APPLICATION NOTE

SOPHiA DDM™ Platform for Exome Analysis

Mitochondrial genome analysis with SOPHiA DDM™ Whole-exome Sequencing

Rare diseases are diverse and heterogeneous, affecting multiple organ systems. Whole-exome sequencing (WES) is often the most appropriate next-generation sequencing (NGS) technique for the identification of variants associated with diverse and overlapping phenotypes in rare diseases, including mitochondrial diseases, especially in combination with familial variant analysis (FVA) to determine inheritance patterns.

SOPHiA GENETICS™

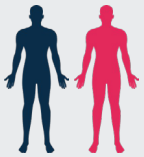
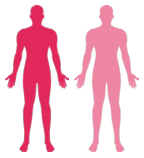
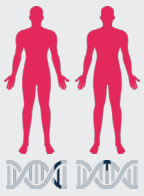
© SOPHiA GENETICS 2023- GL-GN-2200110-r1

Table of Contents

A brief introduction to mitochondrial diseases

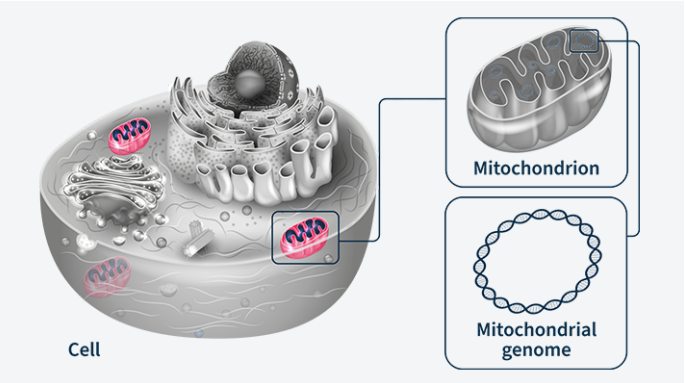
Mitochondrial diseases are the largest class of inborn errors of metabolism, affecting 1.6 per 5,000 individuals,[1] a substantial portion of the rare disease population. Mitochondrial disorders are a clinically heterogeneous group of disorders typically associated with an overlapping spectrum of phenotypes,[2,3] making WES an appropriate first genetic test in the search for a causative variant, which could be located in either the nuclear or mitochondrial genome. Phenotypes associated with these disorders can involve the brain and nervous system, muscles, heart, endocrine, or exocrine systems, with typical symptoms including mental disorders, vision loss, hearing disorders, weakness, hypotonia, cardiomyopathy, and cardiac arrhythmia.

Genetic variants associated with mitochondrial diseases have several attributes that contribute to the observed broad clinical spectrum:[3]

Diverse penetrance	Specific mutations in mitochondrial genes may cause a disease phenotype in some individuals but not others.	
Variable expressivity	The same mutations in mitochondrial genes may result in different disease severity in different patients.	
Genetic and allelic heterogeneity	Similar phenotypes can be caused by mutations in different mitochondrial or nuclear genes, or the same variant can have different clinical manifestations.	

These attributes are driven by unknown genetic, environmental, and lifestyle factors.

A brief introduction to the mitochondrial genome



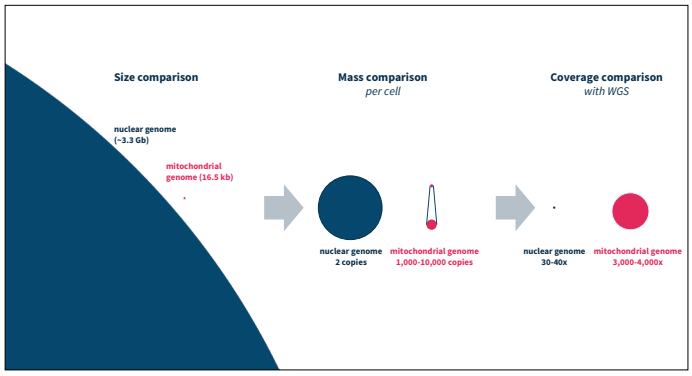
Distinct mtDNA genetic code

	Translation	
Codon	Nuclear DNA	mtDNA
AUA	Isoleucine	Methionine
UGA	Stop	Tryptophan
AGA	Arginine	Stop
AGG	Arginine	Stop

Unlike the nuclear genome, the mitochondrial genome is circular, haploid, and there are many copies per cell.[4,5] In addition, mitochondrial DNA (mtDNA) concentration can vary substantially across tissues, with those requiring more energy (such as the heart and brain) typically containing more mtDNA per cell. The mitochondrial genome comprises 16,569 base pairs arranged into 37 genes which encode 2 rRNAs and 22 tRNAs, along with 13 proteins involved in oxidative phosphorylation. mtDNA is gene dense (93% of the sequence is coding) with genes having no introns, and the genetic code differs slightly from that of the nuclear genome. The mitochondrial genome is maternally inherited and mutates faster than the nuclear genome, potentially due to the close proximity to reactive-oxygen species and/or higher error rates of mitochondrial DNA polymerase.

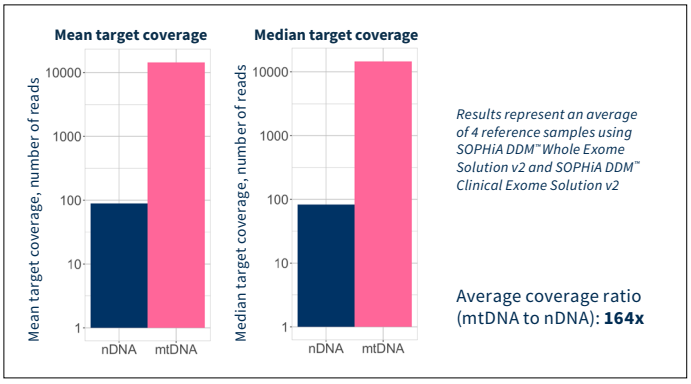
Solving the unique challenges associated with genotyping mtDNA

mtDNA copy number exceeds nuclear DNA by several-fold



Although the mitochondrial genome is a lot smaller than the nuclear genome (16.5 kb vs ~3.3 Gb), there are as many as 10,000 copies per cell, compared with 2 copies of the nuclear genome.[6] This means that in a blood sample, **coverage of the mitochondrial genome is ~3-4,000x compared with 30-40x for the nuclear genome**.[7] This coverage difference can be even greater (up to 10,000-fold) in samples from mitochondria-dense tissues. High copy numbers can pose challenges, particularly when sequencing both the nuclear and mitochondrial genomes in a single assay.

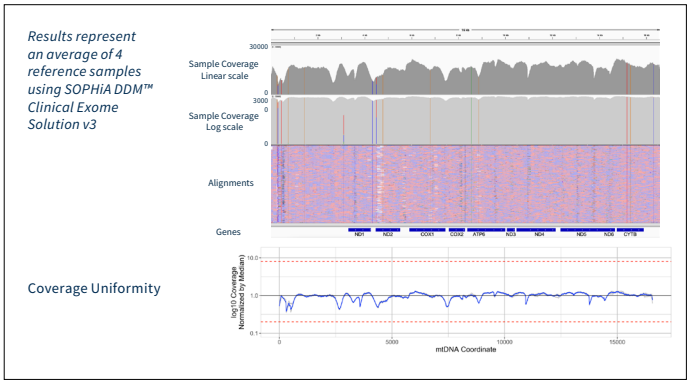
SOPHiA GENETICS™ hybridization capture enrichment technology allows for balanced coverage of mtDNA and nuclear DNA (nDNA) by mixing mtDNA- and nDNA-targeting probes at a ratio that allows for >10x higher coverage of mtDNA. This results in mtDNA accounting for 1-5% of reads from blood-derived DNA.



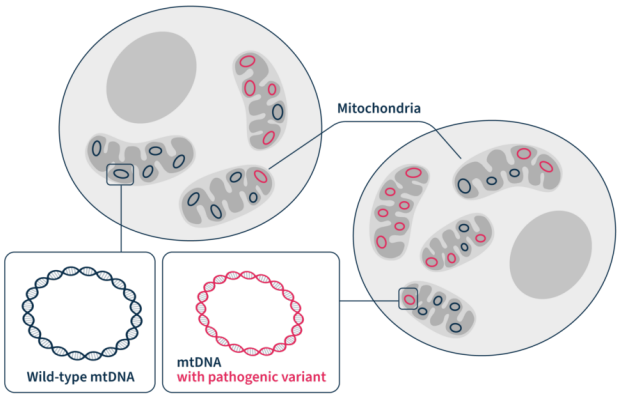
Nuclear mitochondrial DNA sequences can confound mtDNA mapping

Nuclear mitochondrial DNA sequences (NUMTs) are nDNA regions that are highly homologous to mtDNA regions. The high homology is because they originated from the transposition of cytoplasmic mtDNA into the nuclear genome. NUMTs are difficult to distinguish from mtDNA during sequencing.

The SOPHiA GENETICS™ robust probe design limits the capture of homologous mtDNA-like nDNA regions, supporting uniform coverage of the entire mitochondrial genome despite known NUMTs.

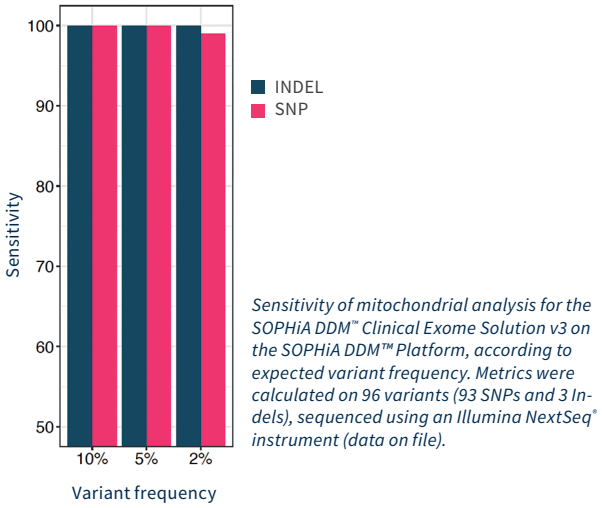


mtDNA exhibits heteroplasmy



Heteroplasmy is the presence of multiple non-identical mitochondrial genomes, resulting in variation in the abundance of pathogenic variants across mitochondria, cells, and tissues. In combination with the size of an mtDNA deletion, heteroplasmy level has a significant role in defining disease phenotype and clinical progression.

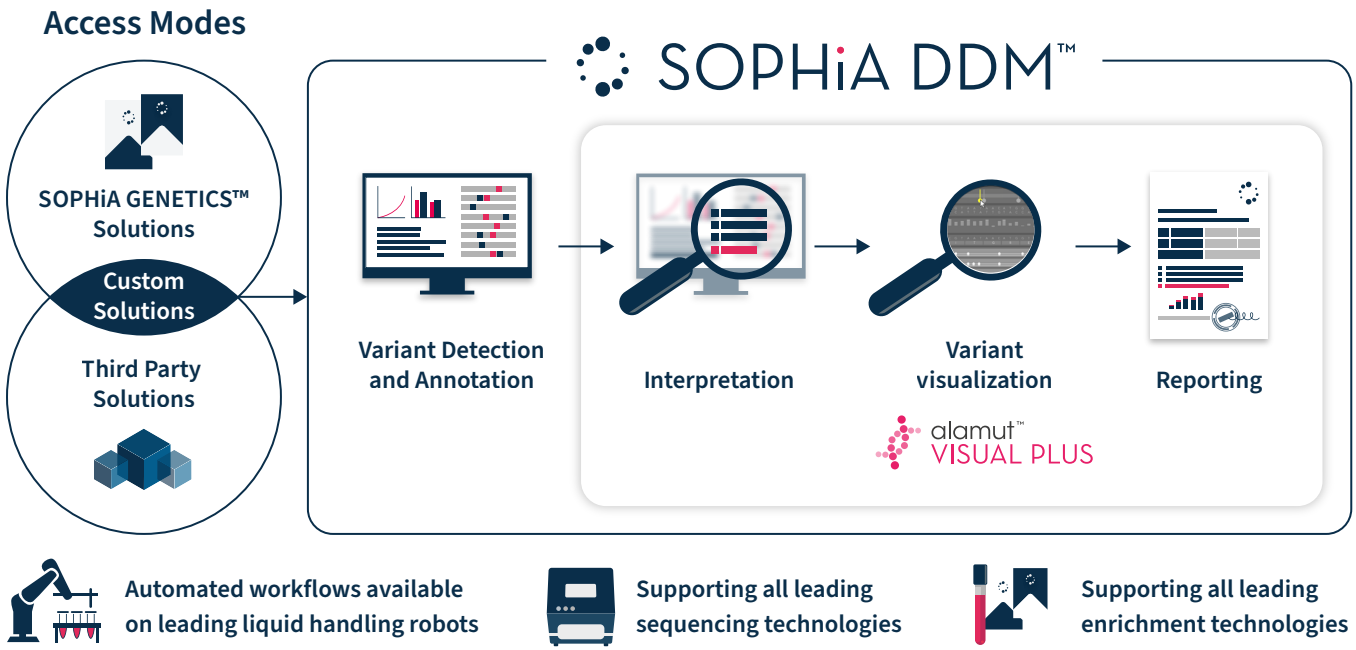
As heteroplasmy means that mtDNA variants may be present at any fraction, SOPHiA GENETICS™ utilize somatic-like variant calling in germline exome solutions to give 100% sensitivity for mitochondrial SNVs/Indels down to 5% variant frequency. The variant fraction annotation in SOPHiA DDM™acts as a measure of heteroplasmy.



Sensitivity of mitochondrial analysis for the SOPHiA DDM™ Clinical Exome Solution v3 on the SOPHiA DDM™ Platform, according to expected variant frequency. Metrics were calculated on 96 variants (93 SNPs and 3 Indels), sequenced using an Illumina NextSeq® instrument (data on file).

SOPHiA DDM™ exome sequencing workflow for mitochondrial diseases

The SOPHiA DDM™ Platform complemented by Alamut™ Visual Plus offers streamlined WES workflows (from FASTQ file to variant report) for the efficient identification of genomic variants associated with rare and inherited diseases. In a single workflow, the SOPHiA DDM™ Platform analyzes variants in mtDNA alongside SNVs, Indels, and CNVs in both coding and disease-relevant non-coding regions of nuclear DNA (nDNA).

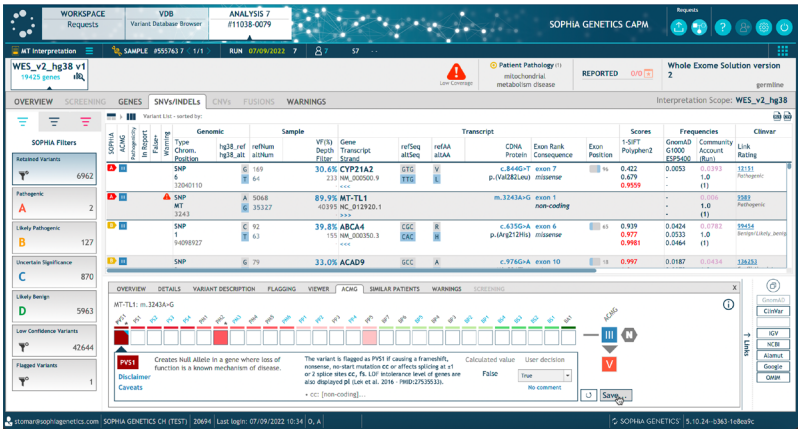


Here, we demonstrate how the analytical capabilities and advanced features of the SOPHiA DDM™ platform can be used in combination with Alamut™ Visual Plus to identify and interpret variants in mitochondrial DNA as part of a WES workflow.

In this research case, a 5-year-old boy, Stephen, had presented with Leigh Syndrome-like symptoms including epilepsy, severe dystonia, global developmental delay, hip subluxation, and a history of rhabdomyolysis. His parents were clinically normal. Stephen's brain CT scan portrayed diffuse mild hypodensity in the region of the left medial and lateral lenticulostriate arteries, suspicious for areas of ischaemia/acute infarcts. This also involved regions of the left caudate head, globus pallidus and putamen. The MRI/MRS scan showed bilateral basal ganglia signal abnormality with associated elevated lactate levels: lactate 1.7, glucose 3.8, protein 0.32, WBC 5, and RBC 3. Due to Stephen's overlapping spectrum of phenotypes potentially indicating Leigh syndrome, it was determined that he was a good research candidate for SOPHiA DDM™ WES v2 including mitochondrial analysis.

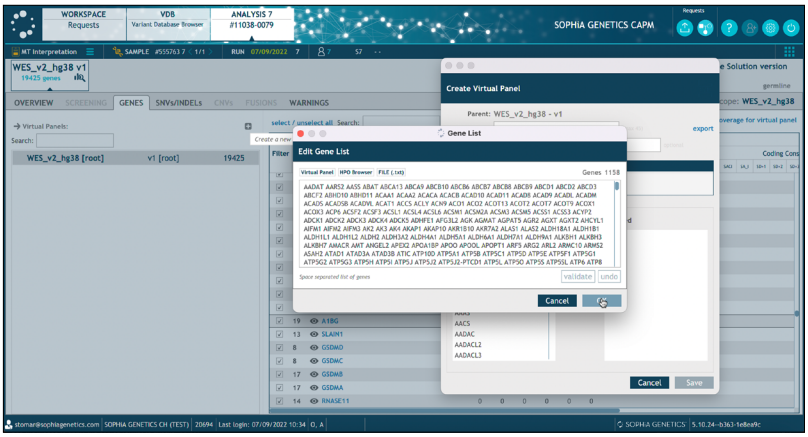
1. Get a list of the sample's genomic variants

The SOPHiA DDM™ Platform analyzes the VCF file from WES v2 and provides a table of the called variants in the SNVs/INDELs and CNVs tabs, with the most likely pathogenic variants at the top of the lists. Variants in mtDNA are labelled with "MT". The variant lists can be adapted to show information of interest such as variant fraction, transcript information, missense and splicing predictor scores, population frequency, and links out to databases such as ClinVar. Due to the proband having a suspected mitochondrial disease, we proceed to explore mitochondrial variants in both nDNA and mtDNA.



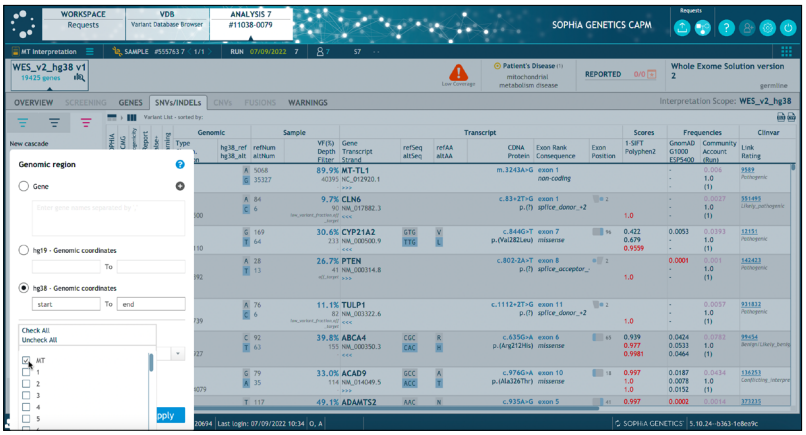
2. Filter mitochondrial variants in nDNA using Virtual Panels

First, we focus on nuclear genes that encode mitochondrial proteins. We create a Virtual Panel in SOPHiA DDM™ by clicking on the Settings button, the "Virtual Panels" tab, and the + in the grey square in the top-right corner. After clicking on "Gene List", we have the option to directly access the human phenotype ontology (HPO) browser and search for the term "mitochondrial" to automatically capture all genes related to mitochondrial phenotypes. In this instance, we choose to copy the list of 1136 mitochondrial localization genes that encode mitochondrial proteins from the Human MitoCarta3.0 download into the Virtual Panel.



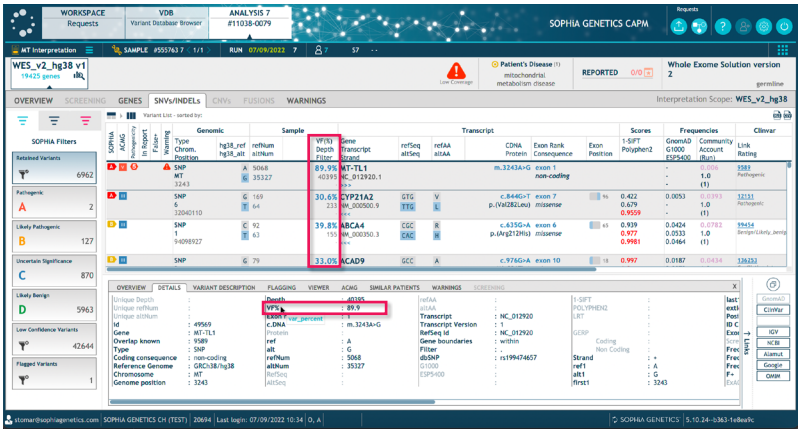
3. Filter mtDNA variants using Cascading Filters

To analyze variants in the mitochondrial genome, we use cascading filters (pink filter icon) to select for coding variants in any of the 37 mitochondrial genes. To do this, we click "+ Add more filters", "Genomic region", and then select the MT chromosome and "Apply" to limit analysis to variants in mitochondrial genes only. Variants are automatically prioritized by the algorithms in SOPHiA DDM™ and according to ACMG scoring. The top candidate in the filtered list of variants is m.3243A>G in the *MT-TL1* gene.



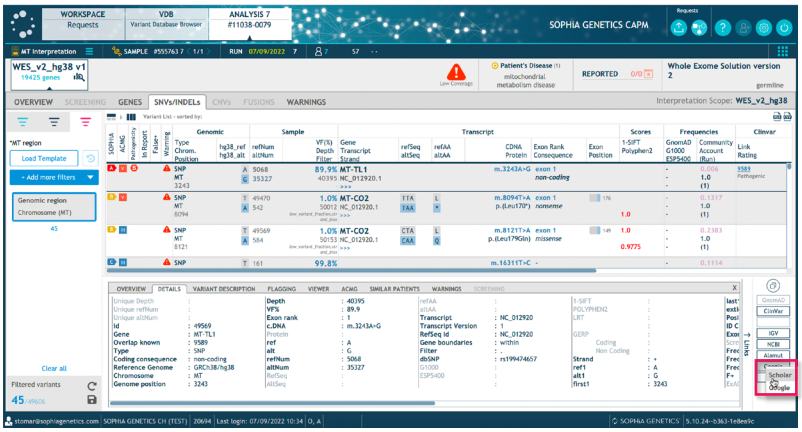
4. Use variant fraction as a measure of heteroplasmy

Clicking on the suspected variant in the SNVs\Indels or CNVs tab reveals additional details in a panel at the bottom of the screen, such as depth of coverage and variant fraction (VF). For the analysis of variants in the mitochondrial genome, we can use the VF annotation as a measure of heteroplasmy, which may be related to the severity of the observed phenotype. Our candidate mitochondrial variant (m.3243A>G in the *MT-TL1* gene) has 89.9% VF, and is flagged for further investigation.



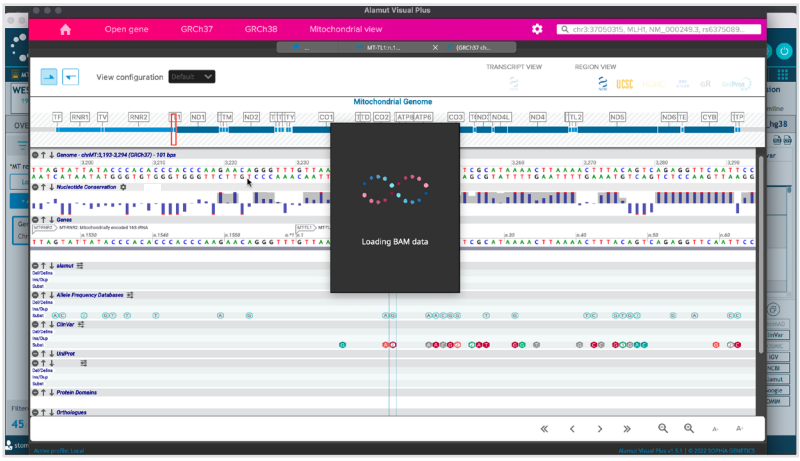
5. Link directly to curated databases for supportive literature

In the variant panel of our candidate mutation, it is possible to link out to key databases such as Google Scholar. When clicking on “Google” and “Scholar”, the Google Scholar search bar is automatically populated with information about our candidate variant, providing a comprehensive list of all relevant publications. These publications can be explored to gain further information about the prevalence, functional impact, and disease heterogeneity associated with this variant. This *MT-TL1* variant represents the most common heteroplasmic mtDNA disease genotype and is associated with a broad range of clinical features, making disease prognosis extremely difficult to predict.[8]



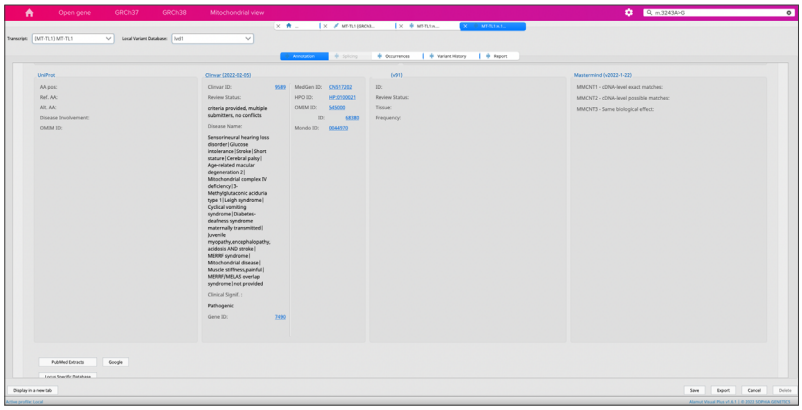
6. Link directly to Alamut™ Visual Plus to visualize the BAM data

Clicking on the “Alamut” button takes you to Alamut™ Visual Plus for further analysis of the variant and coverage data. There is the option to open the software “With” or “Without” the BAM file; directly loading the BAM data from SOPHiA DDM™ avoids the risk of manual errors. We use Alamut™ Visual Plus to view the comprehensive variant annotations in a full-genome browser for deep exploration of the suspected variant in the mitochondrial genome.



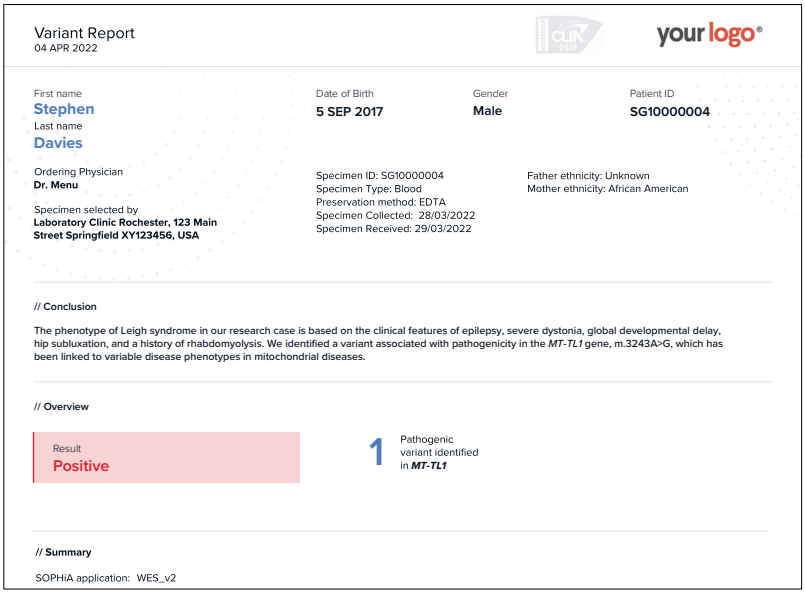
7. Access comprehensive variant annotation information in Alamut™ Visual Plus

The m.3243A>G variant in the *MT-TL1* gene was further investigated by opening the variant viewer in Alamut™ Visual Plus. The “Annotation” tab of the variant viewer contains detailed information such as pathogenicity and disease information reported in ClinVar and population frequency reported in gnomAD. Links are also provided to other curated databases of gene- and disease-specific information, such as MedGen, HPO, OMIM, and Mondo. Together, the data support that this variant in *MT-TL1* is associated with Stephen’s phenotype (Research Use Only).



8. Create a comprehensive genomic report

To complete the analysis, we add our suspected variant, along with its supporting references, to a comprehensive germline genomic report. We can customize the report to show our lab contact information and logo, which will be displayed alongside details of Stephen’s phenotype, the pathogenic variant m.3243A>G in *MT-TL1*, the mitochondrial disease manifestation, and the sequencing platform used, supplemented by any additional information from the variant annotation, or through additional experiments, such as mtDNA testing in urine and muscle samples, or Sanger confirmation (when feasible).



Variant Report
04 APR 2022

your logo

First name
Stephen
Last name
Davies

Date of Birth
5 SEP 2017

Gender
Male

Patient ID
SG10000004

Ordering Physician
Dr. Menu

Specimen selected by
Laboratory Clinic Rochester, 123 Main Street Springfield XY123456, USA

Specimen ID: SG10000004
Specimen Type: Blood
Preservation method: EDTA
Specimen Collected: 28/03/2022
Specimen Received: 29/03/2022

Father ethnicity: Unknown
Mother ethnicity: African American

// Conclusion

The phenotype of Leigh syndrome in our research case is based on the clinical features of epilepsy, severe dystonia, global developmental delay, hip subluxation, and a history of rhabdomyolysis. We identified a variant associated with pathogenicity in the *MT-TL1* gene, m.3243A>G, which has been linked to variable disease phenotypes in mitochondrial diseases.

// Overview

Result
Positive

1 Pathogenic variant identified in ***MT-TL1***

// Summary

SOPHiA application: WES_v2

Conclusion

In this research case, the pathogenic variant m.3243A>G in the *MT-TL1* gene was associated with Stephen's Leigh-Syndrome-like symptoms. The variant annotations in SOPHiA DDM™ and Alamut™ Visual Plus reveal that it is associated with a wide range of clinical features. The m.3243A>G mutation has been linked to a Leigh Syndrome-like phenotype and ~15% of probands present with mitochondrial encephalopathy, lactic acidosis, and mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes (MELAS); the remainder display a wide variety of clinical features such as diabetes, deafness, ataxia, and isolated myopathy.[8] The m.3243A>G mutation encodes mitochondrial tRNAL$^{eu(UUR)}$, with the A>G change in the D-loop domain of the tRNA leading to the reduction of mitochondrial DNA (mtDNA)-encoded proteins and oxidative phosphorylation activity. The m.3243A>G *MT-TL1* variant (rs199474657) is rare (<0.1%, gnomAD: 6/56383 total alleles) and has been reported in ClinVar. The presence of other nuclear genetic factors has been demonstrated to influence clinical outcomes in m.3243A>G-related diseases,[8] providing rationale for further genomic analysis.

The comprehensive SOPHiA DDM™ whole-exome sequencing workflow for mitochondrial variant analysis, complemented by Alamut™ Visual Plus, enabled the efficient identification of a variant associated with a mitochondrial disease, which as a group are phenotypically and genetically diverse. Both nuclear and mitochondrial genomes were analyzed and the level of **heteroplasmy** was determined in a single experiment. In combination with the comprehensive variant annotation and enhanced variant visualization in both **nuclear and mitochondrial genomes**, the efficient sample-to-report workflow in SOPHiA DDM™ and Alamut™ Visual Plus effectively led to a quick conclusion.

References

1. Schlieben LD, Prokisch H. Front Cell Dev Biol. 2020;8:600079.
2. Russell O, Turnbull D. Exp Cell Res. 2014;325(1):38-43.
3. Gorman GS, Schaefer AM, Ng Y, et al. Ann Neurol. 2015;77(5):753-759.
4. Anderson S, Bankier AT, Barrell BG, et al. Nature 1981;290:457-65.
5. Scitable by Nature Education. mtDNA and Mitochondrial Diseases. https://www.nature.com/scitable/topicpage/mtdna-and-mitochondrial-diseases-903/. Accessed Oct 2022.
6. Data on file, SOPHiA GENETICS, 2022.
7. Davis RL, Raj R Kumar K, Puttick C, et al. Neurology 2022;99(7):e730-e742.
8. Pickett SJ, Grady JP, Ng YS, et al. Ann Clin Transl Neurol. 2018;5(3):333-345.

About SOPHiA GENETICS

SOPHiA GENETICS (Nasdaq: SOPH) is a software company dedicated to establishing the practice of data-driven medicine as the standard of care and for life sciences research. We are the creator of the SOPHiA DDM™ Platform, a cloud-native platform capable of analyzing data and generating insights from complex multimodal data sets and different diagnostic modalities. The SOPHiA DDM™ Platform and related applications, modules, and services are currently used by a broad network of hospital, laboratory, and biopharma institutions globally. **Where others see data, we see answers.**

Want to know more?

Contact us at:

info@sophiagenetics.com

For Research Use Only. Not for use in diagnostic procedures.

© 2023 SOPHiA GENETICS. All rights reserved. All trademarks are the property of SOPHiA GENETICS and/or its affiliate(s) in the U.S. and/or other countries. All other names, logos, and other trademarks are the property of their respective owners