR products were generated from 5 ng of genomic DNA isolated (Qiamp DNA Mini kit) from the proband’s blood, fibroblasts, muscle and liver, and from the blood of all other family members. Standard curves were generated from ten-fold dilutions of the mutant and wildtype amplicons inserted into pCR<sup>TM</sup>TOPO vector (Life Technologies). The cycle conditions used were: 95°C for 12 minutes, 95°C for 10 s, 66°C for 15 s, 72°C for 10 s×35 cycles.

Western Blot

Tissue samples were lysed as previously described [11], and whole cell lysates from fibroblasts (15 μg of protein), and from muscle and liver (5 μg of protein), were separated by SDS-PAGE.
on a 4–12% polyacrylamide gel. A PVDF membrane was probed for 2 h at RT with 1:1,000 MitoProfile Total OXPHOS Human WB Antibody Cocktail (ab110411, Abcam). Anti-mouse IgG HRP (GE Healthcare) at 1:2,500 was used as the secondary antibody. Membranes were developed with ECL and exposed on Hyperfilm. Films were scanned on a Microtek ScanMaker 8700 and analysed using ImageQuant (GE Healthcare).

Results

In this case of a patient with typical LS features we undertook a tiered investigation of increasing resolution that consisted of (i) high-density SNP arrays, (ii) exome sequencing, (iii) targeted ultra-deep mitochondrial sequencing on two different platforms, and (iv) targeted multi-tissue sequencing. Analysis of the SNP array data revealed no potentially pathogenic copy number or genomic rearrangements. Exome sequencing was performed at high depth - a total of 71 gigabases were generated from a single rapid HiSeq 2500 sequencing run, yielding a non-redundant depth of 69x with 96% of targeted bases covered at 18-fold or greater. Both an unbiased genome-wide screen of all nuclear encoded genes and a manual investigation of known nuclear-encoded mitochondrial disease associated genes, however, failed to reveal any putatively disease causing homozygous, compound heterozygous or de novo mutations.

Taking advantage of the high-depth exome sequencing, we used off-target mitochondrial genome sequencing reads to survey for putative disease-causing variants and identified a low-coverage (6x) m.10134C>A mutation that was unique to the proband. This variant induces a p.Gln26Lys change in MT-ND3 (NADH-ubiquinone oxidoreductase chain 3), a gene that has previously been associated with LS [12–14]. The mutation (NC_012920.1:m.10134C>A; ClinVar SCV000155651) was predicted to be damaging by SIFT with an annotation of “Deleterious” with a score of −3.800, and “Possibly damaging”
by PolyPhen2 with a score of 0.611 [15,16]. The m.10134C>A variant is not recorded in dbSNP, was absent from the mtDB database, which lists mtDNA variants in 2704 sequenced mtDNA genomes [17], and was not detected in any of over 300 control exomes previously sequenced by our group.

We performed targeted ultra-deep mitochondrial investigations of the m.10134C>A mutation on the both the MiSeq (~20,000-fold) and the Ion Torrent (~5,000-fold), which revealed that the father and all three unaffected siblings carried only the wild-type allele, and that both the proband and the mother carried the m.10134C>A allele at levels of ~99.9% and 1% respectively (Table 1). These data strongly indicated that the p.Gln26Lys nonsynonymous MT-ND3 change was likely to be the cause of LS in this patient.

To assess the functional consequences of the m.10134C>A (p.Gln26Lys) mutation on MT-ND3, spectrophotometric enzyme assays of Complex I were performed, which revealed that Complex I was 24%, 17% and 13% of normal relative to protein, citrate synthase and Complex II, respectively, in patient skeletal muscle; and that Complex I was low relative to CS (50%) and Complex II (34%), but normal relative to protein in liver. Additionally, Complex II, citrate synthase and Complex IV activities were elevated in both muscle and liver, and Complex III was elevated in liver (Table 2), suggesting some degree of compensatory mitochondrial proliferation. Analysis by Western blot mirrored the enzyme assay results, showing a reduction in Complex I protein levels in muscle, but normal levels in liver (Figure 1D).

To investigate the tissue-specific load of the m.10134C>A mutation in the proband, ultra-deep targeted MiSeq sequencing and qPCR of DNA from muscle, liver and fibroblasts was performed. Sequencing revealed homoplasy for the mutation in all tissues tested (Table 1), which was confirmed by qPCR. This suggests that the differences in relative Complex I activity in muscle and liver are not the result of developmental mitochondrial mosaicism, and may simply reflect patient- and tissue-specific variability on the impact of the mutation on complex I enzyme activity, as noted for another MT-ND3 mutation, m.10191T>C [13]. These data also highlight the fact that homoplasmic mitochondrial mutations are not by definition embryonic lethal, and may have variable and tissue-specific consequences which are likely to be tied to background genetics (see below).

### Discussion

To resolve the genetics underpinning LS in this case we employed a tiered investigation using technologies of increasing resolution that resulted in the identification of a novel mutation in MT-ND3 (m.10134C>A, p.Gln26Lys), which facilitated patient inclusion into the phase 2B clinical trial of EPI-743. The SNP arrays, exome and targeted mitochondrial ultra-deep sequencing were completed in less than three and a half weeks, a timeframe comparable to commercially available single gene diagnostic tests. Other orthogonal approaches may have resolved the case in a similar timeframe (e.g. Sanger sequencing of the mitochondrial genome partnered with targeted nuclear gene investigations), but we note that the approach used here provided both maximum depth and breadth and therefore the highest likelihood of causal mutation detection. Access to genetic material from all members of the immediate family virtually eliminated all false positives. Amongst nuclear encoded genes we found only one putative de novo mutation, and no homozygous or compound heterozygous mutations that segregated with the proband, thus reducing the overall period of the investigation.
Three other LS MT-ND3 mutations have been reported - m.10158T>C, m.10191T>G, m.10197G>A [12–14] - and may shed light on some of the results described here. For example, a recent meta-analysis of patients carrying the m.10191T>G mutation demonstrated that there is no clear correlation between mutant load and Complex I activity in muscle and liver [3]. This is consistent with the observed marked decrease in Complex I activity in muscle, but little overall reduction in liver, despite the fact that both tissues are homoplasmic by both qPCR and targeted next-generation sequencing for the m.10134C>A mutation. The biology underpinning such tissue specific variation in enzyme activities and compensatory responses is unresolved, but is likely to be linked to the mutation’s affect on MT-ND3 activity in the context of a specific molecular background. Indeed, prior work investigating m.10191T>G has shown a reduction in Complex I catalytic activity [13], which is consistent with the results presented here that show a reduction in enzyme activity (Table 2) while protein levels are relatively unaffected (Figure 1D). Higher residual activity of Complex I in liver compared with muscle is not unique to MT-ND3 mutations. We have found similar differences with residual activity of Complex I being 2- to 5-fold higher in liver than muscle for patients with near-homoplasmic MT-ND6 mutations and in patients with autosomal recessive mutations in the NDUFS6, NDUFS8 and NDUFS1 Complex I subunit genes (unpublished data). Similarly, there does not appear to be a simple correlation as to why Complex I subunit mutations appear to induce mitochondrial proliferation in some tissues and patients but not others. While it would be interesting to study markers of mitochondrial biogenesis in patient biopsies, this is usually impractical due to the limitation in size of biopsies, at least for paediatric patients.

The clinical utility of an un-constrained high-throughput sequencing approach, particularly in cases with an unresolved diagnosis, is now clearly established. It is also evident, that the approaches used in such cases must be largely free of systematic biases. We note that in our study that the m.10134C>A mutation created a four base adenine homopolymer (Figure 1C) that resulted in a 27% indel call rate in the proband’s Ion Torrent targeted sequencing data, and masked the low-level mutation in the maternal sample (Table 1). This highlights the need for multiple validated and orthogonal sequencing approaches in clinical and diagnostic investigations.

We note that identification of the causal mutation in the proband was successful despite the initial dependence on low-coverage off-target mitochondrial sequencing reads, and potential issues stemming from nuclear copies of mitochondrial DNA (NUMTs), suggesting that next-generation sequencing approaches may be particularly useful in the diagnosis of mitochondrial disorders. This is likely to be particularly true if more targeted, e.g. directed mtDNA sequencing, or broader approaches, e.g. whole genome sequencing (WGS), as opposed to WES, are employed. Indeed, although WGS is currently cost-prohibitive for most research and diagnostic investigations, its use would have dramatically reduced the time required to resolve this case. Whole genome sequencing libraries can be generated in less than 24 hours, and running all six family members would have taken a little more than a week and yielded at least 1,000x coverage of the mitochondrial genome. We anticipate that as sequencing costs continue to fall the utility of WGS for resolution of genetically mediated mitochondrial disease will increase. Consistent with previous reports [18–20], in this case we observed ultra-low level maternal heteroplasmy of the m.10134C>A mutation and homoplasm in the proband, underscoring the importance of using high-throughput strategies for mitochondrial disease gene identification in apparently sporadic cases.

**Acknowledgments**

We thank the Queensland Center for Medical Genomics and the IMB Sequencing Core for their assistance. We also thank Wendy Fagan for assistance with enzyme assays.

**Author Contributions**

Conceived and designed the experiments: DKM CS SMG JC RJT. Performed the experiments: DKM MJM CS. Analyzed the data: DKM MJM CS LGR STC SMG DRT JC RJT. Contributed reagents/
materials/analysis tools: DKM CS SMG JC RJT. Wrote the paper: DKM MJM CS LGR STC SMG DRT JC RJT.

References