The Australian Genomics Mitochondrial Flagship: A National Program Delivering Mitochondrial Diagnoses.

Rocio Rius[1+], Alison G. Compton[2+], Naomi L. Baker[3], Shanti Balasubramaniam[4], Stephanie Best[5], Kaustuv Bhattacharya[6], Kirsten Boggs[7], Tiffany Boughtwood[8], Jeffrey Braithwaite[9], Drago Bratkovic[10], Alessandra Bray[11], Marie-Jo Brion[12], Jo Burke[13], Sarah Casauria[14], Belinda Chong[15], David Coman[16], Shannon Cowie[17], Mark Cowley[18], Michelle G. de Silva[19], Martin B. Delatycki[20], Samantha Edwards[21], Carolyn Ellaway[22], Michael C. Fahey[23], Keri Finlay[24], Janice Fletcher[25], Leah E. Frajman[26], Ann E. Frazier[27], Velimir Gayevskiy[2]8, Roula Ghaoui[29], Himanshu Goel[30], Ilias Goranitis[31], Matilda Haas[32], Daniella H. Hock[33], Denise Howting[34], Matilda R. Jackson[35], Maina P. Kava3[6], Madonna Kemp[37], Sarah King-Smith[38], Nicole J. Lake[39], Phillipa J. Lamont[40], Joy Lee[41], Janet C. Long[42], Mandi MacShane[43], Evanthia O. Madelli[44], Ellenore M. Martin[45], Justine E. Marum[4]6, Tessa Mattiske[47], Jim McGill[48], Alejandro Metke[49], Sean Murray[50], Julie Panetta[51], Liza K. Phillips[52], Michael C.J. Quinn[53], Michael T. Ryan[54], Sarah Schenscher[55], Cas Simons[56], Nicholas Smith[57], David A. Stroud[58], Michel C. Tchan[59], Melanie Tom[60], Mathew Wallis[61], Tyson L. Ware[62], AnneMarie E. Welch[63], Christine Wools[64], You Wu[65], John Christodoulou[66*], David R. Thorburn[67*]

+ Contributed equally * Corresponding

[1]Centre for Population Genomics, Garvan Institute of Medical Research, Sydney, Australia; Murdoch Children's Research Institute, Melbourne, Australia; The University of Melbourne, Melbourne, Australia
[2]Murdoch Children's Research Institute, Melbourne, Australia; The University of Melbourne, Melbourne, Australia; Victorian Clinical Genetics Services, Melbourne, Australia
[3]The University of Melbourne, Melbourne, Australia; Victorian Clinical Genetics Services, Melbourne, Australia
[4]Sydney Children's Hospitals Network, Westmead, Australia; University of Sydney, Sydney, Australia
[5]Australian Institute of Health Innovation, Macquarie University, Sydney, Australia; Peter MacCallum Cancer Centre, Melbourne, VIC, Australia; Victorian Comprehensive Cancer Centre, Melbourne, VIC, Australia; Australian Genomics, Murdoch Children's Research Institute, Melbourne, VIC, Australia; University of Melbourne, Melbourne, VIC, Australia
[6]Sydney Children's Hospitals Network, Westmead, Australia
[7]Sydney Children's Hospitals Network, Westmead, Australia
[8]Australian Genomics; Murdoch Children's Research Institute, Melbourne, Australia
[9]Australian Institute of Health Innovation, Macquarie University, Sydney, Australia
[10]Women's and Children's Hospital, Adelaide, Australia
[11]Sydney Children's Hospitals Network, Westmead, Australia
[12]Australian Genomics; QIMR Berghofer Medical Research Institute, Brisbane, Australia
[13]Tasmanian Clinical Genetics Service, Hobart, Australia;
The University of Tasmania, Hobart, Australia
[14]Australian Genomics; Murdoch Children's Research Institute, Melbourne, Australia
[15]Victorian Clinical Genetics Services, Melbourne, Australia
[16]Queensland Children's Hospital, Brisbane, Australia; Wesley Hospital, Brisbane, Australia
University of Queensland, Brisbane, Australia
[17]Victorian Clinical Genetics Services, Melbourne, Australia
[18]Children's Cancer Institute, Australia
[19]The University of Melbourne, Melbourne, Australia; Victorian Clinical Genetics Services, Melbourne, Australia; Murdoch Children's Research Institute, Melbourne, Australia
[20]The University of Melbourne, Melbourne, Australia; Victorian Clinical Genetics Services, Melbourne, Australia
[21]Harry Perkins Institute of Medical Research; University of Western Australia, Perth, Australia

[22]Sydney Children's Hospitals Network, Westmead, Australia; University of Sydney, Sydney Australia
[23]Royal Melbourne Hospital, Melbourne, Australia
[24]Australian Genomics; Murdoch Children's Research Institute, Melbourne, Australia
[25]Australian Genomics; Murdoch Children's Research Institute, Melbourne, Australia
[26]Murdoch Children's Research Institute, Melbourne, Australia; The University of Melbourne, Melbourne, Australia
[27]Murdoch Children's Research Institute, Melbourne, Australia; The University of Melbourne, Melbourne, Australia
[28]Garvan Institute, Australia
[29]Royal Adelaide Hospital, Adelaide, Australia
[30]John Hunter Hospital, Newcastle, Australia
[31]The University of Melbourne, Melbourne, Australia Australian Genomics
[32]Australian Genomics; Murdoch Children's Research Institute, Melbourne, Australia
[33] The University of Melbourne, Melbourne, Australia; Bio 21 Molecular Science and Biotechnology Institute, Melbourne, Australia; Murdoch Children's Research Institute, Melbourne, Australia; Victorian Clinical Genetics Services, Melbourne, Australia
[34]Harry Perkins Institute of Medical Research, Perth, Australia
[35]Australian Genomics, Parkville, Victoria, Australia; Department of Genetics and Molecular Pathology, SA Pathology, Adelaide, Australia
[36]Perth Children's Hospital, Perth, Australia
[37]The Australian e-Health Research Centre, CSIRO
[38]Australian Genomics, Parkville, Victoria, Australia;Department of Genetics and Molecular Pathology, SA Pathology, Adelaide, Australia
[39]Murdoch Children's Research Institute, Melbourne, Australia; Yale School of Medicine, New Haven, CT, USA
[40]Perth Children's Hospital, Perth, Australia;Royal Perth Hospital, Perth, Australia
[41]The University of Melbourne, Melbourne, Australia; Victorian Clinical Genetics Services, Melbourne, Australia; Royal Children's Hospital, Melbourne, Australia
[42]Australian Institute of Health Innovation, Macquarie University, Sydney, Australia
[43]Harry Perkins Institute of Medical Research, Perth, Australia
[44]Australian Genomics; Murdoch Children's Research Institute, Melbourne, Australia
[45]Sydney Children's Hospitals Network, Westmead, Australia
[46]Victorian Clinical Genetics Services, Melbourne, Australia; Peter MacCallum Cancer Centre, Melbourne, VIC, Australia
[47]Australian Genomics
[48]Queensland Children's Hospital, Brisbane, Australia
[49]The Australian e-Health Research Centre, CSIRO
[50]Mito Foundation, Sydney, Australia
[51]Royal Melbourne Hospital, Melbourne, Australia
[52]Mater Hospital, South Brisbane, Australia
[53]Australian Genomics; Genetic Health Queensland, Brisbane, Australia
[54]Monash University, Melbourne, Australia
[55]Women's and Children's Hospital, Adelaide, Australia
[56]Centre for Population Genomics, Garvan Institute of Medical Research, Sydney, Australia; Murdoch Children's Research Institute, Melbourne, Australia; The University of Melbourne, Melbourne, Australia
[57]Department of Neurology and Clinical Neurophysiology, Women's and Children's Hospital, Adelaide, South Australia, Australia; Discipline of Paediatrics, University of Adelaide, Adelaide, South Australia, Australia
[58]Murdoch Children's Research Institute, Melbourne, Australia; The University of Melbourne, Melbourne, Australia; Victorian Clinical Genetics Services, Melbourne, Australia;
Bio 21 Molecular Science and Biotechnology Institute, Melbourne, Australia
[59]Faculty of Medicine and Health, University of Sydney, Australia; Department of Genetic Medicine, Westmead Hospital, Westmead, Australia
[60]Genetic Health Queensland, Brisbane, Australia

[61] Tasmanian Clinical Genetics Service, Hobart, Australia; School of Medicine and Menzies Institute for Medical Research, University of Tasmania, Hobart, Australia; Royal Hobart Hospital, Hobart, Australia
[62]Royal Hobart Hospital, Hobart, Australia
[63]Murdoch Children's Research Institute, Melbourne, Australia
[64]Royal Melbourne Hospital, Melbourne, Australia
[65]The University of Melbourne, Melbourne, Australia
[66]Murdoch Children's Research Institute, Melbourne, Australia; The University of Melbourne, Melbourne, Australia; Victorian Clinical Genetics Services, Melbourne, Australia
[67]Murdoch Children's Research Institute, Melbourne, Australia; The University of Melbourne, Melbourne, Australia; Victorian Clinical Genetics Services, Melbourne, Australia

Abstract

Purpose. Families living with mitochondrial diseases (MD) often endure prolonged diagnostic journeys and invasive testing, yet many remain without a molecular diagnosis. Through a national team of clinicians, diagnostic, and research scientists, the Australian Genomics Mitochondrial disease flagship conducted a prospective study to identify the diagnostic utility of singleton genomic sequencing using blood samples as a first step to diagnose MD. **Methods.** 140 children and adults living with suspected MD were recruited using modified Nijmegen criteria (MNC) and randomised to either exome + mtDNA sequencing (ES+mtDNAseq) or genome sequencing (GS).

Results. Diagnostic yield was 55% (n=77) with variants in nuclear (n=37) and mtDNA (n=18) MD genes, as well as phenocopy genes (n=22). A nuclear gene aetiology was identified in 77% of diagnoses, irrespective of disease onset. Diagnostic rate was higher in paediatric-onset (71%) than adult-onset (31%) cases. For children, higher MNC scores correlated with increased diagnostic yield and fewer diagnoses in phenocopy genes. Additionally, three adult patients had a mtDNA deletion discovered in skeletal muscle that was not initially detected in blood.

Conclusion. Genomic sequencing from blood can simplify the diagnostic pathway for individuals living with suspected MD, especially those with childhood onset diseases and high MNC scores.

A) INTRODUCTION

Mitochondrial diseases (MD) are a heterogeneous group of disorders caused by pathogenic variants in nearly 400 genes leading to mitochondrial dysfunction and impaired ability of cellular energy generation.[1] The phenotypic spectrum of MD is very broad and can affect many different organs, including the brain, heart, muscles, and the nervous system. They are the most common group of inherited metabolic disorders with a prevalence of at least 1 in 5,000 live births.[2,3]This group represents the highest mortality in the paediatric population among all inherited metabolic disorders.[4]

Identification of an underlying molecular diagnosis for patients and their families living with suspected MD is crucial for informing clinical management, gaining insight about prognosis, and allowing families to make informed reproductive decisions. These diagnoses can also facilitate further mechanistic research, which may ultimately lead to the development of novel treatments.

The diagnosis of MD has traditionally been based on a combination of clinical criteria, biochemical and genetic testing, which often varies depending on the clinical presentation. However, the complexity and variability of these diseases has made it challenging to accurately diagnose them due to various factors, including the fact that they can be caused by pathogenic variants in either nuclear DNA (nDNA) or mitochondrial DNA (mtDNA), the different modes of inheritance, phenotypic variability, and the levels of mtDNA heteroplasmy. Often families with MD have visited multiple specialists, been misdiagnosed, or required extensive evaluations, including biopsies to perform biochemical, histological, and enzyme evaluations.[5]

Exome (ES) and genome (GS) sequencing have emerged as powerful tools for diagnosing MD. These sequencing technologies allow for the simultaneous testing of multiple genes and have improved the diagnostic yield and the identification of novel disease genes.[6,7]

ES has gained widespread adoption due to its lower cost and ability to target nearly all coding regions and flanking intronic nucleotides. To screen for mtDNA variants, "off-target" reads from ES can be analysed. However, this method depends on the specific exome kit employed as the entire mtDNA is often not equally captured and can be limited in its ability to detect and accurately quantify low levels of heteroplasmy, as the sequencing depth may be variable.[8,9] Alternatively, additional targeted mtDNA sequencing (mtDNAseq) could be performed.[10,11] Genome Sequencing (GS) can comprehensively interrogate both nDNA and mtDNA in a single test detecting variants in the coding and non-coding regions of the genome.[6]

For individuals with suspected MD, using blood samples for genomic testing (GS/ES+/-mtDNAseq) is particularly attractive as it could also potentially obviate the need for invasive testing. The diagnostic yield of GS and ES in the context of MD ranges from 31% to 70%.[12,13] The variability in diagnostic yields may be due to differences in the stringency of inclusion criteria, previous testing, study design, and the inherent heterogeneity among patient cohorts. To better understand the clinical diagnostic utility of these technologies in individuals living with MD, Australian Genomics established the *Mitochondrial disease flagship*, assembling a national team of clinicians, diagnostic, and research scientists who conducted a prospectively designed study by selecting children and adults living with suspected MD using modified Nijmegen criteria (MNC) (Supplementary table S1)[6,14] and randomized for testing through ES+mtDNAseq or GS using DNA extracted from blood as a first step.

B) MATERIALS AND METHODS

Study participants

Prospectively identified individuals with a “probable” (score 5–7) or “definite” (score 8–12) diagnosis of MD based on MNC were eligible for recruitment.[14] Thirteen individuals with a score of 4 (“possible diagnosis”) and without a previous muscle biopsy were accepted because there was consensus by an expert clinical intake committee of warranting investigation. Patients were excluded if they had a previously confirmed molecular diagnosis, previous testing through ES or GS, or an indication that there is another likely non-MD diagnosis from other investigations as determined by the intake review committee.

A total of 140 individuals were recruited between 2017-2020 from the states of New South Wales (n=29), Queensland (n=40), South Australia (n=12), Victoria (n=40), Tasmania (n=2) and Western Australia (n=17). Individuals were randomised to be studied through singleton ES+mtDNAseq or singleton GS using DNA extracted from blood as a first step.

Genetic analysis

The genetic analysis iteratively developed over the course of the study (Figure S1). Initially individuals underwent GS or ES+mtDNAseq from samples in blood as follows:

Exome and mitochondrial DNA sequencing

ES was performed at the Victorian Clinical Genetics Services (VCGS) using the Agilent Sureselect QXT CREv1 and CREv2 kit on Illumina sequencing instruments, with a targeted mean coverage of 100x and a minimum of 90% of bases sequenced to at least 15x. Data were processed using Cpipe[15] to generate annotated variant calls within the target region (coding exons +/- 2bp), via alignment to the reference genome (GRCh37). SNV analysis in the ES cohort was performed using an in-house analysis pipeline. CNV analysis from exome

sequencing data was performed in selected individuals using an internal tool CxGo[16] when a gene of interest was identified.

mtDNAseq was performed if initial ES analysis was negative (n=59/72). For mtDNAseq, the whole mitochondrial DNA (mtDNA) of 16.5kb was amplified with a single long-range PCR, followed by Illumina Nextera® XT library preparation and sequencing on a MiSeq using v2 chemistry at VCGS.[10,11] with a minimum coverage of 1000-fold. Raw sequencing data were analysed with MiSeq Reporter (v2-5-1), which was used to align sequencing reads to the revised Cambridge Reference Sequence (rCRS) mitochondrial genome (NC_012920.1) and to generate both BAM and VCF files, as well as assay quality metrics. A custom in-house analysis pipeline was used to annotate the VCF file with variant information, which was used to perform variant filtration and prioritisation. This assay is clinically validated to detect SNV with heteroplasmy >3%. The BAM file was used to generate coverage and split read plots for detection of large (>1kb) deletions.

Genome sequencing

TruSeq Nano libraries were prepared and loaded onto a HiSeq X Ten sequencer (Illumina; control Software v3.0.29.0) and 2 x 150 bp paired-end sequencing was performed at the Kinghorn Centre for Clinical Genomics (KCCG). Raw sequencing data were converted to FASTQ format using Illumina's bcl2fastq converter (v2.15.0.4) and read quality was evaluated using FASTQC. Sequences were aligned to the b37d5 human reference genome using Burrows-Wheeler Aligner (BWA, v0.7.12-r1039), coordinate-sorted using Novosort (v1.03.04, Novocraft Technologies Sdn Bhd, Selangor, Malaysia), and improved using GATK (v3.4-46-gbc02625) indel realignment and base recalibration to generate BAM files. Variants were called using GATK (v3.4-46-gbc02625) HaplotypeCaller followed by joint variant calling with GenotypeGVCFs and VariantRecalibration.[17] The resulting multi-sample VCF file was annotated

using ENSEMBL's Variant Effect Predictor (v74) and converted to an SQLite database using gemini (v0.17.2).[18] Gemini databases were imported into Seave[19] which was used to perform variant filtration and prioritization for initial GS analyses.

Mitochondrial SNV and indels were identified using mity[20] optimised to identify low heteroplasmy variants (<1%), with an average coverage of 3000-fold of the mitochondrial genome. Structural variants (SV) including copy number variants (CNVs) were investigated using ClinSV.[21]

Updated Genome and Exome analysis

Expanded analyses of the GS and ES data were performed using updated pipelines at the Centre for Population Genomics; in brief the reads were realigned to the UCSC GRCh38/hg38 reference genome using Dragmap (v1.3.0). Cohort-wide joint calling of single nucleotide variants (SNVs) and small insertion/deletion (indel) variants was performed using GATK HaplotypeCaller (v4.1.4.1) with "dragen-mode" enabled.[17] Variants were annotated using VEP (v105) and loaded into the web-based variant filtration platform, Seqr.[22] Sex was inferred from the genotypes using the Somalier tool.[23]

Variant filtration and prioritization were performed using gene lists from PanelApp (Australia)[24], initially using *mitochondrial diseases* (Version 0.850) and *mendeliome* (Version 1.571) gene lists. If a diagnosis was not reached, an expanded analysis was performed using a custom *mitoexome* gene list, which includes genes related to mitochondrial function (Supplementary table S2). Variant curation was based on the American College of Medical Genetics and Genomics (ACMG) guidelines[25], and Variants of Uncertain Significance (VUS) were further sub-classified as being of potential clinical relevance (class 3A), uncertain significance (class 3B) or with low clinical relevance (class 3C).

C) RESULTS

One hundred and forty individuals were recruited into this study, and their characteristics for each sequencing arm are summarized in Table 1.

Table 1. Mitochondrial Flagship individual characteristics

Characteristic	All individuals	ES +mtDNAseq	GS	ES+mtDNAseq vs GS p value
Sex (n, %)				0.608
Male	57 (41%)	31 (43%)	26 (38%)	
Female	83 (59%)	41 (57%)	42 (62%)	
Age of onset (n, %)				0.7315
Adult	55 (39%)	29 (40%)	26 (38%)	
Paediatric	85 (61%)	43 (60%)	42 (62%)	
Modified Nijmegen score (median, IQR)	6 (5 - 7)	6 (5 - 7)	6 (5 - 7)	0.7591

IQR denotes interquartile range. Paediatric individuals were younger than 16 years of age at onset of clinical symptoms.

Human Phenotype Ontology (HPO) terms were extracted from the phenotypic data entries using the CSIRO Fast Healthcare Interoperability Resources terminology server [26] and by manually inspecting intake forms and clinical summaries. In the cohort, each individual exhibited a range of 4 to 23 distinct HPO terms, resulting in a total of 121 unique entries. The HPO terms totalled to 1503 occurrences across all individuals in the cohort. The most frequent terms were HP:0001324 muscle weakness (n=81, 58%), HP:0003546 exercise intolerance (n=75, 54%), HP:0002151 increased serum lactate (n=75, 54%), HP:0001249 intellectual disability (n=54, 39%), HP:0001263 global developmental delay (n=47, 34%), HP:0002376 developmental

regression (n=47, 34%), HP:0001250 seizure (n=43, 31%), HP:0000407 sensorineural hearing impairment (n=40, 29%), HP:0000508 ptosis (n=39, 28%) (Figure 1, supplementary table S3).

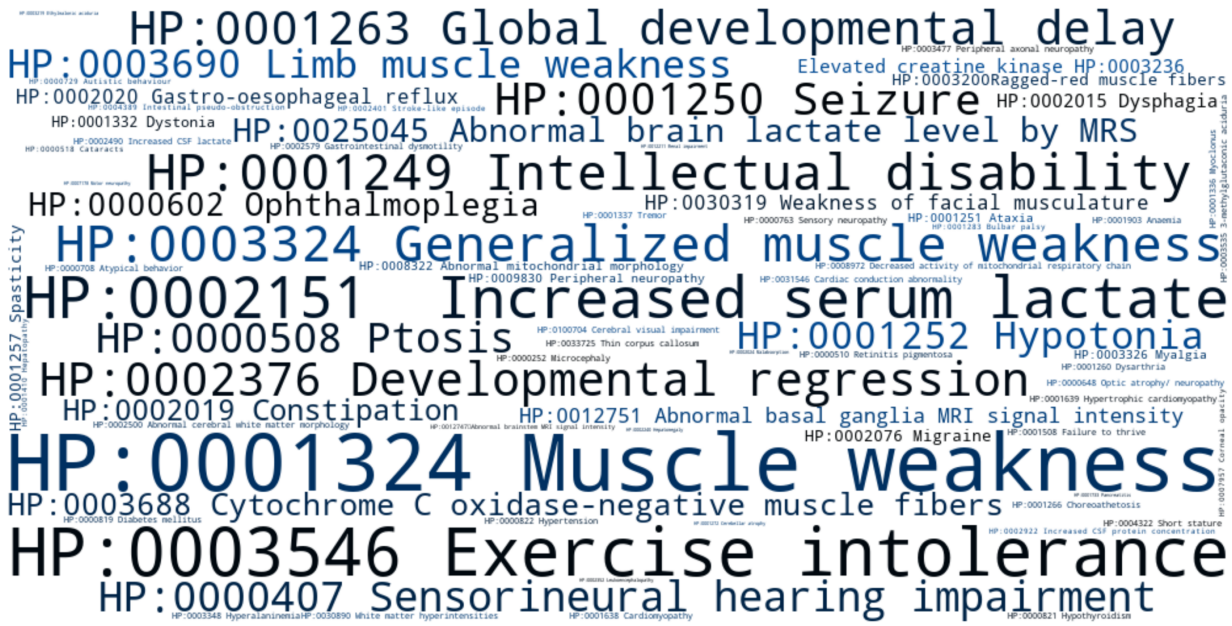

Figure 1. Word Cloud of HPO Term Frequencies (HPO) observed in the cohort.

Size and darkness of each term within the cloud represent its prevalence within the cohort, illustrating the most common phenotypes.

A likely molecular diagnosis was identified in 55% of individuals in the cohort (n=77), which includes 7 individuals with “strong candidate” diagnoses in known disease genes and 2 with novel disease gene associations. Ongoing work, including functional studies, is being conducted to confirm their causality. Seventy one percent (n= 55) of the total cohort diagnoses were in known MD genes, of which 67% (n=37) were nuclear and 33% (n=18) mitochondrial genome in origin (Figure 2). For 29% (n=22) of the diagnoses, the causative genes were not known to have a mitochondrial function (i.e., a phenocopy). Most of the diagnoses were due to single nucleotide variants (SNV) (n=68); other types of variants included 3 duplications involving the

ATAD3 gene cluster[27], 4 single large mtDNA deletions, and intragenic deletions that were identified *in trans* with a SNV in 2 individuals (P3- *SERAC1*; P135 – *AARS2*).

Three individuals had a dual diagnosis, two in non-mitochondrial disease genes and one with a MD and non-mitochondrial disorder. The dual non-mitochondrial disease diagnoses were in *MOGS1* and *CRYAA* in individual P119, and in *MYH9* and *USH2A* in individual P5[28]. Individual P47 had pathogenic variants in the mtDNA encoded MD gene *MT-TL1,* explaining most of his symptoms, and in *SORD* (a non-mitochondrial disease gene) contributing to some of the phenotype. One additional individual, P117, had a partial diagnosis identified in *MYH7*.

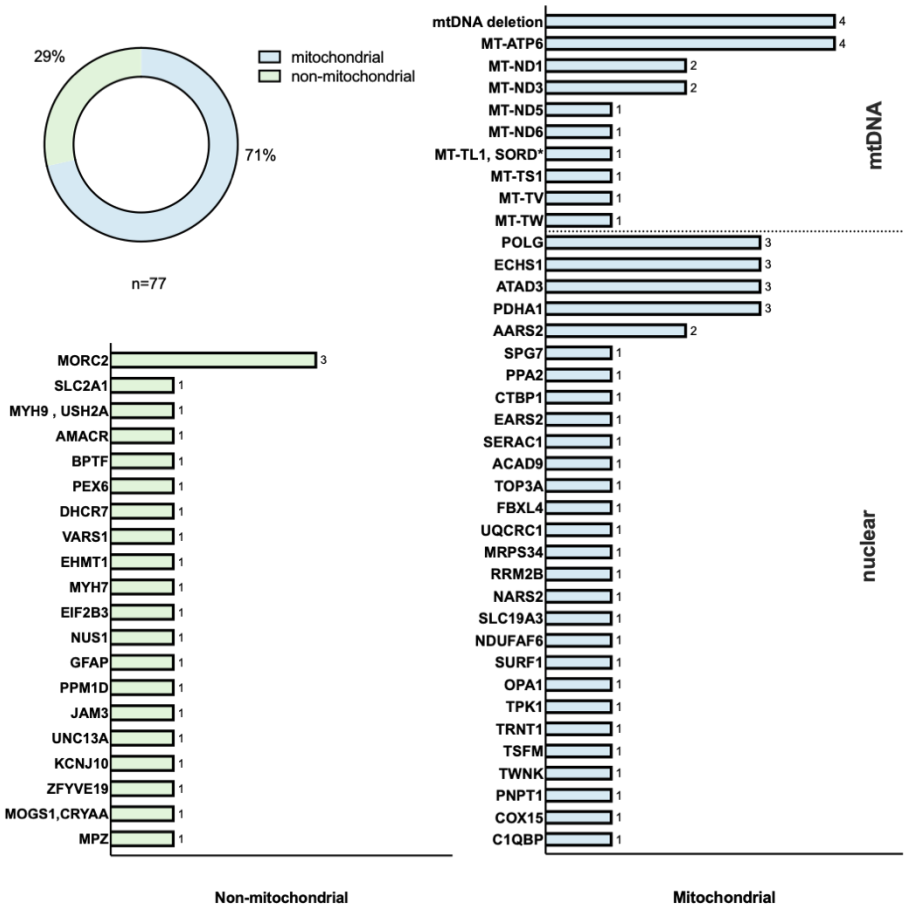

Figure 2. **Number of individuals with a likely molecular diagnosis in mitochondrial and non-mitochondrial disease genes.**

**SORD* is not a mitochondrial disease gene; however, the individual is listed in the mtDNA group as most of the phenotype is explained by the *MT-TL1* pathogenic variant.

During the first stage of analysis, the GS data were interrogated for variants in coding regions of known disease genes (*mitochondrial disease* and *Mendeliome* gene lists). Given this focus on coding regions, it is plausible to assume that if the exome sequencing had robust coverage of these regions, most of the SNVs in the GS arm could have been identified by ES+mtDNAseq. In the expanded GS, non-coding regions of known and candidate disease genes were interrogated, which was not technically possible in the ES+mtDNAseq group.

We subsequently performed secondary GS in 14 individuals from the ES+mtDNAseq cohort where there was a high diagnostic suspicion (such as a single variant of interest in a gene associated with an autosomal recessive disease) and where DNA was available. This resulted in an additional probable diagnosis in P67 as secondary GS identified a deep intronic “second hit” NM_003365.3:c.707-186G>A (spliceAI Donor gain 0.23) in *UQCRC1* currently undergoing functional studies.

In addition, during the expanded analysis, muscle mtDNA testing was suggested if the phenotype was compatible with a mtDNA deletion or a low heteroplasmy variant was identified in blood. Of the 12 individuals who had mtDNAseq in muscle after enrolling in the study, one adult was confirmed to have higher heteroplasmy levels of a pathogenic SNV in *MT-TL1* first identified in blood (m.3243A>T 2% in blood; 69% in muscle), and a single large mtDNA deletion was identified in muscle from 3 individuals.

The diagnostic rate of individuals who started the diagnostic pathway with GS from blood was 56% (n=38). In one individual from this arm, the molecular diagnosis was identified through mtDNAseq in muscle (P56) following non-diagnostic GS. The diagnostic yield of individuals who started their diagnostic trajectory with ES+mtDNAseq was 54% (n=39). However, two individuals were diagnosed with a mtDNA deletion in muscle that was not initially identified in

blood (P8, P124), and in one individual, the presumed molecular diagnosis was achieved after secondary GS (P67). After excluding these 4 individuals, the diagnostic yield was 54% (n=37) for GS and 50% (n=36) for ES+mtDNAseq (p=0.86). The diagnostic pathway and final diagnostic method for the individuals in the cohort are summarised in Figure 3.

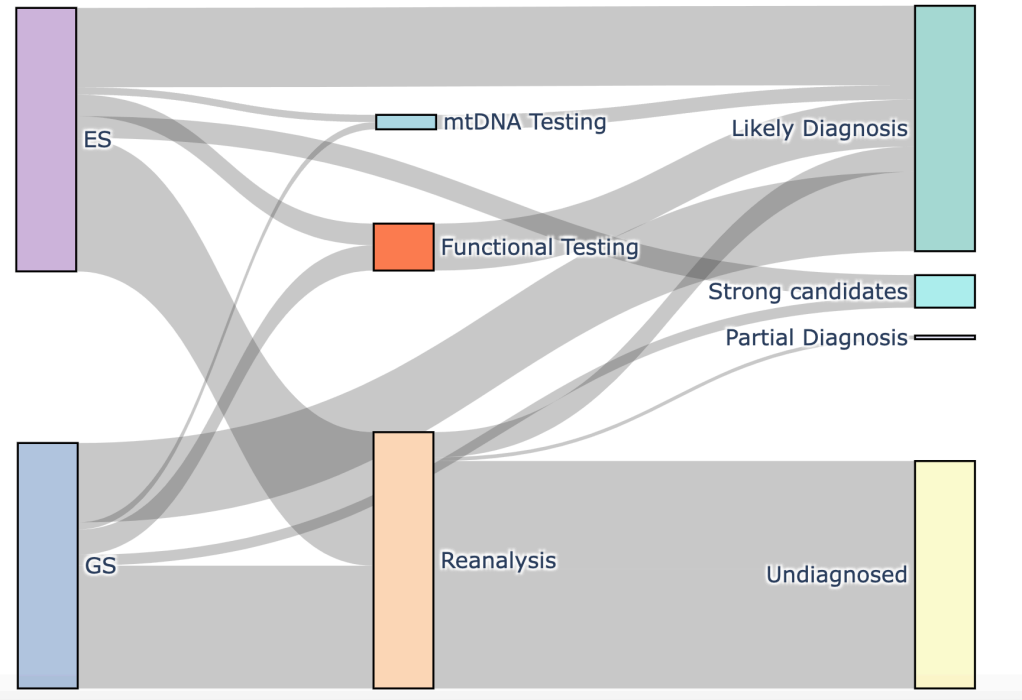

***Figure 3*. Diagnostic status by genomic testing pathway**

Sankey diagram representing the diagnostic trajectory of individuals from the Mitochondrial Flagship cohort. The arc's thickness represents the proportion of individuals transitioning from analysis groups and diagnostic status.

Of the 140 patients, a higher proportion of molecular diagnoses was achieved in the paediatric-onset group 71% (n=60), compared to the adult-onset group 31% (n=17). The paediatric group had a higher MNC score (median 6 IQR 3) than the adult group (median 5 IQR 2) (p= 0.0005), and a higher MNC score was associated with a greater rate of molecular diagnosis in the paediatric but not in the adult participants. In addition, the MNC scores were higher in

individuals with a likely diagnosis in a MD gene than in a non-mitochondrial disease gene or in the undiagnosed group (Figure 4).

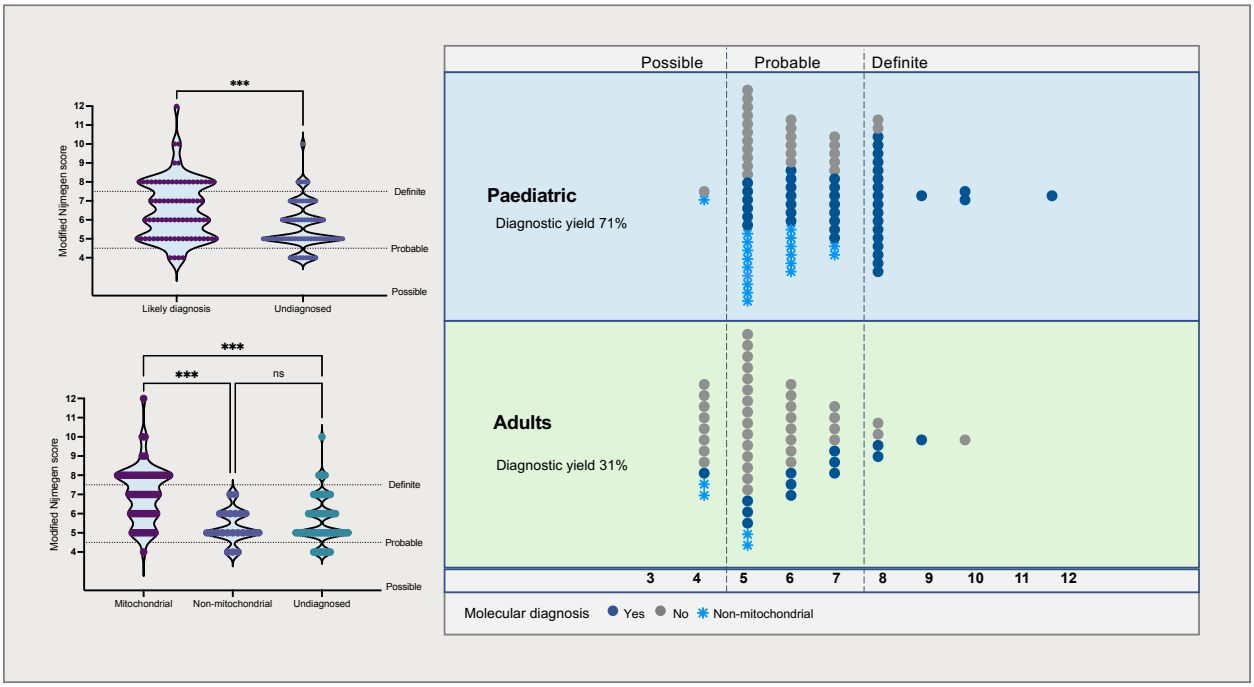

Figure 4 **Modified Nijmegen score and diagnostic outcomes**

The modified Nijmegen scores were higher in individuals with a genetic diagnosis. A higher score was associated with a mitochondrial diagnosis than a non-mitochondrial diagnosis or those who remained undiagnosed. Modified Nijmegen scores were higher in individuals with a likely molecular diagnosis in the paediatric group (median 7 IQR 3 vs 6 IQR 2 p=0.01), but not in the adult individuals (median 6 IQR 2 vs 5 IQR 1 p=0.30). *** <0.001; ns non-significant; IQR Interquartile Range.

D) DISCUSSION

Our results show the diagnostic utility of starting the diagnostic pathway with genomic sequencing (GS or ES+mtDNA) from blood for the diagnosis of MD. Interestingly, the diagnostic yield was higher in individuals with paediatric than adult onset (71% vs 31%, p<0.0001). Several factors likely contributed to this outcome.

Firstly, in adult blood, heteroplasmy levels for some mtDNA variants can decline with age along with mtDNA deletions becoming undetectable due to the positive selection of hematopoietic

stem cells that harbor zero or a low amount of deleted mtDNA.[29-31] The use of blood as a source of DNA testing could also contribute to only 29% (n=5) of adult-onset individuals having a molecular diagnosis due to primary mtDNA variants, which is lower than estimates of up to 75% of adult-onset MD being caused by mtDNA variants from previous retrospective studies.[7] Skeletal muscle tissue was available from 12 individuals in our cohort who lacked a confirmed diagnosis after genomic testing in blood, and three of these had single mtDNA deletions detected by muscle mtDNAseq. A fourth (P47), had the m.3243A>T SNV detected in blood at 2% heteroplasmy, which we regarded as too low to be diagnostic but its presence at 69% heteroplasmy in muscle confirmed the genotype/phenotype relationship. Testing muscle in further individuals from the cohort could help clarify the proportion of patients in whom a diagnosis was missed due to blood-derived DNA being the initial source for testing. A recent cohort of individuals with adult-onset MD achieved a diagnostic yield of 54% (130/242).[32] In 62% (n=80/130) of those diagnoses, the cause was mtDNA in origin. Seven of those individuals had mtDNA deletions detected in muscle that were not detected in blood when tested using GS. All mtDNA SNVs were detected in blood, albeit 9 at heteroplasmy levels of <=3% while all mtDNA SNVs and some deletions were detected in blood.[32]

A second contributor to the lower diagnostic yield in adults could be the selection of individuals, whereby adults with well-defined mitochondrial phenotypes may have already undergone targeted molecular testing rather than being recruited into this study. Targeted testing for the most common pathogenic mtDNA SNVs (e.g., m.3243A>G, m.1555A>G, m.11778G>A, m.14484T>C, m.3460G>A) has been available for decades but no individuals with these SNVs were detected in the Mitochondrial Flagship cohort.[2] Therefore, it is likely that individuals with pathogenic mtDNA SNVs causing common mtDNA disorders such as Mitochondrial myopathy, Encephalopathy, Lactic Acidosis, and Stroke-like episodes (MELAS, MIM # 540000) or Leber

Hereditary Optic Neuropathy (LHON, MIM # 535000) had prior testing and were not recruited to this cohort.

A third contributor to the lower diagnostic yield in adults is that the MNC score used in our inclusion criteria appears less useful for adult-onset than paediatric-onset individuals. The original Nijmegen criteria were developed as a diagnostic tool to evaluate the likelihood of a child having MD and neither the original criteria or MNC have been validated for adults.[14] In a recent cohort of adult-onset MD, a higher Nijmegen criteria score was not found to be associated with a diagnosis.[32] In the Mitochondrial Flagship cohort, the MNC score was a useful tool to prioritize which paediatric-onset individuals would be more likely to receive a molecular diagnosis. However, different diagnostic and prioritization criteria may be required for adults and further research involving larger cohorts is necessary to develop appropriate screening tools for this population.

The MNC score appears to be a useful tool for identifying individuals who have a higher likelihood of a molecular diagnosis in genes related to mitochondrial function. For the 77 individuals with a likely molecular diagnosis, the cause was in a known MD gene rather than a phenocopy gene in 100% (24/24) with a MNC score >7 (“definite”), compared to 61% (30/49) in the probable and 25% (1/4) in the possible groups.

Expanding beyond mitochondrial disease genes in individuals with MNC score <8 resulted in the identification of 29% (n=22) of the molecular diagnoses, which is comparable with findings from another highly selected cohort of 40 paediatric individuals with suspected mitochondrial disease, where non-mitochondrial disease genes accounted for 18% (7/40) of diagnoses.[6] In cohorts with less stringent inclusion criteria, non-mitochondrial disorders were even more common than mitochondrial disorders (63% of diagnoses).[12]

A diagnosis in a non-mitochondrial disease gene was identified even in individuals where imaging or biochemical evidence was suggestive of a MD. Three children were diagnosed with *MORC2*-neurodevelopmental disease (MIM 619090) and one adult with Alpha-methylacyl-CoA racemase (*AMACR*) deficiency (MIM 614307). *De novo* variants in *MORC2* have recently emerged as a mitochondrial phenocopy gene, with some individuals having Leigh syndrome-like lesions on brain MRI.[33] Similarly, *AMACR* variants have also been recognised in multiple adults with suspected MD.[34] In addition, an individual with persistent 3-methylglutaconic aciduria (3MGA), a biomarker often associated with phospholipid remodelling or mitochondrial membrane-associated disorders,[35] was diagnosed with Kleefstra syndrome (MIM 607001), which is associated with a gene (*EHMT1)*, not known to cause mitochondrial disease.[36] Testing for 3MGA in additional patients with Kleefstra syndrome could clarify if there is an underlying secondary mechanism associated with the persistent 3MGA. Overall, these examples highlight the utility of non-targeted sequencing approaches and expanding analyses to include genes without a known mitochondrial function.

A molecular diagnosis is yet to be identified in 45% of individuals who were part of this study. A combination of factors likely contributes to this; technological limitations make it difficult to identify certain types of genetic variations particularly in ES data (such as SV, short tandem repeats, and variants in non-coding regions). However, many challenges are related to limitations of variant interpretation. For instance, GS can technically identify variants in non-coding regions, but our current variant interpretation tools are still limited. Bioinformatic approaches are improving rapidly; for instance, when the Mitochondrial Flagship program first began, SpliceAI, which is used to analyse and interpret genetic variation that can affect the splicing process, was not available. Now, it is considered a standard tool for variant filtration and prioritization.[37] The re-analysis of existing genomic data, as more variant interpretation tools

become available and novel disease genes are discovered, is expected to be a valuable tool for increasing diagnostic yield.[38,39]

Combining GS with other methodologies, such as transcriptome, proteome, metabolome, lipidome, and glycome analyses, may help to overcome some of the limitations of using ES/GS alone. For example, RNA sequencing (RNAseq) can detect abnormal gene expression, mono-allelic expression, or splicing defects, while quantitative proteomics can detect changes in protein abundance for different variant types, including missense, intronic and copy number variants as well as downstream effects of these variants on pathways and complexes.[27,40-43] Similarly, studies of metabolites, lipids, and glycans can detect characteristic metabolite profiles and biomarkers.[44]

Therefore, individuals with high MNC scores who are still molecularly undiagnosed are currently being recruited to other research projects to incorporate systematic reanalysis and other -omic technologies with the aim to provide more patients and families with a molecular diagnosis. Building on the effort to extend molecular diagnoses to more individuals, the Mitochondrial Flagship has contributed to shaping standard care in Australia. By providing data to Australia's Medical Services Advisory Committee (MSAC) recommendation application 1675,[45] it has played a role in the establishment of new Medicare Benefits Schedule (MBS) item numbers (73456, 73457, 73458, 73459, 73460, 73461 and 73462)[46]. This initiative is a major step forward in promoting diagnostic accuracy and equal access to genomic testing through public funding.

Data Availability

The datasets supporting the current study have not been deposited in a public repository due to consent restrictions. De-identified genomic and associated data from this study are available for ethically approved research. The online access application process is administered by the Australian Genomics Data Access Committee.

Acknowledgments

We would like to thank all the families who participated in this study, and the clinical teams involved in their care. The research conducted at the Murdoch Children's Research Institute was supported by the Victorian Government's Operational Infrastructure Support Program. The Chair in Genomic Medicine awarded to JC is generously supported by The Royal Children's Hospital Foundation. We acknowledge the Bio21 Mass Spectrometry and Proteomics Facility (MMSPF) for the provision of instrumentation, training, and technical support.

Funding Statement

The Mitochondrial Flagship project was funded by Australian Genomics Health Alliance (Australian Genomics) NHMRC Targeted Call for Research grant GNT1113531 and supported by NHMRC grants 1164479, 1155244, 1159456, and 2009732, and the US Department of Defense Congressionally Directed Medical Research Programs PR170396. We acknowledge the Australian Mito Foundation for funding support. We are grateful to the Crane, Perkins, and Miller families for their generous financial support.

Author Contributions

Conceptualization- J.C., D.R.T

Data curation- R.R., A.G.C., N.L.B., L.E.F., N.J.L., J.E.M.

Formal analysis- R.R., A.G.C., L.E.F., A.E.F., A.E.W., D.H.H., D.A.S., I.G, Y.W.

Funding acquisition- J.C., D.R.T., T.F.B.

Investigation-S.Be, J.Br., J.C.L., K.Bo., J.Bu., A.B., S.E., M.G.dS, K.F., D.H.H., D.H., M.R.J., M.K., S.K-S., M.M., E.M.M., D.A.S., J.L.,M.C.J.Q., S.S., T.F.B., S.C., M.H.

Methodology- M.J.C, C.S., M.T.R.

Project administration- N.L.B., M.G.dS, T.F.B., T.M., S.M.,

Clinical intake committee- S.Ba., D.B., K.Bh., D.C., J.C., M.B.D., C.E., J.F., M.C.F., R.G., H.G., M.P.K., P.J.L., J.L., J.McG., J.P., L.K.P., N.S., D.R.T., M.C.T., M.W., T.L.W., C.W.

Resources- J.C., D.R.T, C.E

Software- M.J.B., A.M-J., M.J.C, C.S., V.G.

Supervision- J.C., D.R.T, A.G.C.

Validation- B.C., S.C.

Visualization- R.R.

Writing-original draft- R.R.

Writing-review & editing- J.C., A.G.C., D.R.T. and all authors

Ethics Declaration

This study was conducted in accordance with the revised Declaration of Helsinki and following the Australian National Health and Medical Research Council statement of ethical conduct in research involving humans. The Mitochondrial Flagship study was reviewed and approved through our lead Human Research Ethics Committee (HREC), Royal Melbourne Hospital (formerly known as Melbourne Health) (HREC/16/MH/251). Sites that were not covered at the time by the Australian National Mutual Acceptance system were reviewed and approved by the Western Australian Child and Adolescent Health Service HREC (RGS0000000086), Tasmanian HREC (H0016443) and Queensland UnitingCare Health HREC (1717).

Conflict of Interest

JC is an approved pathology provider for the Victorian Clinical Genetics Service.

References

1. Rahman S. Mitochondrial disease in children. *J Intern Med*. Jun 2020;287(6):609-633. doi:10.1111/joim.13054
2. Gorman GS, Chinnery PF, DiMauro S, et al. Mitochondrial diseases. *Nat Rev Dis Primers*. Oct 20 2016;2:16080. doi:10.1038/nrdp.2016.80
3. Skladal D, Halliday J, Thorburn DR. Minimum birth prevalence of mitochondrial respiratory chain disorders in children. *Brain*. Aug 2003;126(Pt 8):1905-12. doi:10.1093/brain/awg170
4. Goel H, Lusher A, Boneh A. Pediatric Mortality Due to Inborn Errors of Metabolism in Victoria, Australia: A Population-Based Study. *JAMA*. 2010;304(10):1070-1072. doi:10.1001/jama.2010.1259
5. Grier J, Hirano M, Karaa A, Shepard E, Thompson JLP. Diagnostic odyssey of patients with mitochondrial disease: Results of a survey. *Neurol Genet*. Apr 2018;4(2):e230. doi:10.1212/NXG.0000000000000230
6. Riley LG, Cowley MJ, Gayevskiy V, et al. The diagnostic utility of genome sequencing in a pediatric cohort with suspected mitochondrial disease. *Genet Med*. Jul 2020;22(7):1254-1261. doi:10.1038/s41436-020-0793-6
7. Stenton SL, Prokisch H. Genetics of mitochondrial diseases: Identifying mutations to help diagnosis. *EBioMedicine*. Jun 2020;56:102784. doi:10.1016/j.ebiom.2020.102784
8. Wagner M, Berutti R, Lorenz-Depiereux B, et al. Mitochondrial DNA mutation analysis from exome sequencing-A more holistic approach in diagnostics of suspected mitochondrial disease. *J Inherit Metab Dis*. Sep 2019;42(5):909-917. doi:10.1002/jimd.12109
9. Griffin HR, Pyle A, Blakely EL, et al. Accurate mitochondrial DNA sequencing using off-target reads provides a single test to identify pathogenic point mutations. *Genet Med*. Dec 2014;16(12):962-71. doi:10.1038/gim.2014.66
10. Akesson LS, Eggers S, Love CJ, et al. Early diagnosis of Pearson syndrome in neonatal intensive care following rapid mitochondrial genome sequencing in tandem with exome sequencing. *Eur J Hum Genet*. Dec 2019;27(12):1821-1826. doi:10.1038/s41431-019-0477-3
11. Rius R, Compton AG, Baker NL, et al. Application of Genome Sequencing from Blood to Diagnose Mitochondrial Diseases. *Genes (Basel)*. Apr 20 2021;12(4)doi:10.3390/genes12040607
12. Schon KR, Horvath R, Wei W, et al. Use of whole genome sequencing to determine genetic basis of suspected mitochondrial disorders: cohort study. *BMJ*. Nov 3 2021;375:e066288. doi:10.1136/bmj-2021-066288
13. Haack TB, Haberberger B, Frisch EM, et al. Molecular diagnosis in mitochondrial complex I deficiency using exome sequencing. *J Med Genet*. Apr 2012;49(4):277-83. doi:10.1136/jmedgenet-2012-100846
14. Morava E, van den Heuvel L, Hol F, et al. Mitochondrial disease criteria: diagnostic applications in children. *Neurology*. Nov 28 2006;67(10):1823-6. doi:10.1212/01.wnl.0000244435.27645.54
15. Sadedin SP, Dashnow H, James PA, et al. Cpipe: a shared variant detection pipeline designed for diagnostic settings. *Genome Med*. 2015;7(1):68. doi:10.1186/s13073-015-0191-x
16. Sadedin SP, Ellis JA, Masters SL, Oshlack A. Ximmer: a system for improving accuracy and consistency of CNV calling from exome data. *Gigascience*. Oct 1 2018;7(10)doi:10.1093/gigascience/giy112
17. Van der Auwera GA, O'Connor BD. *Genomics in the Cloud: Using Docker, GATK, and WDL in Terra*. 1st ed. O'Reilly Media; 2020.
18. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. *Genome Biol*. Jun 6 2016;17(1):122. doi:10.1186/s13059-016-0974-4

19. Gayevskiy V, Roscioli T, Dinger ME, Cowley MJ. Seave: a comprehensive web platform for storing and interrogating human genomic variation. *Bioinformatics*. Jan 1 2019;35(1):122-125. doi:10.1093/bioinformatics/bty540
20. Puttick C, Kumar KR, Davis RL, et al. mity: A highly sensitive mitochondrial variant analysis pipeline for whole genome sequencing data. *bioRxiv*. 2019;doi:10.1101/852210
21. Minoche AE, Lundie B, Peters GB, et al. ClinSV: clinical grade structural and copy number variant detection from whole genome sequencing data. *Genome Med*. Feb 25 2021;13(1):32. doi:10.1186/s13073-021-00841-x
22. Pais LS, Snow H, Weisburd B, et al. seqr: A web-based analysis and collaboration tool for rare disease genomics. *Hum Mutat*. Jun 2022;43(6):698-707. doi:10.1002/humu.24366
23. Pedersen BS, Bhetariya PJ, Brown J, et al. Somalier: rapid relatedness estimation for cancer and germline studies using efficient genome sketches. *Genome Med*. Jul 14 2020;12(1):62. doi:10.1186/s13073-020-00761-2
24. Martin AR, Williams E, Foulger RE, et al. PanelApp crowdsources expert knowledge to establish consensus diagnostic gene panels. *Nat Genet*. Nov 2019;51(11):1560-1565. doi:10.1038/s41588-019-0528-2
25. Richards S, Aziz N, Bale S, et al. 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. *Genet Med*. May 2015;17(5):405-24. doi:10.1038/gim.2015.30
26. Metke-Jimenez A, Lawley M, Hansen D. FHIR OWL: Transforming OWL ontologies into FHIR terminology resources. *AMIA Annu Symp Proc*. 2019;2019:664-672.
27. Frazier AE, Compton AG, Kishita Y, et al. Fatal perinatal mitochondrial cardiac failure caused by recurrent de novo duplications in the ATAD3 locus. *Med (N Y)*. Jan 15 2021;2(1):49-73. doi:10.1016/j.medj.2020.06.004
28. Ratnayake C, Rius R, Wallis M, Raj R, Christodoulou J. Any symptom, in any organ, at any age: A case report of multiple genetic diagnoses mimicking mitochondrial disease in an adult with kidney disease. *Nephrology (Carlton)*. Jul 2022;27(7):640-641. doi:10.1111/nep.14023
29. Grady JP, Pickett SJ, Ng YS, et al. mtDNA heteroplasmy level and copy number indicate disease burden in m.3243A>G mitochondrial disease. *EMBO Mol Med*. Jun 2018;10(6)doi:10.15252/emmm.201708262
30. Rahman S, Poulton J, Marchington D, Suomalainen A. Decrease of 3243 A-->G mtDNA mutation from blood in MELAS syndrome: a longitudinal study. *Am J Hum Genet*. Jan 2001;68(1):238-40. doi:10.1086/316930
31. Yanagihara I, Inui K, Yanagihara K, et al. Fluorescence in situ hybridization analysis of peripheral blood cells in Pearson marrow-pancreas syndrome. *J Pediatr*. Sep 2001;139(3):452-5. doi:10.1067/mpd.2001.116296
32. Davis RL, Kumar KR, Puttick C, et al. Use of Whole-Genome Sequencing for Mitochondrial Disease Diagnosis. *Neurology*. Aug 16 2022;99(7):e730-e742. doi:10.1212/WNL.0000000000200745
33. Guillen Sacoto MJ, Tchasovnikarova IA, Torti E, et al. De Novo Variants in the ATPase Module of MORC2 Cause a Neurodevelopmental Disorder with Growth Retardation and Variable Craniofacial Dysmorphism. *Am J Hum Genet*. Aug 6 2020;107(2):352-363. doi:10.1016/j.ajhg.2020.06.013
34. Schon KR, Chinnery PF. Whole-genome sequencing for mitochondrial disorders identifies unexpected mimics. *Pract Neurol*. Oct 17 2022;doi:10.1136/pn-2022-003570
35. Wortmann SB, Duran M, Anikster Y, et al. Inborn errors of metabolism with 3-methylglutaconic aciduria as discriminative feature: proper classification and nomenclature. *J Inherit Metab Dis*. Nov 2013;36(6):923-8. doi:10.1007/s10545-012-9580-0

36. Willemsen MH, Vulto-van Silfhout AT, Nillesen WM, et al. Update on Kleefstra Syndrome. *Mol Syndromol*. Apr 2012;2(3-5):202-212. doi:10.1159/000335648
37. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, et al. Predicting Splicing from Primary Sequence with Deep Learning. *Cell*. Jan 24 2019;176(3):535-548 e24. doi:10.1016/j.cell.2018.12.015
38. Robertson AJ, Tan NB, Spurdle AB, Metke-Jimenez A, Sullivan C, Waddell N. Re-analysis of genomic data: An overview of the mechanisms and complexities of clinical adoption. *Genet Med*. Apr 2022;24(4):798-810. doi:10.1016/j.gim.2021.12.011
39. Ewans LJ, Schofield D, Shrestha R, et al. Whole-exome sequencing reanalysis at 12 months boosts diagnosis and is cost-effective when applied early in Mendelian disorders. *Genet Med*. Dec 2018;20(12):1564-1574. doi:10.1038/gim.2018.39
40. Lake NJ, Webb BD, Stroud DA, et al. Biallelic Mutations in MRPS34 Lead to Instability of the Small Mitoribosomal Subunit and Leigh Syndrome. *Am J Hum Genet*. Aug 03 2017;101(2):239-254. doi:10.1016/j.ajhg.2017.07.005
41. Kremer LS, Bader DM, Mertes C, et al. Genetic diagnosis of Mendelian disorders via RNA sequencing. *Nat Commun*. Jun 12 2017;8:15824. doi:10.1038/ncomms15824
42. Helman G, Compton AG, Hock DH, et al. Multiomic analysis elucidates Complex I deficiency caused by a deep intronic variant in NDUFB10. *Hum Mutat*. Nov 10 2020;doi:10.1002/humu.24135
43. Amarasekera SSC, Hock DH, Lake NJ, et al. Multi-omics identifies large mitoribosomal subunit instability caused by pathogenic MRPL39 variants as a cause of pediatric onset mitochondrial disease. *Human Molecular Genetics*. 2023;32(15):2441-2454. doi:10.1093/hmg/ddad069
44. Wortmann SB, Oud MM, Alders M, et al. How to proceed after "negative" exome: A review on genetic diagnostics, limitations, challenges, and emerging new multiomics techniques. *J Inherit Metab Dis*. Jul 2022;45(4):663-681. doi:10.1002/jimd.12507
45. Medical Services Advisory Committee. 1675 – Whole Genome Sequencing for the diagnosis of mitochondrial disease. Australian Government Department of Health and Aged Care. Accessed November 03, 2023, http://www.msac.gov.au/internet/msac/publishing.nsf/Content/1675-public
46. Medicare Benefits Schedule Online. Medicare updates November 2023. Australian Government Department of Health and Aged Care. Accessed November 11, 2023 , http://www.mbsonline.gov.au/internet/mbsonline/publishing.nsf/Content/news-2023-11-01