# Panel-based nuclear and mitochondrial next-generation sequencing outcomes of an ethnically diverse paediatric patient cohort with mitochondrial disease

**Running head**: NGS of paediatric patients with MD

Maryke Schoonen1\*, Izelle Smuts2, Roan Louw1, Joanna L Elson1,3, Etresia van Dyk1, Lindi-Maryn Jonck1, Richard JT Rodenburg4, Francois H van der Westhuizen1

1Human Metabolomics, North-West University, Potchefstroom, South Africa. 2Department of Paediatrics and Child Health, Steve Biko Academic Hospital, University of Pretoria, Pretoria, South Africa. 3Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, United Kingdom. 4Department of Paediatrics, Radboud Centre for Mitochondrial Medicine, Radboud University Medical Centre, Nijmegen, The Netherlands.

**Grant numbers:** This work was in part supported by the Medical Research Council of South Africa under the project title: *Investigating the aetiology of South African paediatric patients diagnosed with mitochondrial disorders*.

\*Corresponding Author:

Maryke Schoonen

Human Metabolomics

North-West University

Potchefstroom, South Africa

Tel: +27 (0)18 299 4074

Fax: +27 (0)18 299 2477

E-mail: mschoonen28@gmail.com

**Word count**:

**Text**:

**Summary**:

**Number of figures and tables**:

**Keywords**: Mitochondrial disease, OXPHOS, next generation sequencing, panel sequencing, African population, ethnically diverse

**Disclosures:** Dr. E. van Dyk contributed to next-generation sequencing. After completion of this study, she has joined ThermoFisher Scientific. The authors declare no conflict of interest or disclosures.

## Abstract

Mitochondrial disease (MD) is a group of rare inherited disorders with clinical heterogeneous phenotypes. Recent advances in next-generation sequencing (NGS) allows for rapid genetic diagnostics in patients who suffer from MD, resulting in significant strides into determining the aetiology of the disease. This, however, has not been the case in many patient populations. We report a molecular diagnostic study using targeted nuclear gene and whole mitochondrial DNA NGS of an extensive cohort of predominantly sub-Saharan African paediatric patients with clinical and biochemically defined MD. Patients in this novel cohort presented mostly with muscle involvement (n=155), developmental delay (n=141) and central nervous system involvements (n=122). Of the original 212 patients in this cohort, a muscle respiratory chain deficiency was identified in 127 cases. Genetic analyses were conducted in these 127 cases based on biochemical deficiencies, using whole mitochondrial DNA (n=123), and panel-based NGS (n=86). As a pilot investigation, whole exome NGS was conducted in a subset of patients (n=8). From these results, a previously reported pathogenic mtDNA variant in MT-ND6, and pathogenic-/likely pathogenic variants were detected in seven nuclear genes (*ETFDH, SURF1, COQ6, RYR1, STAC3, ALAS2, TRIOBP)* in only nine cases - most of which were identified via WES. This study contributes to the lacking knowledge on MD aetiology in an understudied, ethnically diverse African population, highlights inconsistencies in genotype-phenotype correlations of some variants, and proposes future directions of diagnostic approaches in such patient populations.

## Introduction

Mitochondria are ubiquitous in the human body and serve mainly as the energy producing organelle via oxidative phosphorylation (OXPHOS). This metabolic pathway comprises five protein complexes (CI-V), consisting of a total of 92 structural subunits encoded by both mitochondrial DNA (mtDNA) and nuclear DNA (nDNA) genes1, 2. Underlying genetic mutations create disruptions within this system which manifest clinically, often affecting multiple high energy demanding organs simultaneously. The heterogeneous class of clinical phenotypes associated with such mutations is collectively referred to as mitochondrial disease (MD)3. A large number of genes (at least 289), extending beyond the structural OXPHOS genes4, have been identified as being involved in MD.

Traditionally, MDs are diagnosed through extensive clinical evaluation including biochemical tissue analysis, followed by genetic screening for selected mutations. The current suite of next-generation sequencing (NGS) options available for MD diagnosis, and for heterogeneous disease diagnosis in general, includes targeted panel sequencing, unbiased whole-exome sequencing (WES), and whole genome sequencing (WGS); each with particular advantages, disadvantages and considerations5, 6. Recent publications advocate for a “genetics first” diagnostic approach, with the promise of eliminating the need for functional and biochemical analyses in the majority of diagnoses5, 6. This raises some concerns for understudied ethnically diverse populations in developing countries where relatively little progress has been made towards genetic aetiology of MD, and where the genotype-phenotype correlations are poorly understood and inconsistent. To address these limitations, we and others have undertaken various clinical, biochemical and genetic studies on MDs in the South African population — one of the few developing countries to do so7-11 (Figure S1). To date, a traditional diagnostic trajectory of extensive clinical evaluations and functional biochemical diagnosis of referred patients, followed by screening for known, common mutations has been followed12. Due to the limited information on African MD aetiology, published patient data and public genetic- and disease databases from predominantly non-African populations are currently used as reference point.

## In this study, to demonstrate these diagnostic challenges for MD in an understudied ethnically diverse population, we report the outcome of a NGS approach when targeting reported nuclear and mtDNA-encoded genes involved in MD. We tested this approach in a predominantly African cohort of 212 South African paediatric patients selected based on clinical and muscle respiratory chain enzymology data - an approach which would be considered as prudent in a diagnostic setting. We highlight the contrasting outcome of a WES approach in a small subset of this patient cohort and, considering genotype-phenotype correlations of selected cases, propose future diagnostic directions which should be considered for similar understudied population groups.



Figure S1: Diagnostic procedures followed in South African patients clinically suspected to have mitochondrial diseases.

## Materials and Methods

### Patient cohort

### Since 2006, more than 6000 patients with neurological symptoms have been referred to the Steve Biko Academic Hospital, a state funded institution in Pretoria, South Africa, and clinically evaluated according to a MD criteria (MDC) scoring system set forth by Wolf and Smeitink (2002)13, which was further refined by Smuts et al. (2010)7. Currently this cohort consists of 212 paediatric patients who manifested clinically with MD signs or symptoms from as early as the neonatal period. This cohort originated from the northern provinces of South Africa and is predominantly African (64%), with an equal number of males and females. Urine and muscle (*vastus lateralis)* samples were collected from the entire cohort for subsequent biochemical and molecular genetic investigations. Limited availability of samples from parents and patients prevented segregation analysis. This study was approved by the Ethics Committees of the University of Pretoria (No. 91/98 and amendments) and the North-West University (NWU-00170-13-A1).

### Biochemical analyses

Muscle RC enzyme analyses for complexes I-IV (CI-IV, EC 1.6.5.3, EC 1.3.5.1, EC 1.10.2.2, EC 1.9.3.1, respectively), as well as CII+CIII (EC 1.3.5.1+1.10.2.2) were performed and normalised against citrate synthase (CS, EC 2.3.3.1) activity for all 212 patients. This was done in 600 x g homogenates prepared from frozen muscle samples, as previously described7, 14. Other biochemical analyses performed included urine metabolic analysis7, 9, 15, and muscle CoQ10 analysis16 in muscle samples. In total, 127 patients were identified to have a RC deficiency.

### Genetic analyses

#### mtDNA sequencing

Genomic DNA (gDNA) preparation from muscle homogenate was performed using standard protocol, as previously described8. The Qubit 2.0 Fluorometer (ThermoFisher Scientific, MA, USA) was used for quantification of gDNA. The complete mitochondrial genome was sequenced in 123 patients, all of whom had a known combined clinical and biochemical MD profile. Two chemistries were used, i.e. 71 patient samples were sequenced using the 454 GS-FLX platform, and 52 patient samples were sequenced using the Ion PGM™ platform (ThermoFisher Scientific, MA, USA). The first set of sequencing, including library preparation and enrichment was done according to procedures described by van der Walt et al. (2012)8. Sequencing using Ion Torrent technology, including library preparation and enrichment were done according to the manufacturer’s protocol for Ion PGM™. Due to insufficient sample quantity and quality, four patient samples could not be sequenced.

#### Nuclear panel sequencing

For nuclear gene investigations, Ion Torrent™ amplicon panel sequencing was performed on 86 patients, all of whom had a known combined clinical and biochemical MD profile. Patients were included for sequencing based on their MDC score evaluations for clinical presentations and the severity of their biochemical deficiencies, i.e. patients having low enzyme activity, which was lower than the reference range when expressed against at least two of three enzyme markers (CS, CII, or CIV), were considered7. Genes to be included in sequencing panels were selected based on their mitochondrial RC involvement, either direct or indirect (Table S1-3). Three custom gene panels consisting of 136 genes in total were designed and are briefly explained: Panel 1, consisted of 78 targeted genes associated with CI deficiency (HaloPlex™ Target Enrichment System, Agilent Technologies, CA, USA, Table S1). This customised panel had a target region size of 360 091 kbp and 99.6% targeted coverage. In total, 30 patients were sequenced using Panel 1. Panel 2, from Ion AmpliSeq™ Custom DNA Panels (Thermo Fisher Scientific), had 78 targeted genes associated with CI–IV deficiency (CI=38, CII=6, CIII=10, CIV=24), with a total target region size of 157 834 kbp and 98.2% targeted coverage (Table S2). In total, 48 patients were sequenced using Panel 2, of which five patients overlapped with Panel 1, and 10 patients overlapped with Panel 3. Panel 3, from Ion AmpliSeq™, targeted 18 genes known to be involved with primary and secondary CoQ10 deficiency (Table S3). The design size was 61 kbp with a targeted coverage of 98%. In total, 26 patients were sequenced in Panel 3 (6 patients overlapping with Panel 1). The entire coding regions of each gene, including flanking regions of intron-exons were sequenced using the Ion PGM™ platform as per manufacturers’ protocol (HaloPlex ref G9912C and Ampliseq ref 4480441). The selected genes and panels were not African population-specific, as the underlying genetic cause for MD is mostly unclear in South African populations.

#### Whole exome sequencing

As an initial comparison on the outcomes of a panel vs. WES approach, WES was performed on a subset of eight randomly selected African (Haplogroup L) cases where no strong candidate disease-causing variants were identified by initial mtDNA and/or nuclear panel sequencing. WES was performed at the Central Analytical Facilities, Stellenbosch University, South Africa, using the Ion Proton™ sequencer (ThermoFisher Scientific, MA, USA) according to manufactures protocol for Ion Torrent platform™. A ~95% on target coverage was achieved with average depth coverage of ~140.

### Bioinformatics for mtDNA variants

Mitochondrial DNA sequences were aligned against the human mitochondrial revised Cambridge Reference Sequence (rCRS [NC\_012920 gi:251831106]). Haplogroup assignment, variant identification and annotation was performed using mtDNA-Server (v1.20.0)17, MitoMap and MitoMaster18. Homoplasmic and heteroplasmic (levels above 30%) non-synonymous variants were further evaluated based on their allele frequency reported in Genbank19, 20 and appearance on Phylotree21, 22 and those with a population allele frequency below 0.1% were considered significant. Pathogenicity of variants were furthermore evaluated according to the criteria proposed for mtDNA variants23-25; for example the MutPred scoring system was used as supporting evidence to interpret and classify non-synonymous variants in structural subunits of OXPHOS ([*http://mutpred.mutdb.org/help.html*](http://mutpred.mutdb.org/help.html)*,* last accessed April 2018)26. A MutPred score above 0.5 suggests a probable damaging impact on protein function. Scores between 0.75–1.0 indicate functional damage on a protein/amino acid with high confidence. Mitochondrial-tRNA variants were individually evaluated using MitoTIP27 and classified according to a scoring system set forth by McFarland et al. (2004)24, which was refined by Yarham et al. (2011)28. A low Yarham score (below 10) weighs more towards benign or neutral classification, while a score above 10, with substantial evidence from functional tests, weighs more toward a pathogenic classification. Variants were also evaluated using the The American College of Medical Genetics and Genomics (ACMG) guidelines29 where possible.

### Bioinformatics for nDNA variants

Raw sequencing files, obtained from the Ion PGM™, were analysed with the *Torrent Suite* (v5.0.2). The sequence files were aligned against Genome Reference Consortium Human Build 37 (GRCh37, hg19) followed by coverage analysis and variant calling using the *coverageAnalysis* and *variantCaller* plugins (v.5.0) from the Torrent Suite, respectively. The Variant Calling Format (VCF) files were further annotated using the offline *Variant Effect Predictor* from Ensembl.org (v81; last accessed July 2018)30, followed by variant mining using *GEMINI* (v.20)31. The output text files generated from GEMINI contained information on the novel, reported and previously reported pathogenic variants. Further filtering of variants was done using population databases such as Exome Aggregation Consortium(ExAC) and *gnomAD*32 (specifically African allele frequencies), disease-specific databases such as ClinVar and OMIM, and sequence databases such as NCBI Genome and RefSeqGene. As supporting evidence, the missense variants of interest were cautiously evaluated using various *in silico* predictive algorithms (*SIFT*, *Polyphen-2,* and *CADD* )33-35. These algorithms, however, have been shown to have low sensitivity, specificity and accuracy36. Candidate variants of interest were evaluated using ACMG guidelines29 and are classified as pathogenic, likely pathogenic, variants of uncertain significance, likely benign, and benign.

## Results

### Clinical profiles

From the more than 6000 neurological patients referred for clinical assessment, 212 patients (113 male, 99 female) presented with mitochondrial phenotypes, forming the cohort that is described here. These clinically defined MD patients were comprehensively evaluated since 2006 and were predominantly of African ancestry (n=130, Figure 1A). Age of onset presented as early as the first year of life including the neonatal period (n=139, Figure 1B). The most common clinical finding was muscle involvement, which manifested in 73% (n=155) of the cohort. Cardiac involvement and deafness was the least observed at 5% (n=10) and 8% (n=17), respectively (Figure 1C).

### Biochemical profiles

An RC enzyme complex deficiency was identified in 127 cases (64 males and 63 females), where 64 cases had isolated deficiency and the remaining 63 had a combined deficiency. Complex I deficiency, either isolated (n=43) or combined (n=32) was most prevalent, as is frequently reported37, 38, followed by CIII deficiency (n=54, isolated=15 and combined=39), as seen in Figure 2B.

### Genetic analyses

The findings for whole mtDNA, panel nuclear DNA, and whole exome NGS are summarised in Table 1. Whole mtDNA sequencing was successfully applied in 123 cases (79 African and 44 non-African) and a variant of interest was identified in 12 African cases, all of which could not be classified as disease-causing8. An example of one of these reported pathogenic variants, the well-known Leber Hereditary Optic Neuropathy (LHON, OMIM #535000)39 mutation, m.14484T>C, was identified in an African male (S014) who did not present with typical LHON symptoms and was later clinically diagnosed with fucosidosis. An additional 11 mtDNA variants of interest identified in 10 cases, as listed in Table S4, may still prove to have a functional effect on disease initiation or progression. Notably this may still be the case for variants with a comparatively high population allele frequency (above 0.1%) since further investigation may indeed reveal that these variants appear at a higher frequency in these understudied population lineages.

Panel nuclear DNA sequencing of 86 cases (61 African and 25 non-African) revealed four variants of interest in two cases. Two pathogenic compound heterozygous *ETFDH* variants were identified in one case (S057) and two possible pathogenic compound heterozygous *SURF1* variants were identified in the second case (S085). WES sequencing of eight African cases revealed nine variants of interest in six cases, where eight variants in the genes *COQ6,* *RYR1, STAC3,* and *ALAS2* are considered possibly pathogenic and one variant in the gene *TRIOBP* is considered a variant of uncertain significance when applying the ACMG classification.

#### Pathogenic variants

The *ETFDH* (electron-transfer flavoprotein dehydrogenase) compound heterozygous variants (c.1448C>T and c.1067G>A ENST00000307738) were identified in a non-African female (S057) who presented with features of multiple acyl-dehydrogenase deficiency (MADD, MIM #231680). She had severe muscle weakness, exercise intolerance, chronic fatigue and hepatomegaly. Metabolic markers of MADD, such as dicarboxylic acids, ethylmalonic acid, glutaric acid as well as acylcarnitines (butyryl-, isovaleryl- and glutarylcarnitine) and acylglycine (hexanoyl- isobutyryl-, isovaleryl and suberylglycine) conjugates were observed in the patient’s urine. Furthermore, she had CI, CIII and CII+CIII RC deficiency with severely reduced CoQ10 levels in muscle16. These mutations cause structural instability in ETFDH and were confirmed with functional analysis. This and two other cases have recently been described elsewhere40.

#### Likely pathogenic variants

An African female (S085), with clinically confirmed Leigh Disease (LD, MIM #256000) and confirmed mitochondrial CIV deficiency, harboured compound heterozygosity for a missense and frameshift variant (c.575G>A andc.754\_755delAG, ENST00000371974,) in exons 6 and 8 in the gene *SURF1* (Surfeit 1). The patient clinically presented early in life with several CNS involvements (nystagmus, extrapyramidal and pyramidal symptoms) and muscle involvements (myopathy, hypotonia and weakness). Furthermore, changes in the basal ganglia/thalami were observed. Metabolic investigation revealed elevated lactic acid.

An African female (case S002) with confirmed primary CoQ10 deficiency16 was identified carrying compound heterozygous variants (c.41G>A, c.859G>T, ENST00000394026) in the gene *COQ6* (coenzyme Q10 monooxygenase 6). One of the variants identified in a highly conserved region, c.41G>A, results in a premature truncation of the protein with high loss-of-function (LoF) probability. Clinically she presented at four years and four months of age with a history of severe weakness since birth. Other features were noted upon clinical examination and included macrocephaly, severe hypotonia and head lag, pseudo-hypertrophy of calves and triceps, limitation of extension at knees and elbows, and reduced reflexes. SDS-PAGE analysis revealed lower COQ6 protein content when compared to control samples, confirming COQ6 instability16. No other candidate *COQ6* gene variants were identified in this case.

Four variants, associated with minicore myopathy and external ophthalmoplegia (MMEO, MIM #255320) were identified in the gene *RYR1* (ryanodine 1) in two African females (case S032 and case S033) using WES. Wilmshurst et al. (2010)41 identified several African population-specific pathogenic variants. Two of these disease-causing variants were identified here and are considered likely to be pathogenic in our cases. The first reported pathogenic variant41, c.8342\_8343delTA with high LoF confidence as well as a missense variant, c.11926C>T, was identified in the first case, S032. Both variants were detected as heterozygous and are in highly conserved regions. She presented with severe hypotonia, myopathy with myopathic facial features and initially did not show signs of external ophthalmoplegia. She also had an affected sibling with myopathic facial features and external ophthalmoplegia. A muscle biopsy was collected from the sibling who revealed thickened connective tissue between muscle fibers and evidence of fat infiltration. The female reported here had CIV RC deficiency and her brother had CI+CIII deficiency. Both siblings presented with an additional CII+CIII deficiency. The second pathogenic variant (c.14524G>A) with a second splice donor variant (c.11193+1G>A) was identified in the second female (S033), both of which were detected as heterozygous. She presented with severe hypotonia, mild myopathic facial features and dense external ophthalmoplegia.

A homozygous mutation, c.851G>C (p.Trp284Ser, ENST00000332782), was detected by WES in an African female (S011) in the gene *STAC3* (SH3 and cysteine-rich domains 3)and was first described by Horstick et al. (2013)42 in five families with Native American Myopathy (NAM, MIM #255995)*.* The female described here, had correlating phenotypes as the case describe elsewhere42 and presented with severe myopathy, failure to thrive, developmental delay, relative macrocephaly, and ptosis with no external ophthalmoplegia. She had minor dysmorphic features including a low nasal bridge. She was born prematurely and had intra uterine growth restriction. Biochemically she presented with isolated CIII deficiency in muscle.

An African male (S117) with skin and muscle involvement carried a homozygous gain-of-function variant (c.1757A>T, ENST00000330807) in exon 11 of the gene *ALAS2* (5'-aminolevulinate synthase 2)43. This variant was first described in a Spanish patient with erythropoietic porphyria by To-Figueras et al. (2011)43, who states that *ALAS2* acts as a modifier gene in patients with erythropoietic porphyria (X-linked protoporphyria, MIM #30075252). Another manifestation is sideroblastic anaemia type 1 (MIM #300751), which is also as a result of *ALAS2* mutations. However, our patient did not present with symptoms for the latter. He instead presented with swelling of his face, hands and feet and experienced non-specific body pain, symptoms which are more similar with X-linked protoporphyria. He had depigmented skin lesions in his face and on the extensor areas of the fore and upper arms, as well as over the knees and lateral right thigh. Furthermore, he suffered from severe muscle weakness with decreased muscle bulk in all four limbs and muscle histology revealed atypical dermatomyositis. Biochemically he had confirmed CI, CIII and CIV deficiency in muscle.

#### Variants of uncertain significance

A *TRIOBP* (trio- and filamentous-actin-binding protein) homozygous missense variant (c.3232C>T, ENST00000406386) was identified using WES in an African male (S059) who presented with developmental delay, visual impairment, muscle weakness and hypotonia and clinodactyly. Furthermore, he presented with mild facial dysmorphisms, which included an epicanthic fold and low set ears. Most notably the patient had hearing impairment, which was confirmed by abnormal auditory brainstem response. Mutations in this gene are known to cause autosomal recessive deafness (MIM #609823), a feature found in our case presented here. Metabolic profiles revealed significant ketosis associated with dicarboxylic aciduria involving C6–C10 acids (i.e. adipic-, suberic- and sebacic acid), with a normal amino acid profile.

A number of cases (13) presented with variants of uncertain significance and are listed in Table S5. These nuclear variants are classified as likely-pathogenic or pathogenic according to ClinVar but did not adhere to the ACMG standards and guidelines. For example, previously reported pathogenic variants identified in the genes *NDUFA9, SDHA*, *SDHB,* and *POLG* were all detected as heterozygous, while homozygous variants in the genes *TRMU, ACADVL*, and *GLUD2* had high African allele frequencies. These variants however still remain of interest for South African population groups.

## Discussion and conclusions

The genetic diagnosis of MD, and identification of the number of genes involved, has rapidly improved since the first mutations were reported in the late 1980’s44, 45. While clinical scoring such as MDC in addition to biochemical evaluation of RC/OXPHOS function in tissue is still the hallmark of MD diagnostics, it has been increasingly complimented by NGS in recent years, notably exome sequencing46. The main advantage of using NGS in MD diagnoses is that it allows for the discrimination between primary (having a direct genetic aetiology) and secondary MD (caused by non-genetic factors such as environmental toxins)47. A retrospective investigation into genetic causes of MD was conducted for 127 patients with clinically suspected and biochemically confirmed RC deficiency in an understudied population (predominantly African). Here, we report on two high-throughput NGS techniques (whole mtDNA sequencing and panel sequencing) which were used to find common, previously reported pathogenic or likely pathogenic variants in reported genes involved in MD in this understudied ethnically diverse cohort with mostly unknown MD genotype-phenotype correlations.

The mtDNA sequencing data for a sub-section of this cohort have been extensively investigated elsewhere8. The majority of these variants could not be classified as pathogenic due to a number of mtDNA criteria which were not met. Most notably among the observed variants was a well-known LHON-associated mutation (m.14484T>C at 53% heteroplasmy in muscle), identified in an African male (S014). Although this patient had clinical features of eye involvement, his clinical phenotype did not match that expected for LHON (visual failure and optic atrophy). This inconsistency could be ascribed to varied penetrance of the disease, highlighting the importance of investigating the penetrance of this and other pathogenic variants when detected in new populations.

Panel sequencing revealed pathogenic and likely pathogenic variants in two cases: S057 and S085. For S057, compound heterozygous variants were identified in *ETFDH* (c.1448T>C and c.1067G>A). This case has previously been investigated extensively, leading to the classification of these variants as pathogenic with substantial clinical, biochemical, and *in vivo* supporting experimental evidence (from structural and functional analyses)40. For S085, compound heterozygous variants were identified in *SURF1* (c.754\_755delAG and c.575G>A). SURF1 is directly involved with cytochrome c oxidase (COX) maintenance and assembly. LoF mutations in *SURF1* cause major structural instability in COX, and are responsible for the phenotype of LD as clinically diagnosed in the case reported here. The frameshift variant has previously been reported as pathogenic for a different (Japanese) patient, with substantial clinical and experimental supporting evidence and a confirmed LoF mechanism48. Consequently, we classify this variant, which has a low African allele frequency, as likely pathogenic in this case. The missense variant has not previously been associated with a clinical phenotype and the extremely low allele frequency suggests moderate likelihood to be pathogenic.

Since the panel sequencing of genes known to be involved in MD revealed a pathogenic or likely pathogenic variant in only two cases (S057 and S85) of the 86 patients investigated, it was evident that this is not an effective approach to follow in this patient population. Although targeted/gene panel NGS is considered a prudent NGS approach in many diagnostic settings, whilst being aware that this is a heterogeneous, understudied patient population and that an expanded gene panel may not necessarily increase diagnostic yield49, we expanded our genetic investigations to include WES to probe its outcome on a small subset of eight African patients where no initial NGS results were obtained. A variant of interest in the genes *COQ6, RYR1, STAC3, ALAS2,* and *TRIOBP* was identified in six of the eight cases.

The compound heterozygous *COQ6* variants (c.41G>A and c.859G>T) identified in S002 have been extensively investigate elsewhere16, where functional and structural analyses showed significantly decreased levels of COQ6. This protein is directly involved with CoQ10 biosynthesis and mutations in this gene result in primary CoQ10 deficiency50, 51. This correlates well with the clinically and biochemically profiles observed in the case reported here. Based on these experimental findings, the correlation between the observed and reported phenotypes, their allele frequencies and benign classification from several disease databases, these variants are classified as likely pathogenic. It is likely that the disease penetrance in this patient is higher than observed in other individuals.

Two previously reported pathogenic and two reported *RYR1* variants were identified in two compound heterozygotes, S032(c.14524G>A and c.8342\_8343delTA) and S033 (c.11926C>T and c.11193+1G>A). The c.14524G>A and c.11926C>T variants are classified as founder mutations for South African patient populations with centronuclear myopathy (CNM) and minicore myopathy with external ophthalmoplgia (MMEO)42. *RYR1* encodes a homotetrameric calcium channel in skeletal muscle and regulates cytosolic Ca2+ levels52. Dysfunctional RYR1 disrupts the Ca2+ balance, directly affecting different mitochondrial functions such as ATP synthesis regulation and reactive oxygen species generation41, 53, 54, consequently contributing to myopathy and external ophthalmoplegia and ptosis. The two cases reported here had consistent clinical phenotypes to reported cases. For S032 and her brother, the specific type of congenital myopathy is still unclear. Both had neurogenic features which are absent in multi-minicore myopathy. The siblings, whose features correlated better with CNM, showed similar, but also slight differences in their manifestations. For example, the female initially did not have external ophthalmoplegia, while the brother did. Importantly, no homozygous variants have been reported in patients with CNM55. S033 had correlating multi-minicore myopathic features as seen in previously reported cases41, 56 and included mild myopathic facial features with dense external ophthalmoplegia. Both sets of compound heterozygous variants are classified as likely pathogenic in these cases due to substantial supporting clinical and experimental evidence confirming the LoF mechanism41, 56, confirmed founder effect for South African populations with a low or absent allele frequency in several population databases, and multiple *in silico* algorithms supporting a deleterious effect on the gene product for the two frameshift variants.

The previously reported pathogenic *STAC3* variant (c.851G>C)42 was identified as homozygous in S011. *STAC3* encodes a putative muscle-specific adaptor protein, which takes part in excitation-contraction in muscle. Disruptions within this protein are thought to cause reduction in mitochondrial Ca2+, either intra- or extra mitochondrial,which have a direct effect on OXPHOS52. Clinical features, as a result, include congenital myopathy with facial dysmorphic features including ptosis which is consistent with the case reported here. This variant, c.851G>C, was sequenced as part of the 1000 Genomes Project and was undetected in 113 Caucasian controls. Consequently, we classify this variant as likely pathogenic in this case due to previously confirmed pathogenicity with substantial supporting evidence, and a low African allele frequency,

The homozygous variants identified in *ALAS2* (c.1757A>T) and *TRIOBP* (c.3232C>T) in S117 and S059 are classified as variants of uncertain significance and benign or likely benign respectively according to several disease databases. The two cases reported here, S117 and S059, had overlapping but also slight inconstant phenotypical manifestations compared to reported cases.

S117, did not present with typical X-linked protoporphyria43, 57, the primary phenotype associated with *ALAS2* mutations. His primary cause of disease is interplay between two deficiencies; dermatomyositis and X-linked porphyria. The muscle histology from this patient was suggestive of an inflammatory myopathy, but the inflammatory cell infiltrate was perivascular and the typical pattern of peripheral muscle cell atrophy in dermatomyositis was not demonstrated. As the patient had definitive skin involvement, which included swelling and redness, it is suggestive that the patient has porphyria rather than inflamed myopathy. His mitochondrial dysfunction could arise from impaired electron transport as a direct result of bone marrow haem synthetic dysfunction. The clinical significance is still unclear and therefore this variant is classified as a variant of uncertain significance. The variant shows evidence to be likely pathogenic with substantial supporting experimental pathogenicity evidence43, and low African allele frequencies.

S059 had a severe mitochondrial phenotype, where deafness was most notable, The latter is consistent with reported cases harbouring *TRIOBP* mutations58, 59, however the former has not yet been associated with mutations in this gene. The underlying genetic cause for this patient’s mitochondrial phenotype is still unclear. This variant, c.3232C>T, have not been associated with a clinical phenotype and therefor lacks substantial supporting pathogenicity evidence. Consequently, we classify this variant as a variant of uncertain significance.

To conclude, in this unique and predominantly African cohort, by firstly looking at nuclear and mitochondrial genes known to be involved in MD - which is in line with current diagnostic practices - only a relatively small number of pathogenic or likely pathogenic variants could be identified, which is clearly a poor diagnostic outcome. Initial indications from limited WES data are much more promising as nine likely pathogenic variants were identified in six cases using WES compared to five variants identified in three cases using panel NGS and mtDNA sequencing.

 Considering all the cases in this study, in only five cases (S057, S085, S032, S033 and S011) a strong genotype-phenotype correlation could be established and a moderately strong correlation in two cases (S002, S117). . For the remaining twocases S014 and S059, a non-specific correlation was observed. Observations like these serve as a strong motivation that a “genetics first/only” approach without supporting clinical and biochemical investigation is not suitable in such understudied, ethnically diverse populations. Furthermore, when following a genetic approach, we concur that panel sequencing could be an efficient approach in populations where the genotype-phenotype correlations are well-established for specific monogenic diseases. For heterogeneous diseases such as MD, even in such populations, WES/WGS compares significantly better compared to a targeted gene-panel approach4, 5, 60, 61 . Our results are thus in line with proposals that WES should be considered as the primary option for genetic investigations in heterogeneous inherited diseases such as MD and, in fact, may be particularly ideal in understudied, ethnically diverse populations where there is evidence of inconsistencies with documented MD phenotypes. However, in such populations the value of extensive clinical and biochemical (structural and functional) investigations to support molecular genetic data outcomes, should not be neglected and at this time are, in fact, more crucial than in well studied populations.

## Acknowledgments

This work was in part supported by the Medical Research Council of South Africa under project title: *Investigating the aetiology of South African paediatric patients diagnosed with mitochondrial disorders*

## References

1. Chinnery PF, Hudson G: Mitochondrial genetics. British medical bulletin 2013, 106:135-159.

2. Anderson S, Bankier AT, Barrell BG, De Bruijn M, Coulson AR, Drouin J, Eperon I, Nierlich D, Roe BA, Sanger F: Sequence and organization of the human mitochondrial genome. Nature 1981, 290:457-465.

3. Gorman GS, Chinnery PF, DiMauro S, Hirano M, Koga Y, McFarland R, Suomalainen A, Thorburn DR, Zeviani M, Turnbull DM: Mitochondrial diseases. Nat Rev Dis Primers 2016, 2:16080.

4. Wortmann SB, Mayr JA, Nuoffer JM, Prokisch H, Sperl W: A Guideline for the Diagnosis of Pediatric Mitochondrial Disease: The Value of Muscle and Skin Biopsies in the Genetics Era. Neuropediatrics 2017, 48:309-314.

5. Wortmann SB, Koolen DA, Smeitink JA, van den Heuvel L, Rodenburg RJ: Whole exome sequencing of suspected mitochondrial patients in clinical practice. J Inherit Metab Dis 2015, 38:437-443.

6. Craven L, Alston CL, Taylor RW, Turnbull DM: Recent Advances in Mitochondrial Disease. Annu Rev Genomics Hum Genet 2017, 18:257-275.

7. Smuts I, Louw R, Du Toit H, Klopper B, Mienie LJ, van der Westhuizen FH: An overview of a cohort of South African patients with mitochondrial disorders. Journal of inherited metabolic disease 2010, 33:95-104.

8. van der Walt EM, Smuts I, Taylor RW, Elson JL, Turnbull DM, Louw R, van Der Westhuizen FH: Characterization of mtDNA variation in a cohort of South African paediatric patients with mitochondrial disease. European Journal of Human Genetics 2012, 20:650-656.

9. Reinecke CJ, Koekemoer G, Van der Westhuizen FH, Louw R, Lindeque JZ, Mienie LJ, Smuts I: Metabolomics of urinary organic acids in respiratory chain deficiencies in children. Metabolomics 2012, 8:264-283.

10. Meldau S, De Lacy R, Riordan G, Goddard E, Pillay K, Fieggen K, Marais D, Van der Watt G: Identification of a single MPV17 nonsense‐associated altered splice variant in 24 South African infants with mitochondrial neurohepatopathy. Clinical genetics 2018.

11. Van der Watt G, Owen EP, Berman P, Meldau S, Watermeyer N, Olpin SE, Manning NJ, Baumgarten I, Leisegang F, Henderson H: Glutaric aciduria type 1 in South Africa-high incidence of glutaryl-CoA dehydrogenase deficiency in black South Africans. Mol Genet Metab 2010, 101:178-182.

12. van der Westhuizen FH, Sinxadi PZ, Dandara C, Smuts I, Riordan G, Meldau S, Malik AN, Sweeney MG, Tsai Y, Towers GW: Understanding the implications of mitochondrial DNA variation in the health of black southern African populations: The 2014 Workshop. Human mutation 2015, 36:569-571.

13. Wolf NI, Smeitink JA: Mitochondrial disorders A proposal for consensus diagnostic criteria in infants and children. Neurology 2002, 59:1402-1405.

14. Smith Pe, Krohn RI, Hermanson G, Mallia A, Gartner F, Provenzano M, Fujimoto E, Goeke N, Olson B, Klenk D: Measurement of protein using bicinchoninic acid. Analytical biochemistry 1985, 150:76-85.

15. Venter L, Lindeque Z, van Rensburg PJ, Van der Westhuizen F, Smuts I, Louw R: Untargeted urine metabolomics reveals a biosignature for muscle respiratory chain deficiencies. Metabolomics 2015, 11:111-121.

16. Louw R, Smuts I, Wilsenach K-L, Jonck L-M, Schoonen M, van der Westhuizen FH: The dilemma of diagnosing coenzyme Q10 deficiency in muscle. Molecular genetics and metabolism 2018.

17. Weissensteiner H, Forer L, Fuchsberger C, Schopf B, Kloss-Brandstatter A, Specht G, Kronenberg F, Schonherr S: mtDNA-Server: next-generation sequencing data analysis of human mitochondrial DNA in the cloud. Nucleic Acids Res 2016, 44:W64-69.

18. Lott MT, Leipzig JN, Derbeneva O, Xie HM, Chalkia D, Sarmady M, Procaccio V, Wallace DC: mtDNA variation and analysis using MITOMAP and MITOMASTER. Current protocols in bioinformatics 2013:1.23. 21-21.23. 26.

19. Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW: GenBank. Nucleic acids research 2012, 41:D36-D42.

20. van Oven M: PhyloTree Build 17: Growing the human mitochondrial DNA tree. Forensic Science International: Genetics Supplement Series 2015, 5:e392-e394.

21. Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW: GenBank. Nucleic acids research 2017, 45:D37.

22. Van Oven M, Kayser M: Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Human mutation 2009, 30.

23. DiMauro S, Schon EA: Mitochondrial DNA mutations in human disease. American Journal of Medical Genetics Part A 2001, 106:18-26.

24. McFarland R, Elson JL, Taylor RW, Howell N, Turnbull DM: Assigning pathogenicity to mitochondrial tRNA mutations: when ‘definitely maybe’is not good enough. TRENDS in Genetics 2004, 20:591-596.

25. Mitchell AL, Elson JL, Howell N, Taylor RW, Turnbull DM: Sequence variation in mitochondrial complex I genes: mutation or polymorphism? Journal of medical genetics 2006, 43:175-179.

26. Li B, Krishnan VG, Mort ME, Xin F, Kamati KK, Cooper DN, Mooney SD, Radivojac P: Automated inference of molecular mechanisms of disease from amino acid substitutions. Bioinformatics 2009, 25:2744-2750.

27. Sonney S, Leipzig J, Lott MT, Zhang S, Procaccio V, Wallace DC, Sondheimer N: Predicting the pathogenicity of novel variants in mitochondrial tRNA with MitoTIP. PLoS computational biology 2017, 13:e1005867.

28. Yarham JW, Al‐Dosary M, Blakely EL, Alston CL, Taylor RW, Elson JL, McFarland R: A comparative analysis approach to determining the pathogenicity of mitochondrial tRNA mutations. Human mutation 2011, 32:1319-1325.

29. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E: Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in medicine 2015, 17:405.

30. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, Flicek P, Cunningham F: The ensembl variant effect predictor. Genome biology 2016, 17:122.

31. Paila U, Chapman BA, Kirchner R, Quinlan AR: GEMINI: integrative exploration of genetic variation and genome annotations. PLoS computational biology 2013, 9:e1003153.

32. Karczewski KJ, Weisburd B, Thomas B, Solomonson M, Ruderfer DM, Kavanagh D, Hamamsy T, Lek M, Samocha KE, Cummings BB: The ExAC browser: displaying reference data information from over 60 000 exomes. Nucleic acids research 2016, 45:D840-D845.

33. Ng PC, Henikoff S: SIFT: Predicting amino acid changes that affect protein function. Nucleic acids research 2003, 31:3812-3814.

34. Adzhubei I, Jordan DM, Sunyaev SR: Predicting functional effect of human missense mutations using PolyPhen‐2. Current protocols in human genetics 2013:7.20. 21-27.20. 41.

35. Kircher M, Witten DM, Jain P, O'roak BJ, Cooper GM, Shendure J: A general framework for estimating the relative pathogenicity of human genetic variants. Nature genetics 2014, 46:310.

36. Ernst C, Hahnen E, Engel C, Nothnagel M, Weber J, Schmutzler RK, Hauke J: Performance of in silico prediction tools for the classification of rare BRCA1/2 missense variants in clinical diagnostics. BMC medical genomics 2018, 11:35.

37. McFarland R, Kirby DM, Fowler KJ, Ohtake A, Ryan MT, Amor DJ, Fletcher JM, Dixon JW, Collins FA, Turnbull DM: De novo mutations in the mitochondrial ND3 gene as a cause of infantile mitochondrial encephalopathy and complex I deficiency. Annals of neurology 2004, 55:58-64.

38. Kirby DM, Salemi R, Sugiana C, Ohtake A, Parry L, Bell KM, Kirk EP, Boneh A, Taylor RW, Dahl H-HM: NDUFS6 mutations are a novel cause of lethal neonatal mitochondrial complex I deficiency. The Journal of clinical investigation 2004, 114:837-845.

39. Brown M, Voljavec A, Lott M, MacDonald I, Wallace D: Leber's hereditary optic neuropathy: a model for mitochondrial neurodegenerative diseases. The FASEB Journal 1992, 6:2791-2799.

40. van der Westhuizen FH, Smuts I, Honey E, Louw R, Schoonen M, Jonck L-M, Dercksen M: A novel mutation in ETFDH manifesting as severe neonatal-onset multiple acyl-CoA dehydrogenase deficiency. Journal of the Neurological Sciences 2017.

41. Wilmshurst J, Lillis S, Zhou H, Pillay K, Henderson H, Kress W, Müller C, Ndondo A, Cloke V, Cullup T: RYR1 mutations are a common cause of congenital myopathies with central nuclei. Annals of neurology 2010, 68:717-726.

42. Horstick EJ, Linsley JW, Dowling JJ, Hauser MA, McDonald KK, Ashley-Koch A, Saint-Amant L, Satish A, Cui WW, Zhou W: Stac3 is a component of the excitation–contraction coupling machinery and mutated in Native American myopathy. Nature communications 2013, 4:1952.

43. To-Figueras J, Ducamp S, Clayton J, Badenas C, Delaby C, Ged C, Lyoumi S, Gouya L, de Verneuil H, Beaumont C: ALAS2 acts as a modifier gene in patients with congenital erythropoietic porphyria. Blood 2011, 118:1443-1451.

44. Wallace DC, Singh G, Lott MT, Hodge JA, Schurr TG, Lezza A, Elsas LJ, Nikoskelainen EK: Mitochondrial DNA mutation associated with Leber's hereditary optic neuropathy. Science 1988, 242:1427-1430.

45. Holt I, Harding A, Morgan-Hughes J: Deletions of muscle mitochondrial DNA in patients with mitochondrial myopathies. Nature 1988, 331:717-719.

46. Witters P, Saada A, Honzik T, Tesarova M, Kleinle S, Horvath R, Goldstein A, Morava E: Revisiting mitochondrial diagnostic criteria in the new era of genomics. Genetics in Medicine 2018, 20:444.

47. Niyazov DM, Kahler SG, Frye RE: Primary mitochondrial disease and secondary mitochondrial dysfunction: importance of distinction for diagnosis and treatment. Molecular syndromology 2016, 7:122-137.

48. Tanigawa J, Kaneko K, Honda M, Harashima H, Murayama K, Wada T, Takano K, Iai M, Yamashita S, Shimbo H, Aida N, Ohtake A, Osaka H: Two Japanese patients with Leigh syndrome caused by novel SURF1 mutations. Brain Dev 2012, 34:861-865.

49. Plutino M, Chaussenot A, Rouzier C, Ait-El-Mkadem S, Fragaki K, Paquis-Flucklinger V, Bannwarth S: Targeted next generation sequencing with an extended gene panel does not impact variant detection in mitochondrial diseases. BMC medical genetics 2018, 19:57.

50. Alcázar-Fabra M, Trevisson E, Brea-Calvo G: Clinical syndromes associated with Coenzyme Q10 deficiency. Essays in biochemistry 2018, 62:377-398.

51. Heeringa SF, Chernin G, Chaki M, Zhou W, Sloan AJ, Ji Z, Xie LX, Salviati L, Hurd TW, Vega-Warner V: COQ6 mutations in human patients produce nephrotic syndrome with sensorineural deafness. The Journal of clinical investigation 2011, 121:2013-2024.

52. Blackburn PR, Selcen D, Gass JM, Jackson JL, Macklin S, Cousin MA, Boczek NJ, Klee EW, Dimberg EL, Kennelly KD: Whole exome sequencing of a patient with suspected mitochondrial myopathy reveals novel compound heterozygous variants in RYR1. Molecular genetics & genomic medicine 2017, 5:295-302.

53. Görlach A, Bertram K, Hudecova S, Krizanova O: Calcium and ROS: a mutual interplay. Redox biology 2015, 6:260-271.

54. Zhou H, Jungbluth H, Sewry CA, Feng L, Bertini E, Bushby K, Straub V, Roper H, Rose MR, Brockington M: Molecular mechanisms and phenotypic variation in RYR1-related congenital myopathies. Brain 2007, 130:2024-2036.

55. Abath Neto O, Martins CdA, Carvalho M, Chadi G, Seitz KW, Oliveira ASB, Reed UC, Laporte J, Zanoteli E: DNM2 mutations in a cohort of sporadic patients with centronuclear myopathy. Genetics and molecular biology 2015, 38:147-151.

56. Punetha J, Kesari A, Uapinyoying P, Giri M, Clarke NF, Waddell LB, North KN, Ghaoui R, O’Grady GL, Oates EC: Targeted re-sequencing emulsion PCR panel for myopathies: results in 94 cases. Journal of neuromuscular diseases 2016, 3:209-225.

57. Balwani M, Bloomer J, Desnick R, of the NIH-Sponsored PC, Network RDCR: X-linked protoporphyria. 2013.

58. Shahin H, Walsh T, Sobe T, Rayan AA, Lynch ED, Lee MK, Avraham KB, King M-C, Kanaan M: Mutations in a novel isoform of TRIOBP that encodes a filamentous-actin binding protein are responsible for DFNB28 recessive nonsyndromic hearing loss. The American Journal of Human Genetics 2006, 78:144-152.

59. Riazuddin S, Khan SN, Ahmed ZM, Ghosh M, Caution K, Nazli S, Kabra M, Zafar AU, Chen K, Naz S: Mutations in TRIOBP, which encodes a putative cytoskeletal-organizing protein, are associated with nonsyndromic recessive deafness. The American Journal of Human Genetics 2006, 78:137-143.

60. Neveling K, Feenstra I, Gilissen C, Hoefsloot LH, Kamsteeg EJ, Mensenkamp AR, Rodenburg RJ, Yntema HG, Spruijt L, Vermeer S, Rinne T, van Gassen KL, Bodmer D, Lugtenberg D, de Reuver R, Buijsman W, Derks RC, Wieskamp N, van den Heuvel B, Ligtenberg MJ, Kremer H, Koolen DA, van de Warrenburg BP, Cremers FP, Marcelis CL, Smeitink JA, Wortmann SB, van Zelst-Stams WA, Veltman JA, Brunner HG, Scheffer H, Nelen MR: A post-hoc comparison of the utility of sanger sequencing and exome sequencing for the diagnosis of heterogeneous diseases. Hum Mutat 2013, 34:1721-1726.

61. Haack TB, Haberberger B, Frisch E-M, Wieland T, Iuso A, Gorza M, Strecker V, Graf E, Mayr JA, Herberg U: Molecular diagnosis in mitochondrial complex I deficiency using exome sequencing. Journal of medical genetics 2012, 49:277-283.

## Footnote’s Page

**Abbreviations:**

MD Mitochondrial disease

CI to CIV Respiratory chain enzyme complexes I to IV, respectively

NGS Next generation sequencing

OXPHOS Oxidative phosphorylation

mtDNA mitochondrial DNA

nDNA nuclear DNA

WES Whole-exome sequencing

RC Respiratory chain

CNS Central nervous system

CS Citrate synthase

VCF Variant Calling Format

ExAC Exome Aggregation Consortium

SIFT Sorting Intolerant From Tolerant

CADD Combined Annotation Dependent Depletion

LoF Loss-of-function

*MT-ND6* NADH-ubiquinone oxidoreductase chain 6

CoQ10 Coenzyme Q10

*ETFDH* Electron transfer flavoprotein-ubiquinone oxidoreductase

*SURF1* Surfeit locus protein 1

*COQ6* Coenzyme Q6, monooxygenase

*STAC3* SH3 and cysteine rich domain 3

*RYR1* Ryanodine receptor 1

*TRIOBP* TRIO and F-actin binding protein

*ALAS2* 5'-aminolevulinate synthase 2

**Conflict of interest**

All authors declare that they have no conflict of interest

**Ethical approval:**

This study was ethically approved by the Ethics Committees of the University of Pretoria (No. 91/98 and amendments) and the North-West University (NWU-00170-13-A1). Informed parental consent and assent, if applicable, were obtained.

**List of figures:**

Figure 1: Demographic information for 212 paediatric patients. A:Summary of the ethnicity distribution. Caucasian, Asian, and Coloured patients are collectively referred to as non-African patients in text. B: Summary of the age categories representing the age at symptom onset for each of the 212 patients. C:Summary of the clinical features and finding in patient cohort.

Figure 2: Summary for the biochemical RC enzyme deficiency distribution. A: Combined versus isolated enzyme deficiency for 127 patients. B: Summary for CI-CIV enzyme deficiency in 64 patients with isolated RC deficiency.

Figure S1: Summary of the general diagnostic workflow that was followed to establish MD aetiology in an ethnically diverse South African paediatric cohort.

**List of tables:**

Table 1: Summary of mtDNA and nDNA variants of interest identified in South African patients with clinically confirmed mitochondrial disease

Table S1: Targeted panel 1

Table S2: Targeted panel 2

Table S3: Targeted panel 3

Table S4: Mitochondrial DNA variants of importance that is less likely to be disease-causing

Table S5: Nuclear DNA variants of importance that is less likely to be disease-causing

Table 1: Summary of mtDNA and nDNA variants of interest identified in South African patients with clinically confirmed mitochondrial disease

|  |
| --- |
| ***Mitochondrial variants*** |
| **Gene** | **Position** | **Amino acid change** | **Mutpred score** | **ACMG classification\*** | **GenBank frequency, variant allele frequency** | **Patient, ethnicity, gender** | **Clinical profile** | **Biochemical profile** | **Reference** |
| *MT-ND6* | m.14484T>C | p.Met64Val | 0.627 | Likely benign | 0.19, 53% | S014, A, M | DR, CNS, Eye, M, L | CI | Brown et al. (1992) 39 |
| ***Nuclear variants*** |
| **Gene** | **Variant position** | **Amino acid change** | **RefSNP ID** | **ACMG classification\*** | **ExAC frequency** | **Patient, ethnicity, gender** | **Clinical profile** | **Biochemical profile** | **Reference** |
| *ETFDH*† | c.1448C>T (+/-) | p.Pro483Leu | rs377656387 | Pathogenic | None | S057, NA, F | DD, CNS, M, L, S | CI, CIII, CIV, CII+CIII | van der westhuizen et al. (2017)40 |
| c.1067G>A (+/-) | p.Gly356Glu | Novel | Pathogenic | None |
| *SURF1*† | c.575G>A (+/-) | p.Arg192Gln | rs782021521 | Likely pathogenic | 0.00001 | S085, A, F | FTT, DD, DR, CNS, M, E | CIV | Tanigawa et al. (2012)48 |
| c.754\_755delAG (+/-) | p.Ser252HisfsTer39 | rs782007828 | Likely pathogenic | 0.00002 |
| *COQ6*†† | c.41G>A (+/-) | p.Trp14Ter | rs17094161 | Likely pathogenic / risk factor | 0.07047 | S002, A, F | Mac, DD, Eye, M | CII, CII+CIII | Louw et al. (2018)16 |
| c.859G>T (+/-) | p.Ala49Ser | rs61743884 | 0.03162 |
| *RYR1*†† | c.8342\_8343delTA (+/-) | p.Ile2781ArgfsX49 | rs758580075 | Likely pathogenic | None | S032, A, F | FTT, DD, Eye, M | CIV, CII+CIII | Wilmshurst et al. (2010)41 |
| c.11926C>T (+/-) | p.His3976Tyr | rs148772854 | Likely pathogenic | 0.01287 |
| c.14524G>A (+/-) | p.Val4842Met | rs193922879 | Likely pathogenic | 0.001935 | S033, A, F | FTT, DD, Eye, M, R | CI | Punetha et al. (2016)56Wilmshurst et al. (2010)41 |
| c.11193+1G>A (+/-) | - | rs111986316 | Likely pathogenic | None |
| *STAC3*†† | c.851G>C (+/+) | p.Trp284Ser | rs140291094 | Likely pathogenic | 0.0011 | S011, A, F | Mac, FTT, DD, Dys, Eye, M, R | CIII | Horstick et al. (2013)42 |
| *ALAS2*†† | c.1757A>T (+/+) | p.Tyr586Phe | rs139596860 | Likely pathogenic / VUS | 0.009364 | S117, A, M | M, Skin | CI, CIII, CIV | Balwani et al. (2013)57To-Figueras et al. (2011)43 |
| *TRIOBP*†† | c.3232C>T (+/+) | p.Arg1078Cys | rs200359708 | VUS | 0.00030 | S059, A, M | DD, CNS, Eye, D, M, GIT | CI, CIII, CIV, CII+CIII | None to date |

*\*ACMG classifications based on patient evaluations.*

*† Variants identified using panel NGS*

*††Variants identified using WES*

Mac: Macrocephaly; FTT: Failure to thrive; DD: Developmental delay; DR: Developmental regression; CNS: Central nervous system involvement; Eye: Eye involvement; D: Sens. Deafness; M: Muscle involvement; GIT: Gastro-intestinal involvement; R: Renal involvement; C: Cardiac involvement; E: Endocrine abnormalities; L: Liver involvement; S: Skeletal involvement; Skin: Skin involvement; VUS: variant of uncertain significance.