Whole Exome Sequencing to Uncover the Genetic Etiology of a Cohort of Patients with Suspected Mitochondrial Leukoencephalopathies

Travis A.J Moore

Integrated Program in Neuroscience

McGill University, Montreal

August 2019

A thesis submitted to McGill University in partial fulfillment of the degree of Master of Science

Table of Contents

Abstract

The number of identified Mendelian disorders has far surpassed what would be possible without advancements in DNA sequencing technologies, such as whole exome sequencing (WES). In medicine, this advancement has become an invaluable tool for rapidly identifying the underlying molecular causes of rare diseases and is essential in diagnosis, treatment, and further genetic counselling. In the case of leukodystrophies, a group of rare heritable white matter diseases that primarily present in children with distinct Magnetic Resonance Imaging (MRI) patterns, there are still many unsolved genetic cases.

The first section of this project focuses on identifying the etiology of a cohort of patients with unsolved mitochondrial leukoencephalopathies, a non-hypomyelinating leukodystrophy caused by disruption of mitochondrial function. WES analysis was performed on 11 patients using custom gene panels and variant categorization calling using guidelines from the American College of Medical Genetics (ACMG). From this analysis, a large mitochondrial DNA deletion was identified in one patient, while biallelic variants in *MBOAT7, WDR62, PDHA1*, and *AARS* were identified in four other patients. We identified a strong gene candidate in 5/11 patients, i.e. 45% of the patients, a rate that is similar to what has been published in other cohorts.

The second portion of this project focuses on investigating the functional consequences of two novel variants identified in the aminoacyl-tRNA synthetase for alanine (*AARS*) in one of our patients. The patient presented a novel phenotype of AARS-related disorder and the patient-specific variants, (NM_001605.2) c.295G>A; p.E99K & c.778A>G; p.T260A, were proven to be pathogenic. The patient's *AARS* variants were confirmed to be in *trans*, with the c.295G>A being inherited from the mother and

c.778A>G being identified as a *de novo* variant by cDNA amplification and cloning techniques. The functional consequences of these variants on messenger RNA was assessed next, in addition to protein and cellular localization of AARS. A significant increase was seen in mRNA levels, however, the AARS proteins levels were significantly reduced. Finally, AARS typically has a homogenous distribution in the cytoplasm, yet patient's fibroblasts showed an altered distribution with abnormal puncta staining. Assessment of AARS aminoacylation activity in patient-derived fibroblasts, completed by collaborators, revealed a reduction of AARS activity at 16% compared to that of a control. Taken together, these functional tests demonstrate the variants have caused functional defects in AARS protein and that this patient presents with a novel AARS phenotype.

By identifying strong candidates in patients with unsolved suspected mitochondrial leukoencephalopathies, WES continues to be demonstrated as an invaluable tool in the identification of the molecular causes for rare diseases. Investigations into patient-specific AARS variants provide evidence for the phenotypic expansion of AARS-related diseases and the specific functional defects that underlie these late-onset cases. As the pathophysiological understanding of aminoacyl t-RNA synthetases is still limited, our work as well as others' will be invaluable to our foundational understanding of this group of disorders, opening the door to the development of potential therapies.

Résumé

Le nombre de maladies mendéliennes identifiées à ce jour a largement dépassé ce qu'il aurait été possible sans le progrès des technologies de séquençage de l'ADN tel que le séquençage d'exome entier (WES). En médecine, cette avancée est devenue un outil indispensable à l'identification rapide des causes moléculaires des maladies rares, en plus d'être essentielle pour le diagnostic, le traitement et les conseils génétiques associés à ces maladies. Dans le cas des leucodystrophies, un groupe de maladies héréditaires rares affectant la substance blanche, touchant principalement les enfants et présentant un profil spécifique d'imagerie par résonance magnétique (IRM), il existe encore de nombreux cas génétiques non résolus.

La première partie de ce projet porte sur l'identification de l'étiologie d'une cohorte de patients atteints de leucoencéphalopathies mitochondriales non résolues. Ces dernières sont des leucodystrophies non hypomyélinisantes causées par une perturbation de la fonction mitochondriale. L'analyse WES a été effectuée sur 11 patients à l'aide de panels de gènes prédéfinis et suivant les lignes directrices de l'American College of Medical Genetics (ACMG). De ces analyses, une délétion importante de l'ADN mitochondrial a été identifiée chez un patient et des variants bialléliques dans les gènes MBOAT7, WDR62, PDHA1 et AARS ont été identifiés chez quatre autres patients. Nous avons identifié un gène candidat chez 5 des 11 patients, ce qui équivaut à 45% de la cohorte. Ce taux est similaire à celui publié dans les autres cohortes.

La deuxième partie de ce projet porte sur l'étude fonctionnelle de deux nouveaux variants identifiés dans l'aminoacyl-ARNt synthétase de l'alanine (AARS) chez l'un de nos patients. Le patient présente un nouveau phénotype associé à un trouble lié au gène AARS et les variants identifiés (NM_001605.2) c.295G> A; p.E99K & c.778A> G; p.T260A

se sont révélés pathogéniques. Il a été confirmé que les variants de AARS du patient étaient en trans, que la mutation c.295G>A a été héritée de la mère et par les techniques d'amplification et de clonage d'ADN complémentaire que la mutation c788A> G est de novo. Les conséquences fonctionnelles de ces variants sur l'ARN messager ainsi que sur la localisation protéique et cellulaire de AARS ont ensuite été évaluées. Une augmentation significative des niveaux d'ARNm a été constatée. Cependant, les taux de protéines AARS ont été réduits de manière significative. De plus, la protéine AARS est distribuée de façon homogène dans le cytoplasme, mais dans les fibroblastes du patient, la distribution de la protéine AARS est perturbée et est sous forme anormale en *puncta*. L'évaluation de l'activité d'aminoacylation de AARS dans les fibroblastes dérivés du patient, réalisée par des collaborateurs, a révélé une réduction de l'activité de l'AARS de 16% par rapport à celle d'un témoin. Pris ensemble, ces tests fonctionnels montrent que les variants ont provoqué des défauts fonctionnels dans la protéine AARS et que ces défauts causent le nouveau phénotype associé à AARS chez notre patient.

En identifiant des gènes candidats chez les patients atteints de leucoencéphalopathies mitochondriales non résolues, le WES s'avère être un outil précieux pour l'identification des causes moléculaires des maladies rares. L'étude des mutations de AARS, spécifiques à certains patients, permet d'apprécier l'expansion phénotypique des maladies reliées à AARS ainsi que les défauts fonctionnels spécifiques qui sont associés avec les cas d'apparition tardive. Puisque la compréhension pathophysiologique des aminoacyl ARNt synthétases est limitée à ce jour, notre travail ainsi que celui d'autres chercheurs est indispensable à une meilleure compréhension de ce groupe de maladies, ouvrant ainsi la porte au développement de thérapies potentielles.

Acknowledgements

I would first like to thank my thesis supervisor Dr. Genevieve Bernard for her mentorship and guidance throughout my master's training. I am exceedingly thankful that she provided an environment that allowed me to broaden my established techniques and develop a more comprehensive understanding of science. Her dedication to patients and her students, even with a busy schedule, has been inspiring.

Of course, I am incredibly grateful for the wonderful individuals in the Bernard lab. First is Kether Guerrero, who from the beginning was instrumental in teaching me various techniques in the lab. From PCR to cell culturing, he has been a great help to my learning, as well as allowing me to acknowledge that I had a sigh problem, thankfully cured. Thank you to Luan Tran for your exceedingly calm demeanor, I knew that you would be unfazed by any question I could throw at you. I am grateful to have had the assistance from Lama Darbelli, for her passion and drive for science, keeping me going even during the dark times when mycoplasma threatened to destroy everything. I would be remiss if I wasn't thankful to the other students in the lab. Mackenzie Michell-Robinson thank you for being a soundboard of scientific ideas and conversations, also as well as being the other half to the Gilmore girls in the lab. Alexa Derksen, for being one of the sassiest personalities in the lab and being able to enjoy dog photos at the same level as myself. Aaron Spahr thank you for putting up with me calling you A-Aron and being a source of constant positivity. Stefanie Perrier, thank you for accepting my weirdness since I started, you have been a great lab member and friend since I started.

Thank you to past lab members who have been a part of this journey. Especially to Lynda-Marie who after having a child with leukodystrophy decided to help raise awareness and money for research. Her positivity and mentality while in the lab engraved

in my mind why the research was so important. Additionally, I would like to thank Marie-Lou St-Jean for her help with translating my Abstract.

I would also like to thank the many members in Braverman lab, as they have always been willing to help me if I needed outside advice or temporary resource allocation.

I have had the pleasure of being part of the Integrated program in neuroscience that has allowed me to interact with many excellent supervisors. I have been able to learn from Thomas Durcan, Timothy Kennedy, and Eric Shoubridge as a part of my advisory committee, and Adam as my student mentor.

Finally, I would like to thank friends and family back home, and even though they are at the other side of the country, they have been a constant source of support over many, many phone calls. To the friends I have met here in Montreal over the past two years, thank you for helping me adjust to being so far away from family. Also, for your understanding that I became a hermit during my thesis was commendable. Additionally, thank you to my roommate, Brandon Shokoples, allowing me to ramble about maybe more than I had the right to.

Contribution of Authors

Chapter 1: Introduction

TM wrote this section in collaboration with Dr. Genevieve Bernard.

Chapter 2: Hypothesis and Rational

TM wrote this section in collaboration with Dr. Genevieve Bernard.

Chapter 3: Methods

TM designed experiments with guidance from Kether Guerrero and Lama Darbelli

Chapter 4: Results

WES analysis, experiments, and interpretations were performed by TM. Dr. Genevieve Bernard, Kether Guerrero contributed to interpretations of WES analysis. Lama Darbelli provided guidance in analyzing AARS functional tests.

Chapter 5: Discussion

TM wrote this section in collaboration with Dr. Genevieve Bernard.

List of Figures

List of Tables

List of Abbreviations

ATP	Adenosine triphosphate
AA	Amino acid(s)
ARS	Aminoacyl-tRNA synthetase
AcL	Aminoacylation
RARS	Arginyl-tRNA synthetase
DARS2	Mitochondrial Aspartic acid tRNA synthetase
DARS	Aspartyl-tRNA synthetase
AD	Autosomal dominant
CNS	Central nervous system
CT	Computed tomography
DRP1	Dynamin-related protein 1
eIF2B	Eukaryotic initiation factor 2B
ETC	Electron transport chain
FIS1	Fission 1 homologue
EARS2	Mitochondrial glutamic acid tRNA synthetase
EPRS	Glutamyl-prolyl-aminoacyl-tRNA synthetase
HLD	Hypomyelinating leukodystrophy
4H	Hypomyelination, Hypodontia and Hypogonadotropic Hypogonadism
IEM	Inborn errors of metabolism
KSS	Kearns Sayre Syndrome
LS	Leigh's Syndrome
VWM	Leukoencephalopathy with vanishing white matter
MRI	Magnetic resonance imaging
mRNA	Messenger ribonucleic acid
aRS2	Mitochondrial aminoacyl-tRNA synthetases
mtDNA	Mitochondrial deoxyribonucleic acid
MELAS	Mitochondrial encephalopathy, lactic acidosis, and stroke-like episodes
MS	Multiple Sclerosis
MSC	Multisynthetase complex
MBP	Myelin basic protein
Non-HLD	Non-hypomyelinating leukodystrophy
nDNA	Nuclear deoxyribonucleic acid
OPA1	Optic atrophy protein 1
OXPHOS	Oxidative phosphorylation
PNS	Peripheral nervous system
PLP1	Proteolipid protein 1 gene
RRF	Ragged red fibers
ROS	Reactive oxygen species
POLR3	RNA polymerase III
tRNA	Transfer ribonucleic acid

Introduction and Thesis Objectives

Leukodystrophies are a group of rare genetically-determined disorders primarily affecting myelin of the brain. These diseases affect mainly children and are categorized into either hypomyelinating or non-hypomyelinating. Hypomyelinating leukodystrophies are caused by insufficient myelin deposition during development, whereas non-hypomyelinating leukodystrophies are caused by abnormal myelin maintenance. This thesis will focus on a cohort of patients with genetically unsolved mitochondrial leukoencephalopathies, a sub-category of non-hypomyelinating leukodystrophies. The first objective of this thesis is to use whole-exome sequencing to characterize this cohort of well-phenotyped patients to identify novel candidate genes/variants for further biochemical and functional testing. Functional validation of the variants found in one patient was performed as the second objective. In this patient with late-onset neurodegeneration, biallelic variants in the *Alanine t-RNA Synthetase* (*AARS*) were identified and validated with functional and biochemical testing.

Chapter 1: Literature Review

1.1 Leukodystrophies

1.1.1 Leukodystrophy classification

Leukodystrophies are genetically-determined disorders primarily affecting the cerebral white matter (i.e. myelin), with or without involvement of the peripheral nervous system (PNS) [1-3]. The majority of leukodystrophies are neurodegenerative diseases for which no curative therapy is available, and the underlying pathophysiology is not completely understood. Leukodystrophies are classified into hypomyelinating (HLD; abnormal myelin deposition during development) and non-hypomyelinating (non-HLD; abnormal myelin homeostasis), based on distinct magnetic resonance imaging (MRI) characteristics [2, 4]. In HLDs, on T1-weighted images, the white matter signal is normal or almost normal (i.e. hyperintense, isointense or slightly hypointense compared to grey matter structures), while on T2-weighted images, the white matter signal is slightly hyperintense (abnormal), compared to grey matter structures [4, 5]. In non-HLDs, compared to grey matter structures, the white matter signal is significantly hypointense on T1-weighted images (very abnormal) and very hyperintense on T2-weighted images (very abnormal) [4]. Patients with HLDs and non-HLDs typically present with motor regression, such as gait difficulties, and progress to complete loss of ambulation (if achieved) and of most, if not all, voluntary movements, months to year after disease onset. Patients with leukodystrophy also lose the ability to speak, to swallow and, in some instances, to see and hear [6]. The age of onset of leukodystrophies varies widely, from the neonatal period through adulthood, but children are most commonly affected. Collectively, there are more than 30 different types

of leukodystrophies affecting nearly 1 in 8000 individuals, making them at least as common as childhood Multiple Sclerosis (MS) [7].

1.1.2 Hypomyelinating leukodystrophies

HLDs are a group of disorders characterized by a permanent deficit in myelination (Table 1.1). This subgroup of diseases was initially thought to be caused by mutations in genes encoding proteins constituting the myelin sheath. Pelizaeus-Merzbacher disease (PMD) (OMIM #312080), an X-linked disorder, is the prototypical HLD. It is caused by mutations in *PLP1*, encoding *proteolipid protein*, one of the main mammalian structural myelin proteins [8]. In the last 15 years, it has become evident that there are multiple mechanisms underlying hypomyelination and that genes encoding for proteins involved in many other processes can be mutated in patients with HLD. Examples of such processes are transcription and translation. Indeed, mutations in genes encoding different subunits of RNA polymerase III (POLR3) machinery cause what is now thought to be the most common HLD: 4H (Hypomyelination, Hypodontia, and Hypogonadotropic Hypogonadism) or POLR3-related leukodystrophy [9-13]. RNA polymerase III is responsible for the transcription of small non-coding RNA, including all tRNAs. Interestingly, mutations in genes for tRNA synthetases have also been shown to cause HLDs, including *EPRS, RARS* and *DARS,* encoding for glutamyl-prolyl-aminoacyl-tRNA Synthetase, arginyl-tRNA synthetase and aspartyl-tRNA synthetase, respectively [14-16]. The general hypothesis is that mutations in genes encoding POLR3 subunits and tRNA synthetases lead to insufficient transcription during a critical developmental milestone, *i.e.* myelination, resulting in hypomyelination and neurodegeneration.

Table 1.1: Hypomyelinating leukodystrophies. Non-exhaustive list of hypomyelinating leukoencephalopathies. Disorders with significantly associated neuronal pathology were not included in this list.

Disease name	Inheritance	Mutated gene(s)
EPRS-related leukodystrophy	AR	EPRS [14]
Hypomyelination of Early Myelinating Structures	XL	PLP1 [17]
Hypomyelination with Atrophy of the Basal Ganglia and the cerebellum	Sporadic (AD)	TUBB4A [18], UFM1 [19]
Hypomyelination with Congenital cataracts	AR	FAM126A [20]
Hypomyelination with Brainstem involvement and Leg Spasticity	AR	DARS [16]
Hypomyelination with spondylometaphyseal dysplasia	XL	AIFM1 [21]
Oculodentodigital Dysplasia	AD	GJA1 [22]
Pelizaeus-Merzbacher	XL	PLP1 [23]
Pelizaeus-Merzbacher-like	AR	GJC2 [24]
POLR3-related leukodystrophy	AR	POLR3A [10], POLR3B [11], POLR1C [13], POLR3K [12]
RARS-related leukodystrophy	AR	RARS [15]
Sialic acid storage disease	AR	SLC17A5 [25]
SOX10-related disorder	AD	SOX10 [26]
Spastic ataxia-8 with hypomyelinating leukodystrophy	AR	NKX6-2 [27]
TMEM106B-related leukodystrophy	Sporadic (AD)	TMEM106B [28]

1.1.3 Non-hypomyelinating leukodystrophies

Non-HLDs represent a larger group of disorders with heterogeneous clinical and MRI presentations, as well as disease mechanisms. In all these diseases, myelination occurs normally during development and is followed by abnormal myelin maintenance/homeostasis. Common disorders in this category include Metachromatic Leukodystrophy (autosomal recessive, caused by mutation in *ARSA*, leading to Arylsulfatase A deficiency) [29], Krabbe disease (autosomal recessive, caused by

mutations in *GALC,* leading to Galactocerebrosidase deficiency) [30], and Adrenoleukodystrophy (X-linked, caused by mutations in *ABCD1*, leading to abnormal peroxisomal beta oxidation and the accumulation of very long chain fatty acids) [31].

Non-HLD, like HLD, can be the result of defective translational machinery. One example of this is Vanishing White Matter Disease, the most common leukodystrophy [7]. Vanishing White Matter Disease is caused by mutations in any of the five genes (*EIF2B1, EIF2B2, EIF2B3, EIF2B4, EIF2B5)* encoding the five subunits of the eukaryotic translation initiation factor 2B (eIF2B) [32]. eIF2B is responsible for the conversion of the protein synthesis initiation factor 2 (eIF2) from its inactive form (GDP-bound) to an active eIF2B-GTP complex, leading to the formation of the 43S complex required for initiation of mRNA translation [32-34].

Non-HLDs can also be caused by inborn errors of metabolism (e.g. Metachromatic Leukodystrophy, Krabbe, Adrenoleukodystrophy) and defects in intracellular organelles such as the mitochondria. [1, 6]. Mitochondrial leukoencephalopathies are a subgroup of non-HLDs characterized by unique MRI patterns and mitochondria-specific biochemical abnormalities [35]. Mitochondrial leukoencephalopathies are caused by mutations in genes encoding mitochondrial proteins which are either encoded by mitochondrial DNA (mtDNA) or nuclear DNA (nDNA). Examples include mitochondrial DNA depletion disorders (e.g. Kearns-Sayre syndrome [36]), disorders leading to complex I deficiency (e.g. *NDUFS1*, *NDUFA2*, etc. [37]), and disorders caused by mutations in genes encoding the mitochondrial aminoacyl-tRNA synthetases (ARS2) for glutamic acid (*EARS2*) and aspartic acid (*DARS2*) [38, 39].

1.2 Mitochondria

1.2.1 Mitochondrial structure and function

Mitochondria are essential double membrane-bound subcellular organelles that provide metabolic energy for the cell by generating adenosine triphosphate (ATP). Up to 2000 mitochondria can reside within a single somatic cell [40, 41].

There are on average 2.6 double stranded circular mtDNA molecules of 16,569 base pairs within a single mitochondrion, with each mtDNA molecule bearing slight alterations in their sequences [42]. When a variant is identified in mtDNA, it can exist in either of two states: 1) as a homoplasmic variant or 2) as a heteroplasmic variant. At the nucleotide position where the variant is found, if all the mtDNA molecules have the same nucleobase (i.e. adenine, cytosine, guanine, or thymine), it is considered a homoplasmic variant. If there are different nucleobases at that same nucleotide position in different mtDNA molecules, then it is considered a heteroplasmic variant. The percentage of heteroplasmy is the number of the different variants present in reference to homoplasmic state [43].

Mitochondria represent a duality in composition, as they are comprised of genes encoded by both nDNA and mtDNA (Figure 1.1) [40, 41]. There are over 1500 proteins involved in maintenance of mitochondrial homeostasis. mtDNA encodes 37 essential genes and, currently, 1121 nDNA genes are mitochondria-specific [43-45]. As such, mitochondria are subjected to deleterious effects inherited maternally through the mtDNA or by Mendelian inheritance through the nDNA. Of the 37 mtDNA genes, 13 are essential components of the oxidative phosphorylation (OXPHOS) complexes, 2 are ribosomal RNAs (rRNAs), and 22 are transfer RNAs (tRNAs) [43]. Complexes I, III-V are encoded by both nDNA and mtDNA while complex II is solely encoded by nDNA [46]. Localization

and import of mitochondrial proteins are regulated by mitochondrial targeting sequences, and translocation into the mitochondria is facilitated by proteins such as heat shock protein 70 (Hsp70) as well as others [47, 48].

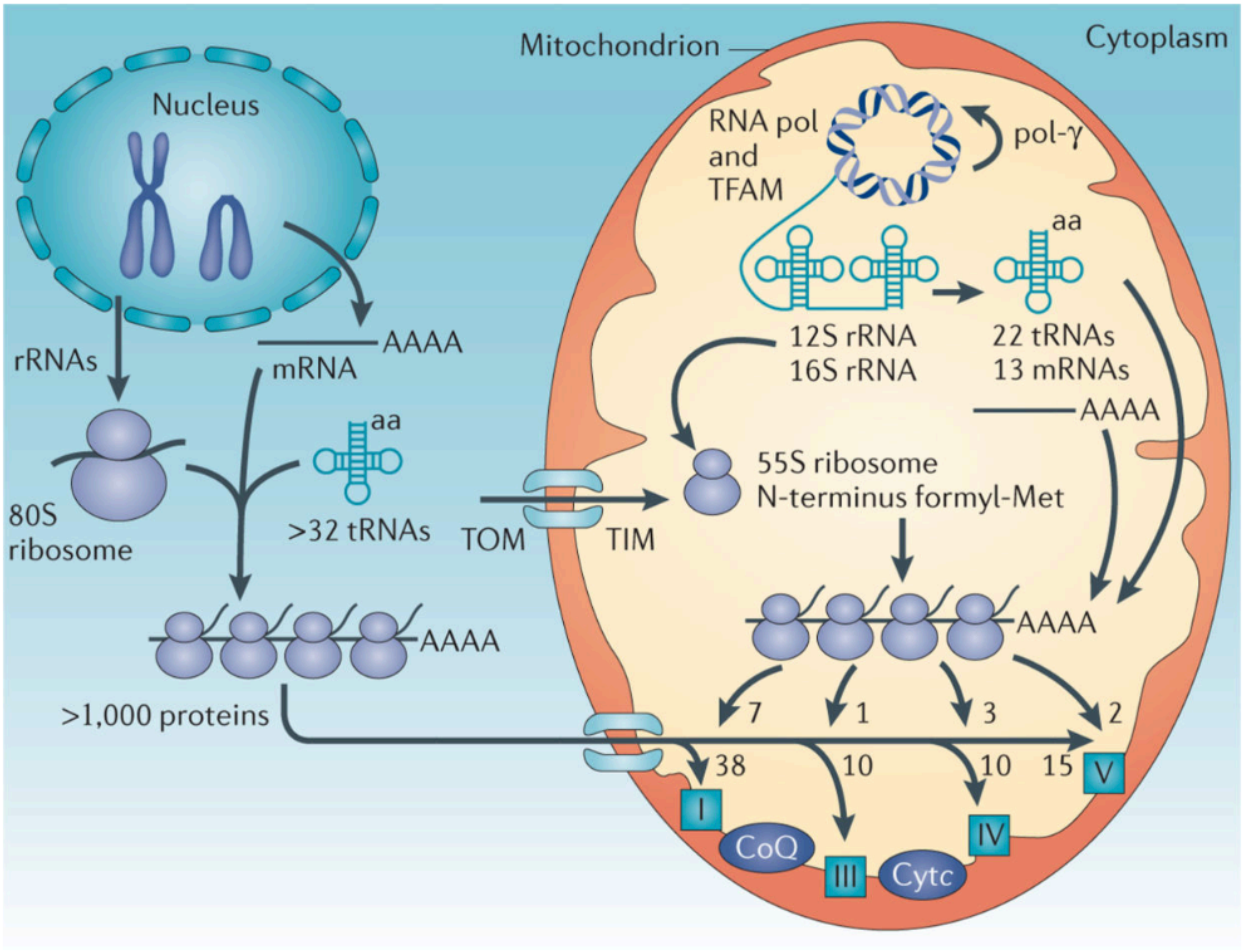


Figure 1.1 The dual encoding of the mitochondrial genome and protein biogenesis. Mitochondrial proteins are encoded by genes found in the nuclear DNA (nDNA) and the mitochondrial DNA (mtDNA). mtDNA encodes 13 proteins for complexes I and III-V, along with 22 tRNAs, and 2 rRNAs (the 12S and 16S subunits). The remaining proteins that function in the mitochondria are imported from the cytoplasm by an outer membrane and inner membrane transporters, TOM and TIM, respectively. This includes approximately 80 oxidative phosphorylation (OXPHOS) proteins that are imported into the mitochondria. mtDNA is replicated by a DNA polymerase γ (POLG) and transcribed by mitochondrial RNA polymerase with the mitochondrial transcription factor A (TFAM). Permission to reprint is granted from Springer Nature.

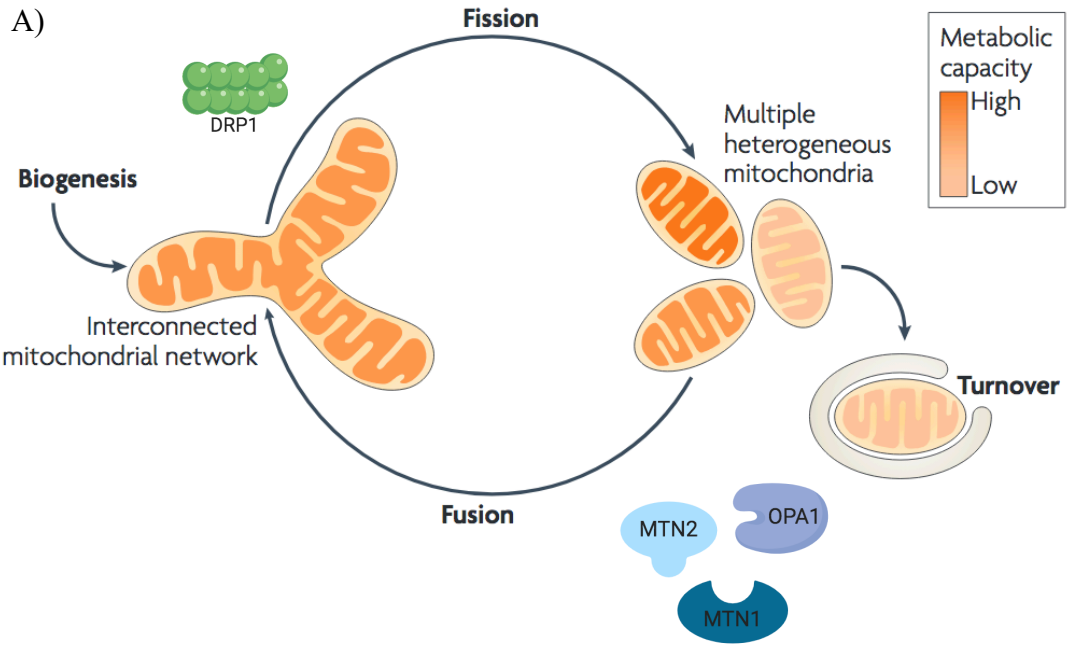
Mitochondrial biogenesis is dynamic, with constant fusion and division (fission) of the organelle (mitochondrial dynamics). These processes are regulated by evolutionarily conserved proteins (Figure 1.2A) [48]. Signalling proteins initiate the events of fusion. Mitofusion 1 and 2 are required for the fusion of the outer membrane whereas optic atrophy protein 1 (OPA1), a dynamin-related GTPase, is required for inner-membrane fusion [48, 49]. On the other hand, fission requires large GTPases to allow for membrane remodeling. Dynamin-related protein 1 (DRP1) and fission 1 homologue (FIS1), along with additional cofactors, are required to initiate the assembly of the fission structures on the mitochondrial membranes [48, 50, 51]. Mitochondrial dynamics are regulated by proteolytic processing [52], ubiquitylation [53], SUMOylation [54], phosphorylation [55, 56] and dephosphorylation [57].

The mitochondrion's primary function is to synthesize ATP (Figure 1.2B), which can be accomplished in two ways: 1) from glucose in aerobic conditions, and 2) as a byproduct during lactic acid fermentation in anaerobic conditions [58]. ATP production, from oxidative phosphorylation (OXPHOS) is accomplished by the movement of electrons through the five multiprotein mitochondrial complexes (I-V) embedded in the folds of the mitochondrial internal membrane, called cristae [59]. The electron transport chain (ETC) allows for the transfer of electrons through the inner membrane, specifically through complex I (CI, nicotinamide adenine dinucleotide (NADH) ubiquinone oxidoreductase), complex II (CII, succinate ubiquinone oxidoreductase), complex III (CIII, ubiquinone cytochrome *c* oxidoreductase), and complex IV (CIV, cytochrome *c* oxidoreductase). Complex V (CV, ATP synthetase) phosphorylates ADP to ATP [60].

The mitochondria are major sites for reactive oxygen species (ROS) generation. Elevated ROS levels has been shown to precede mitochondrial membrane changes that

eventually lead to the release of pro-apoptotic factors [61]. Moreover, the mitochondria can perform other functions including regulation of intracellular calcium homeostasis [62], pyruvate oxidation, the Krebs cycle, and metabolism of fatty acids, amino acids, and steroids.

Mitochondrial homeostasis dysregulation can occur from alterations in oxidative phosphorylation proteins, mitochondrial dynamics machinery, mitochondria-specific tRNAs, and mitochondrial ARS (ARS2) [59, 63-65]. Such dysfunctions result in the loss of ATP production or disturbances in other metabolic pathways. Mitochondria in this dysfunctional state are more likely to activate mitophagy, a mitochondrial specific breakdown, or apoptosis [66, 67]. Loss of too many mitochondria creates a lack of ATP production, and an accumulation of metabolites that can lead to disease.



B)

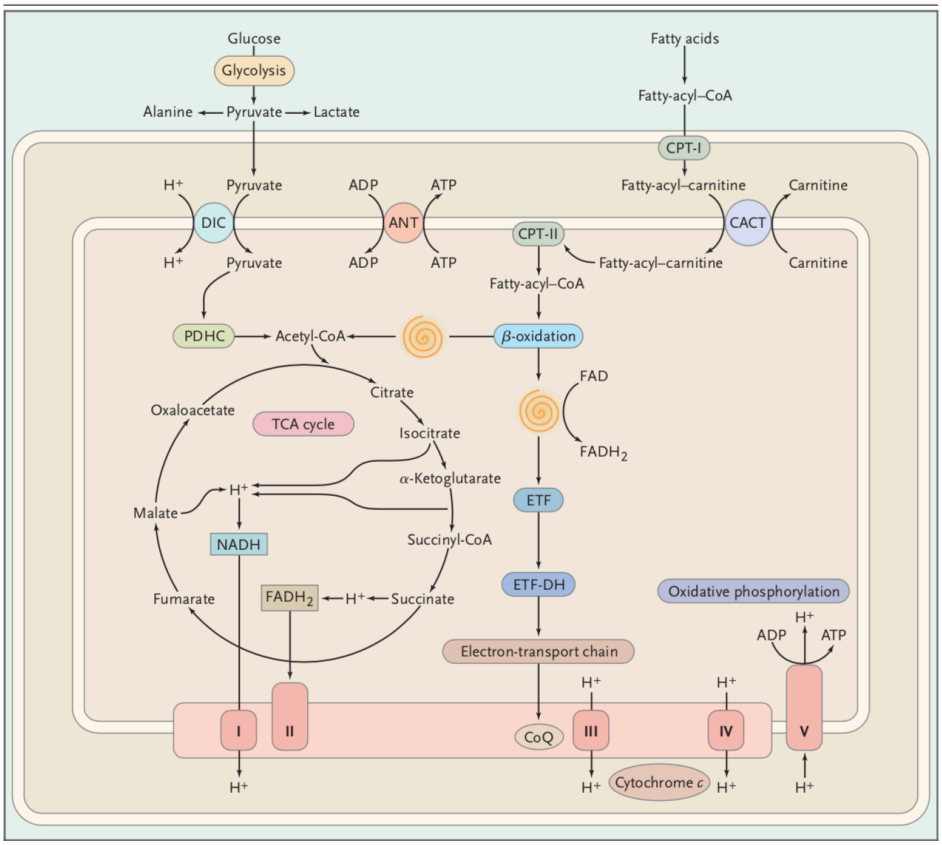


Figure 1.2 Mitochondrial biogenesis. (A) Mitochondria can undergo fission-fusion cycles. Fission is the division of a single organelle into two whereas fusion is the joining of two organelles into one. Fusion leads to an interconnected mitochondrial network that has a high metabolic capacity while fission leads to multiple heterogeneous mitochondria displaying different metabolic capacity. Fusion is predominantly coordinated by dynamin-related GTPases mitofusion 1 and 2 (MTN1 and MTN2) and Optic Atrophy 1 (OPA1). Fission also requires a large GTPase known as Dynamin-related protein 1 (DRP1), along with a collection of receptor proteins such as FIS1 (not shown). Defective mitochondria that have reduced membrane potential are tagged for mitophagy with PINK1 and Parkin (not shown). Permission to reprint is granted from Springer Nature. (B) The movement of electrons and metabolites through mitochondria allows for the diverse and interconnected activities of the mitochondria. Glycolysis, fatty-acid degradation, the tricarboxylic acid cycle (TCA), the electron transport chain, and oxidative phosphorylation are all essential to mitochondrial metabolism. Carnitine palmitoyltransferase I & II (CPT-I/II). Carnitine-acylcarnitine translocase (CACT). Dicarboxylate carrier (DIC). Adenine nucleotide translocator (ANT). Pyruvate dehydrogenase complex (PDHC). Electron transfer flavoprotein (ETF) – dehydrogenase (DC). Coenzyme Q (CoQ). Complex I-V (CI-V). Reproduced with permission from (DiMauro *et al*, 2003, N Engl J Med 2003;348:2656-68), Copyright Massachusetts Medical Society.

1.2.3 Mitochondria and neurological disease

Mitochondrial diseases occur at an incidence of 1:5000, and therefore, are considered rare [68]. These diseases have a heterogeneous multi-organ presentation that is explained by different energy demands of various cell types, with neurons and glia being specifically susceptible due to their high metabolic demands [69]. Other organs typically affected in mitochondrial diseases are: the heart, liver, kidney, skeletal muscle and peripheral nerves.

In mitochondrial diseases caused by mutations in mtDNA, pathogenesis occurs when defective mitochondria have accumulated past a tissue's specific threshold level [70]. The number of dysfunctional mitochondria is known as the mutation load. In most cases, 60-90% mutation load will result in a biochemical phenotype [67]. Cells, and by proxy, a tissue's mutation load, is based on the percentage of mitochondria with pathogenic variants present, over the total mitochondrial DNA. If a 50% level of heteroplasmy is reported, then half of the mtDNA contains an alteration.

Different types of mutations in the mtDNA have been described, including point mutations and large deletions, with strong genotype-phenotype correlation (Figure 1.3). For example, Kearns Sayre Syndrome (KSS), a rare early onset multisystem disorder, is caused by a specific large mtDNA deletion [66, 71].

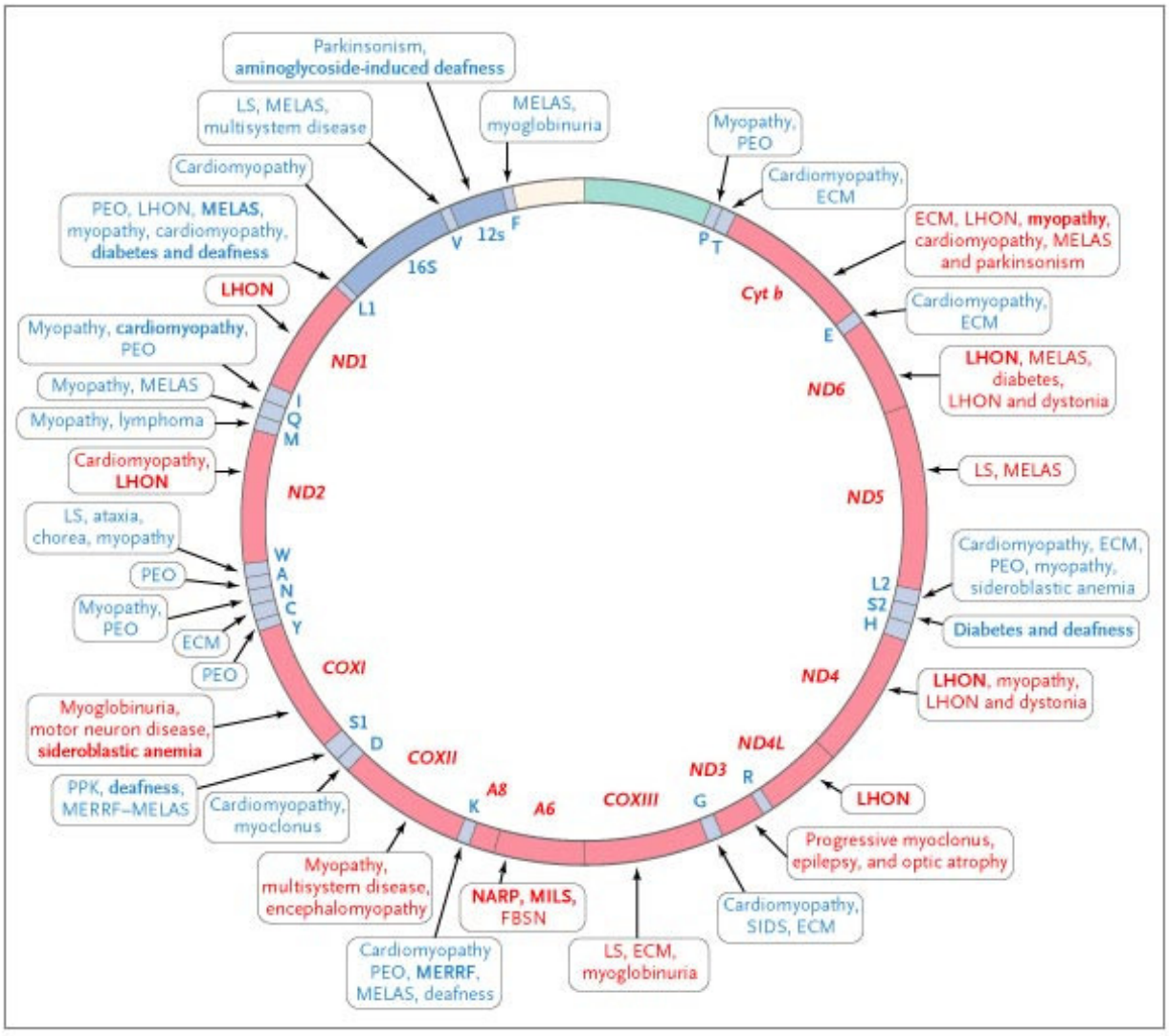


Figure 1.3 Morbid map of mitochondrial DNA. Specific positions on the mtDNA genes can harbor pathogenic variants leading to disease at certain heteroplasmic levels. Diseases shown in red are associated with protein coding mutations, while diseases shown in blue are associated with impaired mitochondrial protein synthesis. If high heteroplasmic variant levels are identified at one of these positions, it is then likely to cause the associated disease. Reproduced with permission from (DiMauro *et al*, 2003, N Engl J Med 2003;348:2656-68), Copyright Massachusetts Medical Society.

Mutations in nDNA and mtDNA genes lead to mitochondrial diseases caused by defects in the OXPHOS complexes. For example, mutations in genes encoding CI, CII, and CIV commonly cause Leigh Syndrome; mutations in genes encoding CII and CIII can cause liver diseases, and mutations in genes encoding CIII-CV can cause different subtypes of encephalopathies [67]. Mitochondrial mobility, which is regulated by OPA1's

interaction with microtubules, can be disrupted and cause an autosomal dominant (AD) optic atrophy [59, 72]. Defects in the localization and incorporation of mitochondrial proteins have also led to disease. One example of this is the X-linked recessive disorder Mohr-Tranebjaerg syndrome caused by mutations in the gene *TIMM8A* encoding the deafness-dystonia protein (DDP1), which is responsible for the import and insertion of inner mitochondrial proteins [73]. The pathogenesis of disorders such as Alzheimer's Disease, Amyotrophic Lateral Sclerosis, Parkinson's Disease and Huntington's Disease, have been partly associated to dysregulation of mitochondrial dynamics [48]. Finally, ROS accumulation has been suggested as a cause of aging-related diseases. Over time, ROS produced as by-products of oxidative phosphorylation cause alteration of the mtDNA [74], thus leading to a gradual accumulation of detrimental changes in the mtDNA, and decreased ATP production.

Mitochondrial leukoencephalopathy, a subcategory of non-HLD, is often suspected on the basis of specific MRI patterns and, in some instances, of elevated serum or cerebrospinal spinal fluid lactate concentration [75]. MRI pattern recognition is not infallible: sometimes, an MRI pattern suggestive of a non-mitochondrial disease may be revealed as mitochondrial, and vice-versa. Examples of this are the MRI pattern similarities seen in tRNA-synthetases-related diseases with mutations in the cytoplasmic aspartic acid *ARS (DARS)* and mitochondrial aspartic acid *ARS2 (DARS2).* Indeed, they both lead to MRI signal abnormalities of the same white matter tracts. Overall, identifying the molecular cause of a mitochondrial leukoencephalopathy requires a detailed analysis of the MRI pattern in combination with biochemical and molecular testing.

1.3 Aminoacyl tRNA Synthetases (ARS)

1.3.1 ARS canonical function and classification

The ARS genes are a family of ubiquitously expressed essential enzymes that are evolutionary conserved, being present in all three phyla: bacteria, archaea, and eukarya [76]. There are two sets of tRNA synthetases, one in the cytoplasm (ARS) and one for the mitochondria (ARS2). Both cytoplasmic and mitochondrial ARS' canonical function is to catalyze the esterification reactions that conjugate a cognate amino acid (AA) to its appropriate transfer RNA (tRNA). This process is known as aminoacylation (AcL) [77]. Once a tRNA is charged with its AA, a ribosome catalyzes the transfer of the AA from the tRNA to a growing polypeptide chain. AcL of a tRNA is done in a two-step reaction process. The first step is the binding of ARS to its cognate AA and an ATP molecule, forming an aminoacyl-adenylate intermediate (AA-AMP). This first step releases a pyrophosphate group. The second step consists of the binding of a tRNA molecule to ARS and the transfer of the AA to the tRNA. This second step is followed by the release of an AMP molecule and subsequently, the release of the charged tRNA (Figure 1.4) [78]. There is one ARS for every amino acid, except for EPRS, which is a bifunctional ARS responsible for the AcL of both glutamic acid and proline. There are 20 cytoplasmic ARS (ARS) and 18 mitochondrial ARS (ARS2). This discrepancy in number is due to the fact that the ARS responsible for the AcL of glycine & lysine, respectively GARS and KARS, function in both the cytoplasm and the mitochondria [63]. Additionally, the mitochondrial glutamine $tRNA^{Gln}$ synthetase has yet to be identified, but it is hypothesized that EARS2 works to mis-aminoacylate mt-$tRNA^{Gln}$ giving it the correct amino acid [63].

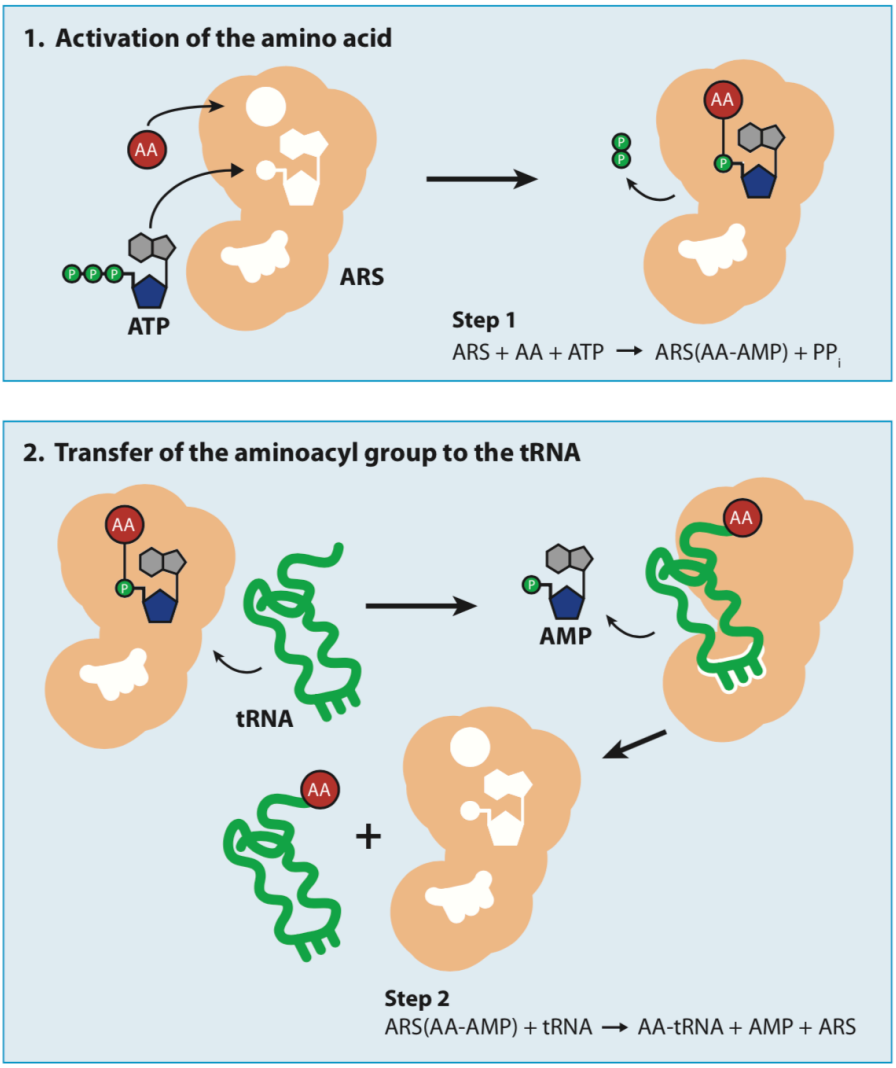


Figure 1.4 Aminoacyl-tRNA synthetase. The mechanism of aminoacylation reaction. First, the AA binds to ARS along with an ATP molecule. Next, an aminoacyl-adenylate intermediate (AA-AMP) is formed and a pyrophosphate is released. In the second step, a tRNA binds to ARS and the AA gets transferred to the tRNA and the AMP molecule is released. The enzyme is then able to repeat the process, and the tRNA is now considered charged. Together the reactions can be seen as: Step 1: ARS + AA + ATP → ARS(AA-AMP) and Step 2: ARS(AA-AMP) + tRNA → AA-tRNA + AMP + ARS. Permission to reprint is granted from Annual Reviews.

Every ARS has domains for catalysis and anticodon recognition. ARSs are grouped into class I and class II based on structural motifs surrounding their catalytic core [79]. The catalytic core of ARS is necessary for recognizing the anticodon of

tRNAs and to allow for the AcL on the 3' hydroxyl region of the tRNA [80]. Structurally, Class I ARSs have a classical Rossmann fold that consists of parallel β-strands, α-helices and signature motifs [79]. Class II ARSs are composed of primarily β-strands with additional low conserved structural motifs [79]. Each class can be further sub-categorized based on whether they can dimerize, which can lead to either homodimers or heterodimers (Figure 1.5). ARSs belonging to Class I can either be monomers or homodimers, while ARSs belonging to class II primarily function as homodimers with the exception of alanyl-tRNA synthetase (AARS), which functions solely as a monomer [81].

Editing activity is the other canonical function that is carried out by ARS proteins. Editing is only possible in the ARS proteins that contain certain structural domains. This function is found in 10 of the cytosolic ARS and it involves the removal of inappropriately charged amino acids from tRNAs, maintaining the high fidelity of translation [82]. There are two types of editing: pre-transfer editing and post-transfer editing [83]. Pre-transfer editing involves the hydrolysis of the incorrectly activated AA-AMP in the first step of AcL [82, 84]. Post-transfer editing occurs after the second step, when there is misacylation of tRNAs [82, 85].

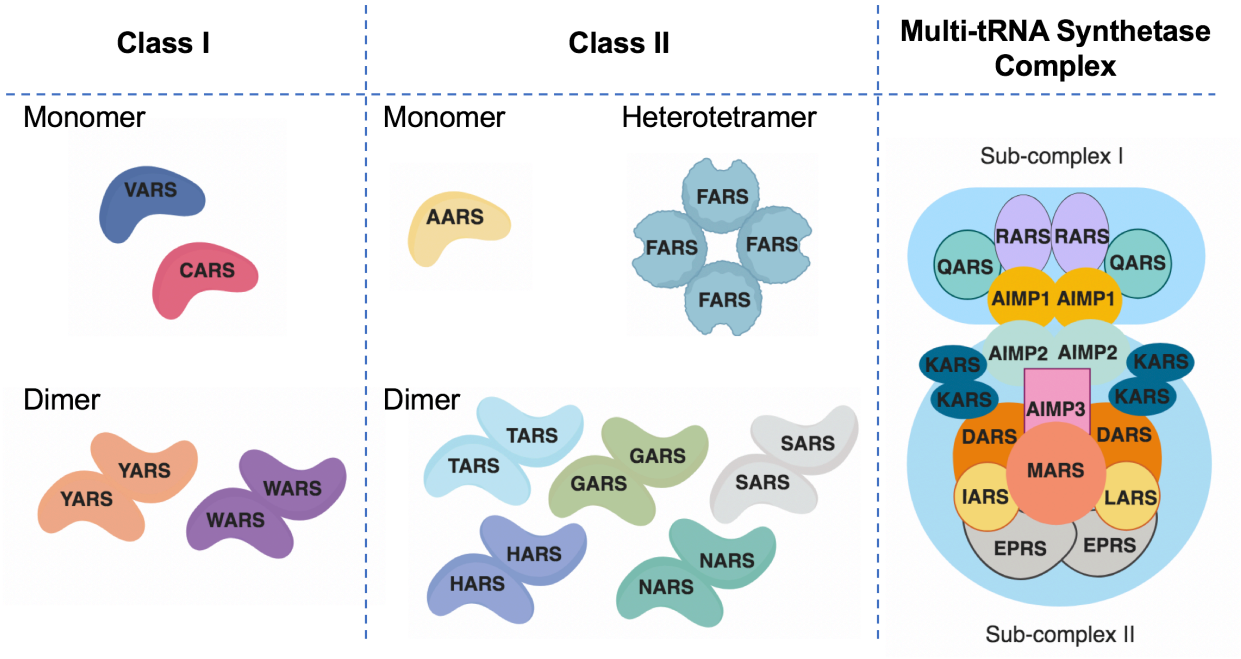


Figure 1.5 Structural Classification of ARS and the MSC complex. ARS can be classified as Class I or Class II based on structural motifs that surround the catalytic core. ARS belonging to Class I can be either monomers or dimers. Class II ARS could either be monomers, dimers or heteromers. Class I ARS proteins have a classical Rossmann fold that consists of parallel β-strands, α-helices and signature motifs. Class II is defined primarily by parallel β-strands. The multi-tRNA synthetase complex is composed of Class I and Class II ARS in subcomplexes I and II along with aminoacyl tRNA-synthetase interacting proteins (AIMP1-3). Only ARS proteins that are part of the MSC have been reported to have non-canonical functions. Permission to reprint is granted from Elsevier.

1.3.2 MSC

The multisynthetase complex (MSC) is formed of nine cytosolic ARS and three accessory proteins, called aminoacyl-tRNA synthetase-interacting multifunctional proteins (AIMP1, 2, 3) (Figure 1.5) [82]. The molecular weight of the MSC has been estimated to range from 1.0 to 1.5 MDa, and x-ray scattering revealed that the complex has an elongated multi-arm shape [86]. Heat-shock protein 90 (Hsp90) and AIMP1, 2 and 3 are important proteins for the MSC assembly [87]. Additionally, p43 [88], p38 [89], and p18 [90] are essential non-enzymatic proteins needed to maintain the structural integrity of the MSC. The ARS for Glutamine (QARS) and Arginine (RARS),

together with AIMP1, form sub-complex I. The ARS for Glutamine and Proline (EPRS), Methionine (MARS), Aspartic acid (DARS), Isoleucine (IARS), Leucine (LARS), and Lysine (KARS), together with AIMP1 and 2, form sub-complex II (Figure 1.5) [91]. RARS, when bound to the MSC, has been shown to aminoacylate arginine more efficiently than when unbounded [92]. This phenomenon is not limited to RARS but extends to the other ARS that are part of the complex [91].

1.3.3 Non-canonical functions

The non-canonical functions of ARSs are predominately based on their involvement in the MSC [91, 93]. The MSC's main function is related to translational efficiency, yet it plays a role in a diversity of non-canonical functions [91]. The MSC is involved in the regulation of gene transcription, rRNA transcription, tRNA maturation, apoptosis, inflammation, and immune response, amongst other roles. Yet, these functions are not fully understood [94-97].

There are 11 cytoplasmic ARS that are not part of the MSC complex: VARS, CARS, YARS, WARS, TARS, GARS, SARS, HARS, NARS, AARS, and FARS. These cytosolic ARS have no known non-canonical functions in humans [94]. On the other hand, these ARS in other organisms are known to have non-canonical functions. This is hypothesized to be due to these organisms not having all human ARS [76]. Even if they are lacking some human ARS, these organisms are able to perform the necessary AcL tasks, as other ARS misaminoacylate the necessary tRNAs, similar to how *EARS2* is believed to aminoacylate both glutamine and glutamic acid in humans [76].

1.3.4 ARS related diseases

Mutations in a number of *ARS* genes lead to a spectrum of neurological diseases: peripheral neuropathies, brain malformations, leukoencephalopathies, etc. The loss of

function of ARS proteins is thought to cause detrimental effects to protein translation: absence of charging of a specific tRNA would impede the biosynthesis of proteins, which, in certain organs and at critical times, would lead to hampered growth, causing disease (Figure 1.6) [91]. Elucidating the exact pathophysiology of ARS variants is complicated by the fact that not all variants lead to hypomorphic changes and it is likely that different diseases and even different pathogenic variants causing a given disease, lead to a different pathogenesis.

Inherited peripheral neuropathies are one of the most frequent causes of neurological disabilities. Charcot-Marie-Tooth (CMT) is the most common group of diseases leading to peripheral neuropathy with an incidence of 1 in 2500 individuals [98]. CMT is a progressive peripheral neuropathy that leads to muscle weakness and decreased peripheral sensations [97]. Currently, mutations in five *ARS* are associated with CMT: *AARS* [99], *KARS* [100], *GARS* [101], *HARS* [102], and *YARS* [103]. Mutations in *KARS* cause the only AR CMT, while the others are all inherited in an autosomal dominant fashion. Mechanisms implicated in CMT caused by mutations in *ARS* gene include alteration in AcL, alteration of non-canonical functions, altered dimerization, and aberrant localization [46].

Autosomal recessive ARS variants, on the other hand, have been shown to cause a heterogeneous group of disorders that is becoming increasingly recognized [77, 91]. All AR ARS-related diseases have CNS and/or cephalic senses (hearing/sight) phenotypes. Cytoplasmic ARS (ARS)-related disorders and mitochondrial ARS (ARS2)-related disorders share common features but also have distinct clinical characteristics [91]. The phenotypic spectrum is wide, much like the heterogeneous phenotypes seen in mitochondrial diseases. This heterogeneity is seen in diseases related to cytosolic ARS

and mitochondrial ARS. Additionally, different diseases involving the ARS part of the MSC have similar characteristics, which could be due to the disruption of translational activities associated with the complex as a whole, instead of the specific function of the single ARS [91].

Since the initial discovery of ARS-related diseases in 2003 using genetic mapping, an increasing number of these diseases have been described, thanks to the advent of next generation sequencing (NGS) [91, 101]. In 2010, only five *ARS* genes were associated to diseases (*DARS2*, *RARS2*, *AARS*, *GARS* and *YARS*) [101, 103-105], compared to now, where 30 *ARS* genes have been associated to diseases (Table 1.2). Alterations in *ARS* genes have been shown to lead to a plethora of leukodystrophies. Of all the *ARS* related to diseases, six out of the fourteen encoding for cytoplasmic ARSs and nine out of sixteen encoding mitochondrial ARSs have been shown to cause a leukoencephalopathy (Table 1.2) [14-16, 38, 106-108]. It has been demonstrated that disruption of canonical and non-canonical functions is involved in the pathophysiology of the diseases [103, 109-111]. ARS-related diseases are predominately caused by hypomorphic mutations, yet alterations to the monomer/dimer conformations, mislocalization, and reduced protein levels can also contribute to disease pathophysiology [110, 112-114]. It is not clear why certain *ARS* variants are transmitted through an AD versus AR mode [91]. In order to elucidate the molecular mechanisms that are disrupted due to mutated ARS proteins, detailed functional testing should be carried out to demonstrate a defect in canonical and/or non-canonical functions, before calling the specific variants pathogenic.

Table 1.2: Diseases Associated with ARS proteins. Cytoplasmic and mitochondrial *ARS* predominately cause a central or peripheral nervous system diseases, Leukodystrophy and genetic leukoencephalopathy associated genes are Underlined.

Gene	Inheritance	Cytoplasmic ARS	Ref
AARS	AD, AR	Epileptic encephalopathy, early infantile, 29 (MIM:616339) and Charcot-Marie-Tooth disease, axonal, type 2N (MIM:613287)	[99, 106]
CARS	AR	Microcephaly, Developmental Delay and Brittle Hair and Nails	[115]
DARS	AR	Hypomyelination with brainstem and spinal cord involvement and leg spasticity (MIM: 615281)	[116]
EPRS	AR	Leukodystrophy, hypomyelinating, 15 (MIM:617951)	[112]
GARS	AD, AR	Charcot-Marie-Tooth disease type 2D (MIM:601472) and Distal Hereditary Motor Neuropathy 5A (MIM:600794)	[101, 117]
HARS	AD, AR	Charcot Marie Tooth Syndrome (MIM:614504)	[118, 119]
IARS	AR	Growth retardation, impaired intellectual development, hypotonia, and hepatopathy (MIM:617093)	[120]
KARS	AR	Autosomal recessive deafness 89 (MIM:613916) and Charcot-Marie-Tooth disease, recessive intermediate B (MIM:613641)	[100, 121]
LARS	AD, AR	Infantile hepatopathy	[122]
MARS	AR	Charcot-Marie-Tooth disease, axonal, type 2U (MIM:616280), Hereditary spastic paraplegia (NA)	[123]
QARS	AR	Microcephaly, progressive, seizures, and cerebral and cerebellar atrophy (MIM:61570)	[124]
RARS	AD, AR	Leukoencephalopathy (MIM:616140)	[125]
VARS	AR	Severe DD, microcephaly, seizures	[126]
YARS	AD, AR	Charcot-Marie-Tooth disease, dominant intermediate C (MIM:608323)	[127]
Gene	**Inheritance**	**Mitochondrial ARS**	**Ref**
AARS2	AR	Ovarioleukodystrophy (MIM:615889) and Combined oxidative phosphorylation deficiency (MIM:614096)	[128, 129]
CARS2	AR	Combined oxidative phosphorylation deficiency 27 (MIM:616672)	[130]
DARS2	AR	Leukoencephalopathy with brain stem and spinal cord involvement and lactate elevation (MIM:611105)	[105]
EARS2	AR	Leukoencephalopathy with thalamus and brainstem involvement and high lactate (MIM:614924)	[38]
FARS2	AR	Alpers encephalopathy (MIM:614946)	[131]
HARS2	AR	Perrault syndrome (MIM:614504)	[132]
IARS2	AR	Cataracts, growth hormone deficiency, sensory neuropathy, sensorineural hearing loss, and skeletal dysplasia (MIM:616007); Leigh Syndrome (MIM:256000)	[133]
LARS2	AR	Perrault Syndrome (MIM:615300)	[134, 135]
MARS2	AR	Autosomal recessive spastic ataxia with leukoencephalopathy (MIM:611390)	[136]
NARS2	AR	Combined oxidative phosphorylation deficiency 24 (MIM:616234)	[137]
PARS2	AR	Alpers syndrome (MIM:618437)	[138]
RARS2	AR	Pontocerebellar hypoplasia type 6 (MIM:611523)	[104]
SARS2	AR	Hyperuricemia, pulmonary hypertension, renal failure and alkalosis (MIM:613845)	[139]
TARS2	AR	Combined oxidative phosphorylation deficiency 21 (MIM:615918)	[140]
VARS2	AR	Combined oxidative phosphorylation deficiency 20 (MIM:615917)	[140]
YARS2	AR	Myopathy, lactic acidosis and sideroblastic anemia (MIM:613561)	[141]

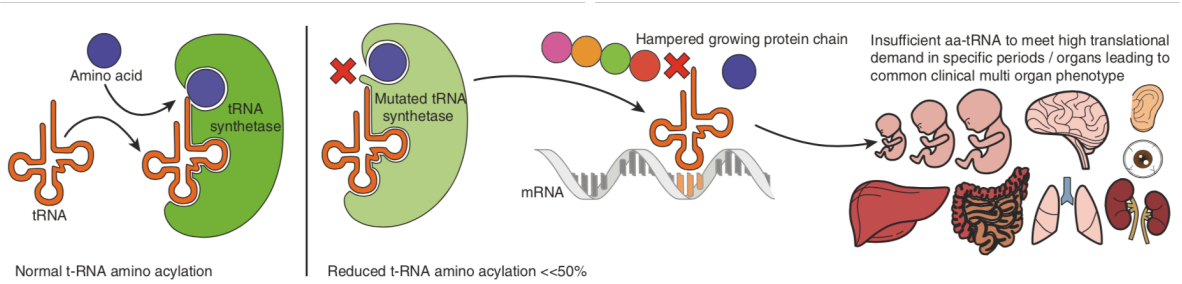


Figure 1.6 Autosomal recessive ARS deficiencies. Bi-allelic hypomorphic mutations in tRNA synthetase genes can lead to defective aminoacylation resulting in defective protein biogenesis. Results of the deficient enzyme activity leads to hampered protein production during critical developmental windows, causing tissue specific deficits. Permission to reprint is granted from Nature Springer.

1.4 Next-Generation Sequencing

1.4.1 Background

Normally, the diagnosis of a leukodystrophy is accomplished by clinical assessment with the identification of a unique MRI pattern [3, 4]. Genetic testing, and, if needed biochemical testing, follows to confirm the suspected diagnosis. In approximately half of the cases, this process of MRI pattern recognition and molecular/biochemical testing confirms the patient's diagnosis and genetic cause of leukodystrophy [142]. However, in the other half, a molecular diagnosis cannot be identified despite extensive investigations. These cases are considered "unsolved". In this project, we used NGS aimed at identifying a molecular etiology for unsolved patients with mitochondrial leukoencephalopathy phenotypes. Identification of a molecular cause allows for patients and families to receive needed information about the disease course while researchers gain better insight into the molecular pathology of the disease.

There are more than 7000 genetic diseases. Genetic diseases affect one in 30 individuals in the United States of America. They cause 1 in 6 children's hospital admissions and 1 in 5 infant deaths. The locations of the mutations in the genome are

known for more than 4000 diseases. The diagnostic odyssey can be long in patients affected by genetic diseases, often taking years. This delay in diagnosis is not without consequences.

Orphan genetic diseases are rare, although collectively, they are estimated to affect 40-82 individuals per 1000 live births. If congenital abnormalities are included, up to 8% of the general population are affected by a genetic disorder.

Traditional molecular testing is done by Sanger sequencing, i.e. amplification of PCR reactions and sequencing gene by gene, exon by exon. To streamline this, multiplexing was developed, allowing for sequencing of a small group of genes at a time. Sanger sequencing is time consuming and limited to the known disease-causing genes.

1.4.2 Molecular identification with Next-Generation Sequencing (NGS)

Developed in 2009, NGS allows for the high-throughput sequencing of genomic data, a more cost-effective method than Sanger sequencing, which requires testing of one gene/exon at a time [143, 144]. NGS allows for rapid interrogation of DNA, up to the entire genome, via massive parallel sequencing. Whole exome sequencing (WES) uses probe primers to amplify exons and intron-exon boundaries [143, 145]. Currently, WES runs for 100-150 base pairs (bp) per amplification and provides a relatively uniform minimum read depth of coverage 30x for ~99% of the exons, with sequences below that threshold possibly leading to false variant calling [146, 147]. When comparing WES, which only amplifies 1% of the genome to the alternative NGS method of whole genome sequencing (WGS), WES is still considered effective and more economical for new discovery rates [144, 146, 147]. The advantages and

disadvantages of WES versus WGS are presented in Table 1.3. NGS has revolutionized the discovery of genes, especially in the unidentified cases of leukodystrophies, which now only have a 20-30% unsolved rate [35, 148].

Table 1.3: Advantages and disadvantages of WES versus WGS.

WES	**WGS**
Deeper coverage, >80-fold coverage.	Lower, ~30-fold coverage
Covers coding regions with some bias	Covers coding and non-coding regions with less bias than WES
More economical	Still very costly
Will uncover incidental findings	Will uncover incidental findings
Allows for the identification of novel causing genes	Allows for the identification of novel causal genes

However, NGS has limitations. Specific technical equipment and knowledge is required, including advanced bioinformatic tools and the ability to store a large amount of data. Variant analysis and interpretation may be challenging, with numerous variants of uncertain significance by NGS dataset. Technical limitations are also important: 1) NGS has an accuracy of 99.99%, i.e. 300000 errors per sample, 2) there is a limited ability to detect certain types of mutations (i.e. deletions, duplications, trinucleotide repeats and methylation defects), and 3) some regions are not well covered, especially GC-rich regions. Even though NGS has been around for ten years now, there is still some education to do to clinicians, patients, and the general public regarding the advantages and disadvantages of performing NGS, including the potential to uncover incidental findings and the potential for discrimination, including by insurance companies.

When performing NGS, in order to maximize the probability of finding the patient's diagnosis, detailed phenotyping is required, in tandem with the hypothesized

inheritance pattern. A trio analysis (i.e. sequencing of the affected individual and his/her parents) allows for segregation analysis of potential variants, therefore helping with the diagnostic process.

WES analysis is performed by following the 2015 American College of Medical Genetics (ACMG) guidelines for variant classification [149]. These guidelines are essential in filtering, prioritizing, removing, and analyzing variants in candidate genes before validation. However, this workflow lacks some specificity, and refinement of these guidelines have been published as semiquantitative, hierarchical evidence-based rules for locus interpretation (Sherloc) [145]. Sherloc helps remove variants of uncertain significance, aiding in the identification of pathogenic variants. By improving analysis methods, researchers are able to better maintain a high degree of confidence in a disease-causing variant and be better overall at identifying novel diseases.

Chapter 2: Rationale and hypothesis

2.1 Rationale

Using NGS has allowed for a rapid interrogation of the human genome and pushed the efficacy of molecular investigations for unsolved leukodystrophy cases from 50-60% to 20-30% [35, 150]. A growing number of genes important for transcription and translation have been associated with leukodystrophies in recent years, such as genes encoding for aminoacyl-tRNA synthetases and subunits of RNA polymerase III [10, 13-16, 35, 106, 108, 151] Mitochondrial leukoencephalopathies, a subcategory of non-HLDs, are caused by pathogenic variants in mtDNA or nDNA encoding for mitochondrial homeostatic genes, and/or genes important for mitochondrial transcription and translation processes. These disorders are clinically and radiologically distinct from other non-HLDs, are associated with abnormal myelin maintenance, and have heterogeneous molecular etiologies [152].

In this study, we pursue NGS analysis of well-phenotyped patients with suspected mitochondrial leukoencephalopathies to identify novel candidate genes/variants. Confirmation of pathogenicity of these candidates requires additional gene specific functional evaluation. Out of the compelling candidates for validation, we chose to pursue a patient with *AARS* variants as a potentially novel autosomal recessive *AARS*-phenotype.

The dysregulation of myelin homeostasis in non-HLDs is based on a diverse set of pathogenic abnormalities in the transcriptional-translational machinery, in addition to inborn errors of metabolism. By utilizing NGS, the identified candidates and the subsequent functional studies allow for better understanding of non-HLD's pathophysiology, benefiting the rate of future molecular discoveries.

2.2 Hypothesis

NGS analysis of a cohort of well-phenotyped patients with suspected mitochondrial leukoencephalopathies will yield novel candidate genes/variants for validation.

2.3 Research objectives

1. To identify novel forms of mitochondrial leukoencephalopathies using whole-exome sequencing.
2. To Functionally characterize patient-specific variants in the *AARS* gene causing a novel phenotype.

Chapter 3: Materials and Methods

3.1 WES Methods and Analysis

3.1.1 Patients and DNA extraction from blood

Informed consent was obtained from all participants in this study. The project was approved by the research ethics board of the McGill University Health Centre and the Montreal Children's Hospital (11-105-PED, 2019-4972) or the patients' local research ethics board, in case of international referrals. The medical charts and brain MRIs of all patients were reviewed.

Genomic DNA was extracted from whole blood samples or fibroblasts using Puregene Blood Core Kit C according to manufacturer instructions (QIAGEN Sciences, Hilden, Germany). DNA quality was assessed on a Nanodrop (Thermo Fisher Scientific), indicated by A260/A280 ratios greater than 1.8, with typical yield between 16-50 µg of genomic DNA per whole blood sample.

3.1.2 Next Generation Sequencing (NGS)-Whole Exome Sequencing (WES)

WES was performed at the Genomic Medical Center of Children's Mercy-Kansas City using genomic DNA after extraction from whole blood. Exome enrichment was performed with Illumina TruSeq PCR Free sample preparation or Nextera Rapid Capture Exome kits according to the manufacturer's protocols [153, 154], with validation performed by real-time PCR. Library samples were sequenced on either the Illumina HiSeq 2000 or 2500 instruments with 2 × 100 nucleotide sequences, providing an average of 80-fold coverage. Sequences were aligned to the human reference genome (UCSC GRCh37/hg19) using the Burrows-Wheeler Aligner algorithm [155]. SAMtools was used to identify single nucleotide polymorphisms and small insertions and deletions, while the

Genome Analysis Toolkit (GATK) was used for variant calling [156, 157]. Variant annotation was done using ANNOVAR [158].

Variant interpretation was done according to the American College of Medical Genetics (ACMG) guidelines [149]. Based on these guidelines, variant prioritization was done using population data, phenotype, mode of inheritance, computational data, and, when available, functional data. More specifically, variants present in healthy population databases (ExAc, EVS and GnomAD) [159-161] with a frequency above 1% were filtered out. Variant priority was refined with the use of Online Mendelian Inheritance of Man (OMIM) and phenotypic relation to the disease or molecular pathway suspected from the clinical and MRI assessment [146, 149, 162]. We used in-house gene panels that were created for the different clinical characteristics (e.g. leukoencephalopathy, microcephaly, intellectual disability). These panel were created using the Genetic testing registry (https://www.ncbi.nlm.nih.gov/gtr/) and detailed literature search. We created a list of mitochondrial genes not yet linked to diseases and genes encoding proteins involved in important pathways for myelin maintenance. Variants that were not part of the panels were filtered based on presumed mode of inheritance. If parents of the affected patients were considered consanguineous, homozygous recessive variants were prioritized. In patients from unrelated parents, compound heterozygous variants were considered more susceptible to be disease-causing. If a dominant mode of inheritance was suspected, X-linked, de novo and heterozygous variants were prioritized. All WES results were screened for compound heterozygous, homozygous, and potential *de novo* heterozygous variants, which were further assessed by Sanger sequencing of the parents. *In silico* prediction tools such as Mutation Tester, SIFT, Provean, and PolyPhen2 were used to predict the pathogenicity of the variants.

The analysis process reduced the variants present in patients down to a few selected candidate genes per patient. The variants were then confirmed by Sanger sequencing in the patient and the parents, to verify segregation of the variants. As recommended by the ACMG guidelines, variants were then categorized as ‘pathogenic’, ‘likely pathogenic’, ‘variants of uncertain significance’, ‘likely benign’, or ‘benign’.

Patients without convincing pathogenic variant(s) but with interesting genes with variants of uncertain significance were submitted to GeneMatcher [163]. GeneMatcher is an online tool that allows clinicians and scientists to post genes of interest and to be connected to other clinicians and scientists who are interested in the same gene.

3.2 Validation of Candidate Genes

3.2.1 Sanger Sequencing and Primer Designs

Primer designs were based on reference sequences from UCSC GRCh37/hg19 genome. Amplicon products of ~500 bps were designed to cover the variants of interest using Primer3 (http://bioinfo.ut.ee/primer3-0.4.0/). Products were forward- and reverse-sequenced at the McGill University and Genome Quebec Innovation Centre using an Applied Biosystems 3730xl DNA Analyzer. Sequences were analyzed with SeqMan (DNAstar), and aligned to their respective genomic sequence, while all mitochondrial sequences were aligned to mtDNA sequence NC_012920 (www.mitomap.org).

3.2.2 Polymerase Chain Reaction of Genomic DNA

Touchdown PCR was used to amplify genomic DNA regions of interest. Primer mixtures were prepared at 40 μM and used at a final concentration of 0.8 μM per sample. 20 ng of total genomic DNA was used for amplification. Cycling conditions were 95°C for

10 minutes of incubation followed by 3 additional conditions. Condition one is 5 cycles; 95°C for 15 seconds for denaturation, 60°C for 20 seconds for annealing, 72°C for 45 seconds for extension and elongation. Condition two is also 5 cycles with annealing reduced to 58°C. Condition three is 40 cycles and annealing is further reduced to 54°C. Finally, there is one cycle at 72°C for 10 minutes for a final extension before cooling. Touchdown PCR was performed using the 2720 Thermo Cycler (LifeTechnologies). Product amplification was confirmed with 2% agarose gel and a 100 bp DNA ladder (FroggaBio) then sent for sequencing.

3.2.3 Mitochondrial DNA multiplex amplification

Suspected mtDNA variants were amplified with a mtDNA multiplex amplification protocol (Ochinnikov et al., 2016) on a 2720 Thermo Cycler (LifeTechnology) [164]. Primer mixtures were prepared at 5 μM, and 20 ng of total genomic DNA per sample. First round cycling conditions were 95°C for 2 minutes for incubation followed by 30 cycles: 95°C for 1-minute denaturation, 58°C for 2 minutes annealing, 72°C for 2 minutes extension and elongation. A final extension step is done at 72°C for 5 minutes. In the second-round conditions, only the extension and elongation at 72°C is reduced to 1 minute. Amplification of products were run on a 1% agarose gel with a 100 bp DNA ladder (FroggaBio) then sent for sequencing.

3.2.4 Long-Range PCR

Long-Range PCR (LR-PCR) is essential for large-scale insertion and deletion (INDEL) amplifications; Touch down-PCR cannot amplify insertions and deletions over roughly 1000 bp (1 kb). The combination of NGS and LR-PCR is efficient, reliable, and cost-effective. LR-PCR primers were designed with Primer3 and used at a concentration

of 40 µM and total genomic DNA amount of 20 ng. Taq Advantage 2 (TaKaRa) was used for LR-PCR as it can reliably process 1 kb/min during the elongation stage without much optimization [165]. Elongations times were adjusted according to the size of the suspected INDEL to allow for the adequate amplification times (ex. 4.5kb deletion, meant 5-minute elongation). Cycling conditions were 95°C for 1 minute for incubation followed by 35 cycles; 95°C for 30 seconds for denaturation, and 68°C for an amount of time based on the number of bp for annealing and extension. This is followed by one cycle at 68°C for 12 minutes for a final extension before cooling. Amplified products were confirmed with agarose gel using an appropriate percentage based on amplicon size with a 1 kb plus DNA ladder (GeneDireX) then sent for sequencing.

3.3 Confirmation of autosomal recessive inheritance of the AARS variants

We identified two candidate variants in *AARS* (NM_001605.2) by WES: c.295G>A; p.E99K (exon 3) and c.778A>G; p.T260A (exon 6). PCR was first performed with genomic DNA (gDNA) from the patient, mother, and father to confirm the inheritance of variants. Sanger sequencing showed that c.295G>A was inherited from the mother, but c.778A>G was not present in either parent, suggesting a potential *de novo* variant. In order to confirm that c.778A>G was a *de novo* variant and that the variants are on separate alleles (*trans),* complementary DNA (cDNA) amplification and cloning was performed. The cloning reactions were performed by using the TOPO pCR®-II vector containing a ß-galactosidase cassette (LifeTechnologies). The PCR products generated using cDNA and specific primers that covered exon 3 to exon 6 (ex3-6) of the *AARS* mRNA were sub-cloned into the TOPO pCR®-II vector. Plasmids with those inserts of *AARS* were transformed into competent cells with a heat shock protocol and spread on

LB agar coated with X-Gal for colony selection. X-Gal is hydrolyzed by ß-galactosidase, which is expressed by TOPO vectors that do not contain an insert and gives colonies with a blue colour when on the plate. On the other hand, vectors containing the *AARS* ex3-6 insert have a disrupted ß-galactosidase cassette, and thus the lack of ß-galactosidase expression, leading to translucent colonies. These colonies were selected and grown overnight in LB broth. Plasmid extraction was performed with a miniprep kit (QIAGEN) according to the supplier's protocol. Purified plasmids were loaded on a 1% agarose gel to test for bacterial DNA contamination and to verify the presence of the cDNA inserts after a restriction enzyme digestion. This digestion was done with the EcoRI-HF restriction enzyme (BioLabs). Two bands are expected from the digestion: one band at the 4kb mark which represents the plasmid backbone and the second band at the 700bp mark which represents the ex3-6 insert. After insert validation, plasmids were analyzed by Sanger sequencing to assess whether variants were present in the insert and if they were present in *trans*.

3.4 Primary Fibroblasts Cell Cultures from Patient and Controls

Patient-derived and age (+/- 2 years) and sex-matched control-derived fibroblasts were cultured in Dulbecco's Modified Eagle's medium (DMEM) (Wisent) and supplemented with 10% fetal bovine serum (FBS) (Wisent) at 37°C in a 5% CO_2 incubator.

3.5 Quantification of *AARS* mRNA Levels

To assess *AARS* transcript level, we performed quantitative PCR (qPCR) using cDNA from RNA extracted from patient and two control samples. Housekeeping gene for hypoxanthine guanine phosphoribosyltransferase 1 (*HPRT1*) was used as a reference gene as it does not have any known interactions with AARS as determined by

the BioGRIS interaction database. Table 3.1 shows the sequences of primers used. Primers' efficiency has been determined using the MIQE guidelines and primers efficiency is between 90-110%.

Table 3.1: Primers Designed Using Primer 3 for Identifying Candidate Variants.

gDNA				
Gene	**Reference #**	**Variant**	**Forward 5'-3'**	**Reverse 5'-3'**
AARS	NM_001605	exon 3: c.295G>A; p.E99K	agcccacgttttaccacatc	agaacccagttcccagtgtg
		exon 6: c.778A>G; p.T260A	gagctcctgacctcagatgg	ccagcccctacactcaaaaa
MBOAT7	NM_024298	exon 3: c.400_412del; p.Phe134ProfsTer12	attgcgtgcgagcaggtaag	cccaatachctacaaggagaag
		exon 4: c.738C>A; p.Tyr246Ter	tgctctcctctcacctctt	tgacaagcgctagcaacaa
PDHA1	NM_000284	exon 7: c.677G>A; p.Arg226His	atgcctgctacttctcctcc	ctctcggacgcacaggatat
WDR62	NM_025160	c.2470C>T; p.Pro824Ser	cggtcagtctcgtgagtcaa	gtgagcctgtgaagggagag
		c.4487T>A; p.Leu1496	ctgagctggtctccaccttc	ccatgctctccactgacaga
cDNA				
AARS	NM_001605	exon 3-6	GCCAATACCCAGAAGTGCAT	GCATCCTCAGCACCAACTTT
HPRT1	NM_021130	exon 2-3	TGCTGAGGATTTGGAAAGGG	ACAGAGGGCTACAATGTGATG

3.6 Assessing AARS Protein Levels

Protein samples were extracted from patient fibroblasts as well as from age- and sex-matched control fibroblasts. Cultured cells were washed once with ice cold PBS. Cells were lysed in RIPA buffer (Thermo Fisher Scientific) with EDTA-free protease inhibitor (Roche). The plates were incubated on ice for 20 minutes. Cell were scraped and harvested into 1.5 ml Eppendorf tubes and spun at 13,000g for 15 minutes at 4°C. Lysates were transferred to new tubes and protein concentrations were determined using Bradford reagent (Bio-Rad). 30 µg protein/well was separated by SDS-PAGE, transferred to nitrocellulose membranes using immunoblot TurboTransfer system (BioRad), blocked with 5% skim milk, and immunoblotted with the indicated primary antibody, AARS (1:1000) or ß-tubulin (1:2000), overnight at 4°C or 2 hours at room temperature. After washing, membranes were incubated with secondary antibody conjugated to horseradish

peroxidase one hour at room temperature and then visualized by ECL Prime reagent (GE Healthcare). Band intensity was measured and quantified by BioRad ImageLab Software.

3.7 Assessing Cellular Localization of AARS Protein

The localization of AARS was determined in the patient's fibroblasts using immunocytochemistry. Cells were grown on poly-D-lysine coated coverslips overnight. The cells were then washed with PBS and fixed with 4% paraformaldehyde. Cells were permeabilized with 0.2% triton-X100. Both patient and control fibroblasts were evaluated for AARS localization with the use of a primary antibody for rabbit anti-AARS (Abcam, 1:200) along with visualization of the cytoplasm with mouse anti-ß-tubulin (Sigma, 1:2000). Following incubation with the primary antibodies, the slides were washed and incubated with secondary IgG conjugated with Alexa Fluor 488 or 594 at 1:1000 dilution (Invitrogen by Thermo Fisher Scientific). DAPI was used for visualization of the nucleus. Images were acquired using EVOS M5000 (Invitrogen by Thermo Fisher Scientific).

3.8 Aminoacylation Ability of Patient AARS (Performed by Collaborators)

Aminoacylation was assessed from patient and controls fibroblast lysates. The lysates were incubated at 37°C for 10 minutes in a reaction buffer containing 50 mmol/L Tris buffer pH 7.5, 12 mmol/L $MgCl_2$, 25 mmol/L KCl, 1 mg/mL bovine serum albumin, 0.5 mmol/L spermine, 1 mmol/L ATP, 0.2 mmol/L yeast total tRNA, 1 mmol/L dithiotreitol, and 0.3 mmol/L labelled alanine. Trichloroacetic acid was used to terminate and wash the sample. Ammonia was used to release the labelled amino acid which was quantified by liquid chromatography-tandem mass spectrometry (LC-MS/MS). Glycine was used as an internal standard. Intra-assay variation was <15%.

3.9 Statistical Analysis

All data presented were produced in at least 3 independent experiments. Statistical analysis was performed using Prism software (GraphPad software). Significance of differences was evaluated with student's t-test. Significance was indicated as *$p<0.05$, **$p<0.01$ and ***$p<0.001$ and ****$p<0.0001$.

Chapter 4: Results

4.1 Description of the Patient Cohort

Patients included in this cohort were referred to Dr. Geneviève Bernard or seen directly in the Montreal Children's Hospital Leukodystrophy Clinic. Patients were suspected to be affected by a mitochondrial leukoencephalopathy based on MRI pattern and/or biochemical abnormalities. Not all cases had explicit family history. General patient phenotypic features are presented in Table 4.1, while more detailed patient histories are presented in Table 4.2.

Patient 1 is female in her teens. She presented with hypotonia and developmental delay. It was already determined that the patient did not have MELAS (mitochondrial encephalopathy, lactic acidosis, and stroke-like episodes), MERRF (myoclonic epilepsy with ragged red fibers), NARP (neuropathy, ataxia, and retinitis pigmentosa), Leigh Syndrome, or Krabbe based on testing done in clinical molecular laboratories.

Patient 2 is female in her 20's with chronic encephalopathy, ataxia, short stature, lack of pubertal changes, lactic acidosis, pigmented retinopathy, hearing loss, exercise intolerance and heart block. The MRI pattern and clinical presentation point to a mitochondrial disorder, most likely KSS, or possibly EARS2-related. There is no family history of mitochondrial disorder. No muscle biopsy was performed.

Patient 3 is male. Symptoms of the patient included progressive spastic diplegia, dystonia, nystagmus, with an MRI pattern highly suggestive of a complex 1 deficiency.

Patient 4 is female who presented at birth with microcephaly, central hypotonia, developmental delay, GI symptoms and tachycardia with slight lactic acid and pyruvic acid imbalances.

Patient 5 is an adult female with a neurodegenerative condition. She presented with developmental delay and motor regression. Ophthalmological examination revealed temporal pallor of the optic disks. Metabolic testing indicated lactic acid imbalance.

Patient 6 is a female presenting with neurodegeneration. Both of her parents are healthy. The family history reveals that the patient has two siblings, one male and one female, both healthy. Therefore, a recessive inheritance is suspected. Up to two years ago, the patient was only considered clumsy. Now at age 28, she has asymmetric spastic diplegia, optic atrophy, and cognitive decline with possible peripheral neuropathy. MRI indicated leukoencephalopathy with an MRI pattern very suggestive of AARS2-related disorder.

Patient 7 is a young male with a neonatal onset of neurodegeneration with hypotonia, and respiratory difficulty. MRI was compatible with a mitochondrial leukoencephalopathy.

Patient 8 is female. She presented in the neonatal period with epilepsy, severe global developmental delay, and dysmorphisms. MRI revealed a leukoencephalopathy while the MRS showed a lactic acid peak. A mitochondrial leukoencephalopathy was therefore suspected. A CGH array revealed a deletion at 1p36.31-p31.33. Since this does not explain the lactic peak on MRS, a second disorder was investigated.

Patient 9 is a male who has passed away. The patient presented at birth with axial hypotonia and severe spasticity. MRI indicated leukoencephalopathy, typical for complex I deficiency.

Patient 10 is a female who passed away at the age of 3 months. The patient showed microcephaly, dysmorphic features and an abnormal metabolic work-up. MRI showed leukoencephalopathy, which was thought to be mitochondrial in origin.

Patient 11 is a male who, after a period of normalcy, presented with rapid neurodegeneration leading to severe developmental regression, epilepsy and hypertonia. The brain MRI revealed a cystic leukoencephalopathy

4.2 List of Candidate Genes for Patients with Suspected Mitochondrial Leukoencephalopathies

WES analysis was performed on the 11 patients with suspected mitochondrial leukoencephalopathies (Table 4.1). A summary of the patients' presentation, candidate genes and Sanger validation of the variants and co-segregation analysis is presented in Table 4.1. We found a strong candidate gene, most likely the cause of the patients' diseases, in 5 out of the 11 (45%) analyzed patients with suspected mitochondrial leukoencephalopathies. For one of these patients (for patient 4 details, see Table 4.2), a known pathogenic variant [166] in a known disease-causing gene has been identified, moving this case to the solved category (1/11 or 9%). The remaining 4 cases (36%) generated strong candidate genes, after other potential candidates failed to pass the validation step (Table 4.3). In the remaining 6 cases (55%), no strong candidate genes were identified after validating with PCR and Sanger sequencing (Table 4.1; Table 4.3).

Excluding patient 4, the 4 cases with a remaining candidate gene were found to have variants that were not previously reported as disease causing (Table 4.4). 4/8 (50%) of the identified variants were not reported in ExAC or GnomAD (Table 4.4). None of the patients had variants in the same genes. Utilizing the ACMG classifications (Table 4.4), categorization of these variants showed that they were all to be likely pathogenic. Overall the inheritance was autosomal recessive in 3/5 (60%) cases, X-linked dominant in 1/5 (20%) case, and mitochondrial (secondary to suspected pathogenic variant in mtDNA) in 1/5 (20%) case. All genes, including those that did not pass validation are presented in Table 4.3.

Table 4.1: WES Analysis for the Suspected Mitochondrial Leukoencephalopathy Patient Cohort

Patient #	Diagnosis	Candidates tested and ruled-out	Status
1	• Female • Hypotonia, developmental delay • MRI: Metabolic, mitochondrial	*MT-ATP6, MT-ND1 MT-TD, TRAP1, ETFDH, ACAD9, AKR1B15*	Strong candidate – *MBOAT7* Put on GeneMatcher
2	• Female • Leukoencephalopathy, progressive spastic diplegia, nystagmus, dystonia • MRI: Kearns-Sayre syndrome (KSS) like	*MT-TQ, MT-RNR2, MT-ND2, POLRMT, MARS, POLR3A, SERAC1*	Strong candidate – break points m.7950-15705
3	• Male • Leukoencephalopathy, progressive diplegia, dystonia, nystagmus • MRI: typical for complex I deficiency	*MT-ND1, EEF1A1*	Strong candidate - *WDR62*
4	• Female • Microcephaly, central hypotonia, developmental delay, GI issues, tachycardia • Lactic acid or pyruvic acid abnormalities	*MT-RNR1, MT-RNR2, MT-CO2, MT-CO3, MT-ND5, MT-ATP6, MT-CYB, PDPR*	Strong candidate – *PDHA1*
5	• Female • Progressive neurodegenerative condition, developmental delay, motor degrative, temporal pallor • MRI: suspected mitochondria related	*PDPR, MT-RNR1, MT-RNR2, MT-CO2, MT-CO3, MT-ND5, MT-ATP6, MT-CYB,*	Unsolved
6	• Female • Late onset leukoencephalopathy acquired spastic diplegia, cognitive regression • MRI: *AARS2*-like	*ISPD*	Strong candidate – *AARS*
7	• Male • Leukoencephalopathy, progressive mental deterioration, hypotonia, respiratory issues	*ABCD2, MIPEP, PMM2, APE41, MRPL15, SLC16A7, KMT2D, GPKOW*	Unsolved
8	• Female • Epilepsy, severe global delay, dysmorphisms • MRI: lactic acid peak indicating possible mitochondrial disease in a patient known for del 1p36.31-p31.33	*VPS13C, FOXD4L3, RGPD3*	No candidate to explain high lactate on MRS
9	• Male • Leukoencephalopathy, axial hypotonia, severe spasticity • MRI: typical for complex I deficiency	*MRPL48, NUBPL*	Unsolved
10	• Female • Microcephaly, facial dysmorphisms • MRI: Metabolic leukoencephalopathy	*PDSS1, SLC25A30*	Unsolved
11	• Male • Normal child until severe developmental regression • MRI: Cystic leukoencephalopathy	*ABCA9, ALKBH7, KMT2C*	Unsolved

Table 4.2: Detailed Patient Notes used for WES analysis.

Patient	Sex	Age of Onset/ Diagnosed	Phenotypic features	Additional Information	Notes	Pedigree*
1	Female	Diagnosed during birth to 2 yrs. Now 9	MRI: suggestive for metabolic disorder and leukodystrophy. Global developmental delay. Born with torticollis, treated by chiropractor till 9 mts. Early postural problems, independent siting, and later standing only with support. Truncal hypotonia. Metabolic abnormal for amino acids: alpha-amino-n-butyric acid, alanine, histidine, serine; acylcarnitine: mild ketosis, mild elevation of hydroxybutyryl carnitine. Muscle coordination problems, hypotonia, serum amino acids, acylcarnitine, ketosis	Not MELAS, MERF NARP, Leigh, Krabbe, SURF	Normal ophthalmology, cardiac, hematology, GI, and lungs. Normocephalic.	Clean pedigree, of parents. Exception diabetes of maternal uncle. Maternal grandmother, thyroid issues. Maternal grandmother's first cousin once removed (6 yrs.) is wheelchair bound due to muscular dystrophy. Maternal grandfathers first cousin once removed had biliary atresia and passed away first yr. Another maternal grandfathers first cousin once removed died at age 16 due to primary carnitine uptake deficiency.
2	Female	At birth, around 20	MRI: showed chronic encephalopathy. Short stature, absence of pubertal change, failure to thrive, pigmentary retinopathy, hearing loss, exercise intolerance, second-degree heart block. Progressive ataxia, and recent mental decline. Lactic acidosis	Muscle biopsy showed unremarkable structural pathology. Normal ETC enzyme activity. Not identified for common mutations or deletions in mtDNA.		Parents are normal
3	Male	Diagnosed around 15 mts, young	MRI: shows leukoencephalopathy. Progressive spastic diplegia, swelling difficulties, nystagmus, other vision problems, frequent colds, dystonia, GI issues. Amino acids: mild increase in BCAAs	Head relatively microcephalic, atraumatic, altricial helixes of ears and mild hypertelorism with a flattened midface	Normal cardiac, respiratory, skin, hepatomegaly, or splenomegaly	Mother blind in one eye. Maternal half-brother to maternal grandmother diagnosed with leukodystrophy and died in his 40s. Maternal great great grandmother brother and three nephews with similar presentation.
4	Female	Onset as young child	MRI indicate leukoencephalopathy. Microcephaly, central hypotonia, developmental delay, failure to thrive, GI issues, tachycardia, possible lactic acid or pyruvic acid imbalance.		Normal respiration, hematology, genitourinary, endocrine, immunology, skin.	Limited on parental side. Maternal side clean. Mother only child.
5	Female	Adult	MRI: suspected mitochondria. Progressive neurodegenerative condition, developmental delay, motor degrative, temporal pallor related motor degeneration, speech, lactic acid imbalance	Ruled out TCTN1		Family history of anxiety and depression, no pedigree available
6	Female	Adult	MRI: suggested AARS2-like. Abnormalities started at 26yrs. Cognitive regression, acquired spastic diplegia of lower extremities, headaches, progressive difficulty walking, dementia, tremors in feet, bilateral inferior altitudinal and superior arcuate field defects, no seizures	Some early symptoms of Poor eye/hand coordination, and running was uncoordinated		Both parents healthy, and both male and female siblings healthy
7	Male	Birth - dec 3 years	MRI: suggestive of mitochondrial leukoencephalopathy. Progressive mental deterioration, hypotonia, respiratory issues, tachycardia, no medical history in family.	Initially thought MS due to appearance of encephalomyelitis		Pedigree not available
8	Female	Few weeks after birth - 4 yrs.	MRI: indicates leukoencephalopathy. Lactic acid peak on MRS indicates possible mitochondrial disease. Patient known a known del at 1p36.31-p31.33. Epilepsy, severe global delay, facial dysmorphisms, spasticity, myopia	Ruled out AARS2, and consanguineous parents		Pedigree going back two generations show no disorders
9	Male	Birth- dec 3 months	MRI indicated a complex 1 deficiency. Not likely metabolic, axial hypotonia, severe spasticity that appeared at birth, diffuse MRI, mild swelling.	EEG showed diffuse continuous epagomenic activity	No grey matter involvement	No pedigree past heathy parents
10	Female	Birth- dec 3 months	MRI show diffuse white matter abnormalities, indicating a metabolic disorder, small head, facial dysmorphic features, normal outside of that, no evidence of lipidica in GI, no significant pathology abnormality, no inclusions, fatty replacements	EEG was abnormal. cerebral atrophy, leukoencephalopathy, Ruled out Aicardi Goutières syndrome, Ruled out tRNA synthetase defects	Pedigree show patents not affected, but other pregnancy did not pass 10 wks. Paternal father has dementia	
11	Male	15 mts old - 5 yrs.	MRI: Cystic leukoencephalopathy acute. Normal child until severe developmental regression, brought on after seizure. Hypertonia, severe dystonia, ataxia, cortical blindness, bradycardia	Metabolic abnormalities with low lactic acid	Extensive pedigree gives diverse history of mental disease but parents healthy	

Table 4.3: All Variants Identified and Tested from WES Analysis

Patient	Gene	Transcript	Genomic position	Variant
1	MBOAT7	NM_001146056	19:54687485	exon 4: c.519C>A; p.Tyr246Ter
			19:54684606	exon 3: c.181_193del; p.Phe61ProfsTer12
	ETFDH	NM_004453	4:159627503	exon 11: c.1448C>T; p.Pro483Leu
	ACAD9	NM_014049	3:128629608	exon17: c.1717G>A; p.Val573Met
	TRAP1	NM_016292	16:3729770	exon 4: c.493A>G; p.Gln165Glu
			16:3714367	exon 12: c.1477G>A; p.Gly493Ser
	MT-ATP6	NC_012920	mtDNA	m.9142G>A; p.Thr206Met
	MT-ND1	NC_012920	mtDNA	m.4025C>A; p.Thr240Met
	MT-TD	NC_012920	mtDNA	m.7521G>A
2	POLRMT	NM_005035	19:633470	exon 1: c.43C>G; p.Arg15Gly
			19:624763	exon 5: c.1096C>T; p.Pro366Ser
	POLR3A	NM_007055	10:79769427	exon 14: c.1771-6C>A
	MT-TD	NC_012920	mtDNA	m.4394C>T
	MT-ND2	NC_012920	mtDNA	m.5178C>A; p.Leu237Met
	MT-RNR2	NC_012920	mtDNA	m.2649T>C
3	WDR62	NM_001083961	19:36587931	exon 21: c.2470C>T; p.Pro824Ser
			19:36595845	exon 32: c.4487T>A; p.Leu1496Gln
	PCDH15	NM_033056	19:55582753	exon 33: c.4733T>C; p.Val1578Ala
			19:55782743	exon19: c.2435T>C; p.Ile812Thr
	EEF1A1	NM_001402	6:74228163	exon 6: c.943G>C; p.Val315Leu
			6:74227951	exon 7: c.1066G>T; p.Gly356Cys
	NADK2	NM_001085411	5:36242196	exon 3: c.438A>C: p.Glu146Asp
			5:36226617	c.-296+2T>C
	TUBB2A	NM_001069	6:3154692	exon 4: c.743C>T; p.Ala248Val
	MT-ND3	NC_012920	mtDNA	m.10144A>G; p.Gly29Asp
	MT-RNR1	NC_012920	mtDNA	m.710T>C
	MT-RNR2	NC_012920	mtDNA	m.2768T>C
4	PDHA1	NM_000284	X:19373540	exon 7: c.677G>A; p.Arg226His
	MT-RNR2	NC_012920	mtDNA	m.3150T>C
	MT-ND1	NC_012920	mtDNA	m.1738T>C; p.Ile10Thr
	MT-TM	NC_012920	mtDNA	m.4418T>C
	MT-CO3	NC_012920	mtDNA	m.9391C>T; p.Thr62met
	MT-ND4	NC_012920	mtDNA	m.11016G>A; p.Ser86Asn
5	PIGX	NM_00116630	3:196439516	exon 1: c.112+1G>T
	PDPR	NM_017990	16:70176181	exon 12: c.G1360A; p.G454S

	MT-RNR1	NC_012920	mtDNA	m.750A>G
	MT-RNR1	NC_012920	mtDNA	m.1438A>G
	MT-RNR2	NC_012920	mtDNA	m.3206C>T
	MT-ATP6	NC_012920	mtDNA	m.8860A>G; p.Thr112Ala
	MT-CYB	NC_012920	mtDNA	m.845A>G; p.Thr194Ala
	MT-CO3	NC_012920	mtDNA	m.9391C>T; p.Thr62Met
6	AARS	NM_001605	16:70310907	exon 3: c.295G>A; p.E99K
			16:70304137	exon 6: c.778A>G; p.T260A
	ISPD	NM_001101417	7: 16298080	exon 7: c.C904A; p.Q302K
7	ABCD2	NM_005164	12:39947925	exon 7: c.A2012T; p.H671L
			12:39980001	exon 10: c.G1745A; p.R582H
	MIPEP	NM_005932	13:24384020	exon 15: c.A1697G; p.K566R
			13:24384062	exon 15: c.C1655A; p.P552Q
	PMM2	NM_000303	16:8898644	exon 3: c.G199A; p.V67M
			16:8898674	exon 3: c.A229C; p.K77Q
	APE41	NM_001252127	15:51207765	exon 3: c.G118T; p.V40L
			15:51285764	exon 17: c.A2063G; p.K688R
	MRPL15	NM_014175	8:55049206	exon 2: c.G244T; p.G82W
			8:55055298	exon 4: c.A505G; p.K169E
	SLC16A7	NM_004731	12:60168736	exon 4: c.660delA; p.P220fs
			12:60173486	exon 2: c.51delA; p.Gfs
	KMT2D	NM_003482	7:151874276	exon 10: c.G2344A; p.V782M
	GPKOW	NM_015698	x:48978863	exon 3: c.A341C; p.K114T
8	VPS13C	NM_020821	15:62277179	exon 17: c.1598C>A; p.Thr533Asn
	FOXD4L3	NM_199135	9:70918118	exon 1: c.251_252del; p.Ser84Ter
			9:70918631	exon 1: c.764C>T; p.Pro255Leu
	RGPD3	NM_001144013	2:107049698	exon 16: c.2249T>C; p.Val750Ala
			2:107074109	exon 19: c.166C>T; p.Arg56Ter
9	MRPL48	NM_016055	11:73536809	exon 5: c.G292T; p.G98X
	NUBPL	NM_025152	14:32142591	exon 5: c.G413A; p.G138D
10	PDSS1	NM_014317	10:26986729	exon 1: c.89G>T; p.Gly30Val
	SLC25A30	NM_001010875	13:45976492	exon 6: c.407C>T; p.Ala136Val
			13:45971443	exon 9: c.784T>C; p.Tyr262His
11	ABCA9	NM_080283	17:67018000	exon 7: c.C916T; p.L306F
			17:67031827	exon 18: c.C2284A; p.L762I
	ALKBH7	NM_032306	19:6374829	exon 4: c.G511A; p.A171T
	KMT2C	NM_170606	7:151945330	exon 14: c.C2189A; p.S730Y
			7:151945349	exon 14: c.A2170T; p.K724X

Table 4.4: Pathogenicity of the Candidate Genes as per ACMG guidelines

Patient	Gene	Transcript ID	GP	Variant	Interpretation of Pathogenicity	MAF (Allele count/ number)	Criteria	CADD/ GERP-++ RS Score
1	*MBOAT7*	NM_001146056	19:54687485	Exon 4: c.519C>A; p.Tyr246Ter	Stop codon – Likely Pathogenic	NR	PM2, PM3, PM4, PP3, PP1	13.155/ 4.59
			19:54684606	exon 3: c.181_193del; p.Phe61ProfsTer12	Frameshift deletion - Likely Pathogenic	NR	PM2, PM3, PM4, PP3, PP1	-
2	mtDNA deletion	NC_012920	-	m.7950_15705del	7,750 bp mtDNA deletion - KSS	-	-	-
3	*WDR62*	NM_001083961	19:36587931	Exon 21: c.2470C>T; p.Pro824Ser	missense - Likely Pathogenic	0.00001213 (3/247252)	PM1, PM2, PM3, PP3, PP1	4.84/ 5.93
			19:36595845	Exon 32: c.4487T>A; p.Leu1496Gln	missense - Likely Pathogenic	0.00000418 3 (1/239090)	PM2, PM3, PP3, PP1	3.834/ 4.89
4	*PDHA1*	NM_000284	X:19373540	Exon 7: c.677G>A; p.Arg226His	missense – Pathogenic (Reported by Shin *et al.* 2017)	NR	PS4, PM1, PP3	3.642/ 5.76
6	*AARS*	NM_001605	16:70310907	Exon 3: c.295G>A; p.E99K	missense - Pathogenic	0.00000397 6 (1/251490)	PS3, PM1, PM3, PP4, PP3, PP1	7.863/ 5.9
			16:70304137	Exon 6: c.778A>G; T260A	missense – Pathogenic	0.00000397 6 (1/251480)	PS2, PS3, PM3, PP4, PP3, PP1	5.728/ 5.91

Abbreviations: GP: Genomic Position; MAF: Minor Allele Frequency; KSS: Kearns-Sayre Syndrome; NR: not reported in ExAC or GnomAD databases. Interpretation of the variant was performed based on the ACMG guidelines. First letter P: pathogenic. Second letter S: strong; M: moderate; P: supportive. Mitochondrial variants' pathogenicity criteria are not described in the ACMG guidelines and this was determined based on MitoWheel (www.mitowheel.org), MITOMAP (www.mitomap.org), MitoBreak (http://mitobreak.portugene.com/) and ANNOVAR annotation tool [158]. CADD: Combined Annotation-Dependent Depletion (raw score); GERP++ RS: Genomic Evolutionary Rate Profiling (rejected substitutions). Both CADD and GERP++ RS cannot interpret INDELs.

4.3 AARS Variants are Present in *Trans*

The MRI pattern of patient 6 strongly suggested AARS2-related disorder, also known as progressive leukoencephalopathy with ovarian failure (OMIM #615889), an

autosomal recessive disorder caused by mutations in the gene encoding the mitochondrial alanyl-tRNA synthetase (AARS2). However, WES analysis did not show any variants in *AARS2*, but instead two variants were found in *AARS*: c.295G>A (p.E99K) and c.778A>G (p.T260A) (Table 4.4). No other strong candidate gene was identified. Since DARS-related and DARS2-related disorders have been shown to have similar involvement on MRI [167], we hypothesized that a similar MRI pattern could be seen in AARS and AAR2-related disorders. Of note, *AARS* is already associated with autosomal dominant Charcot Marie Tooth, Axonal type 2 (OMIM #613287) and autosomal recessive epileptic encephalopathy, early infantile, 29 (EIEE 29, OMIM #616339) [99, 106]. Our patient's phenotype and *AARS* variant were different compared to what has been previously described in these two diseases. We therefore hypothesized that our patient had a mild form of the autosomal recessive leukoencephalopathy found in patients with EIEE 29. An argument in favor of this is that our collaborators have also identified two other patients with late-onset leukoencephalopathy.

Patient 6's *AARS* variants are rare (mean allele frequencies <1%), located in conserved regions, predicted pathogenic, and not reported as disease-causing. PCR amplification and Sanger sequencing of variant 1 c.295G>A and variant 2 c.778A>G is seen in Figure 4.1A. c.295G>A is seen in both the patient and the mother, while c.778A>G is only seen in the patient (Figure 4.1A). Cloning and cDNA amplification of each allele shows that the variants are in *trans* (Figure 4.1B). Figure 4.1B shows that the mother's allele 2 has variant 1 whereas variant 2 is not seen in the parental alleles and is therefore *de novo* (Figure 4.1B). False paternity is not suspected but could not entirely be ruled out because we do not have WES data on the father.

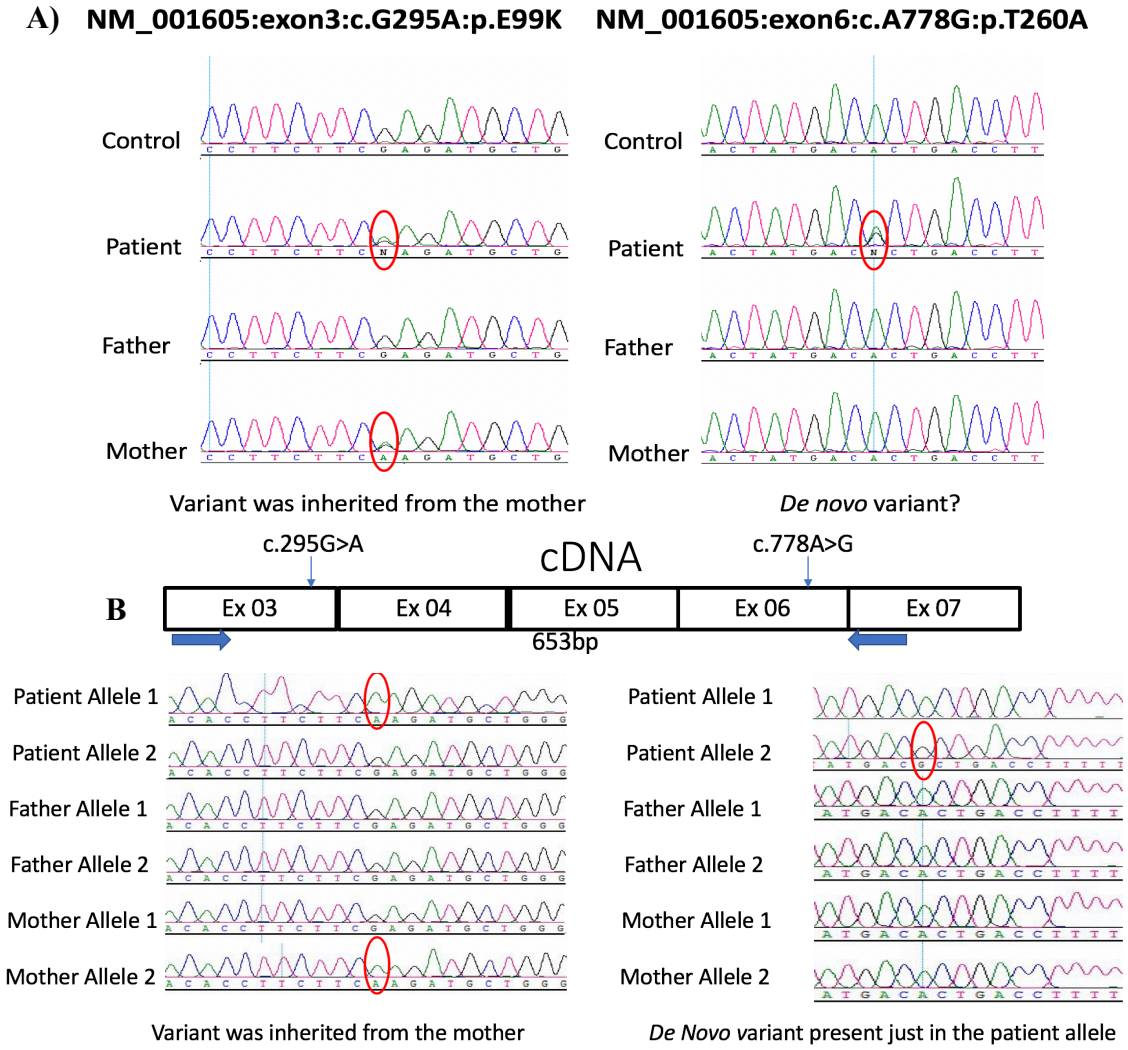


Figure 4.1: Sanger Sequencing of *AARS* genomic sequence and segregation analysis in family members. (A) Patient is heterozygous for a maternally-inherited variant, c.295G>A (p.E99K) located in exon 3 and heterozygous for a potential *de novo* variant, c.778A>G (p.T260A), located in exon 6. (B) cDNA Sanger sequencing confirms that the c.295G>A was inherited from the mother and that the c.778A>G is a *de novo* variant.

4.4 *AARS* mRNA Levels are Increased in our Patient's fibroblasts

Following the identification and validation of these novel *AARS* variants in the patient, the functional consequences for these genomic changes were investigated. AARS is not part of the MSC complex and does not have any known non-canonical functions in humans [106]. Therefore, experiments were designed to assess the impact

of variants on its canonical functions, i.e., aminoacylation (done by collaborators). mRNA transcript levels of AARS was assessed first. RNA from n=3 biological replicates was harvested from patient and control fibroblasts. RNA was DNase I treated to remove any possible DNA contamination, and then converted to cDNA using reverse transcriptase. cDNA at the appropriate dilution (based on standard curve) was analyzed by qPCR using primers specific to *AARS* mRNA as well as *HPRT1* housekeeping genes. AARS mRNA levels were normalized to the average of housekeeping genes. We observe a 2.5-fold increase in the mRNA levels of *AARS* in the patient's fibroblasts compared to the average of the two control fibroblasts we used (Figure 4.2).

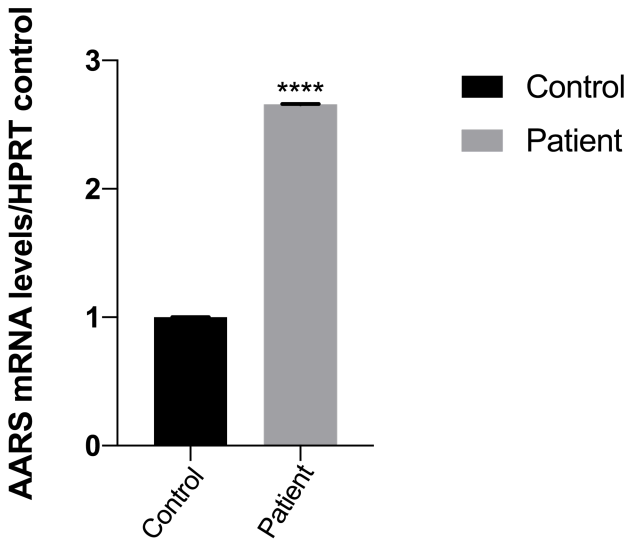


Figure 4.2 AARS shows increased mRNA levels in patient fibroblasts.
RT-qPCR analysis of *AARS* mRNA, from patient and age-sex match controls (n=3). Results are representative of AARS fold change after normalizing to mRNA levels of *HPRT* mRNA. Data values are shown as mean ± standard error of mean (*** denotes $p < 0.0001$).

4.5 AARS protein shows decreased expression

Based on the data we obtained from qPCR analysis of *AARS* mRNA levels we predicted that the variants might lead to improper folding of the protein which in turn might lead to its degradation. Therefore, we predicted that AARS protein levels would be decreased in our patient's fibroblasts. Thus, western blot analysis using protein extracts from patient's fibroblasts as well as two age- and sex-matched control fibroblasts (±1year) was performed. This was done on at least n=3 biological replicates and Beta-tubulin was used as a loading control. Our Western blot shows that there is a decrease in the AARS protein levels in our patient compared to controls (Figure 4.3A). We next quantified the band intensity using ImageJ software from the three biological replicates. Figure 4.3B shows the quantification obtained. A significant reduction in AARS protein levels was seen in patient compared to controls (Student *t*-test, $p<0.05$).

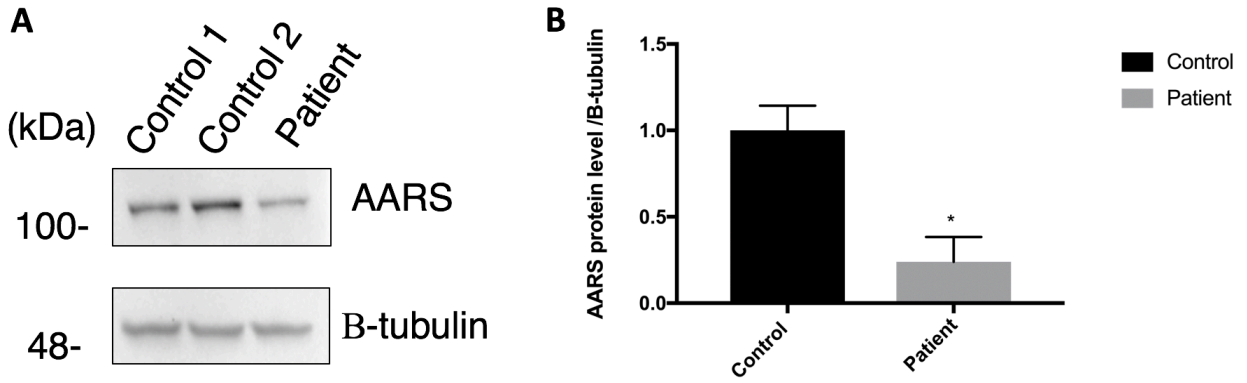


Figure 4.3: AARS protein levels are decreased in patient compared to controls. (A) Immunoblot analysis of patient and age-sex matched control fibroblasts. Lysates were blotted with anti-AARS and anti-β-tubulin antibodies. Molecular mass markers are shown on the left in kilodaltons. (B) Western blot bands from n=3 intendent biological replicates were quantified using ImageJ software, normalized over control (average of both controls) and represented as mean ± standard error of mean (Student *t*-test, *p<0.05).

4.6 AARS protein displays altered pattern of expression in the cytoplasm

AARS normally shows a homogenous distribution in the cytoplasm [41]. Therefore, we sought to investigate whether the variants seen in our patient could lead to improper localization of AARS. We used immunocytochemistry to determine the localization of AARS in patient and control fibroblasts. A homogenous distribution of AARS in the cytoplasm of control fibroblasts was observed, whereas patient cells showed an altered pattern of expression with abnormal puncta staining indicated with white arrows (Figure 4.4).

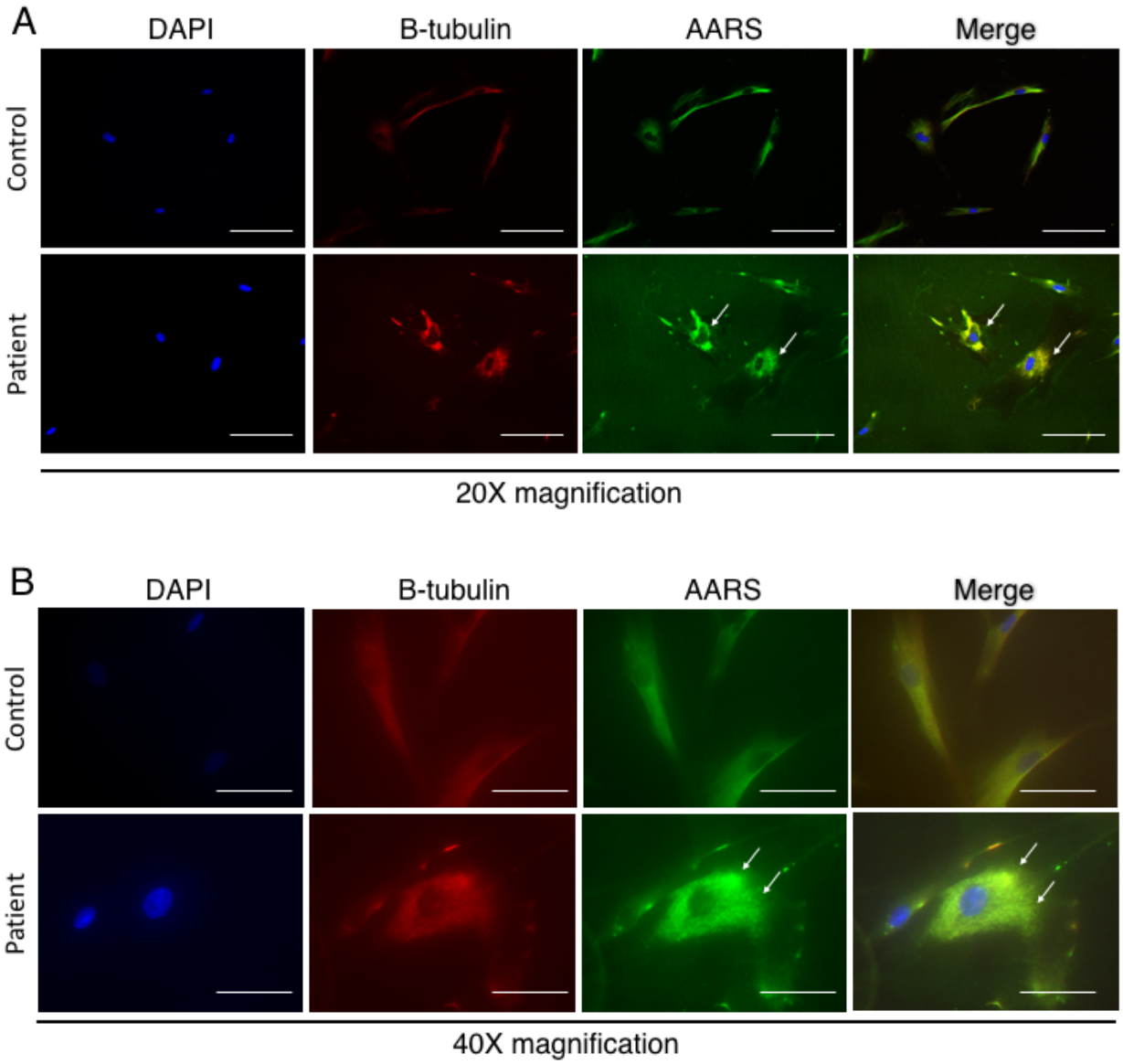


Figure 4.4: AARS shows an altered pattern of expression in patient fibroblasts compared to control.
Evaluation of endogenous AARS expression in control and patient fibroblasts with p.E99K and p.T260A at 20X (A) and 40X (B). Fibroblasts were labeled with anti-AARS and anti-ß-tubulin. DAPI was used to stain the nucleus. Abnormal punctate stain seen in patient fibroblast in indicated by white arrow. Scale bar = 150um.

Chapter 5: Discussion and Conclusion

5.1 General Discussion

5.1.1 WES Analysis for the Molecular Characterization of Novel Forms of Mitochondrial Leukoencephalopathies

NGS has revolutionized and accelerated the medical diagnostic process as it allows for the rapid interrogation of patient genomes via either WES, WGS, or both. In this thesis, we used WES for a cohort of patients with suspected mitochondrial leukoencephalopathy. Out of the 11 patients in this cohort, one was solved, 4 patients had strong candidates identified, while 6 patients remain unsolved.

For patient 1, we found variants in the *MBOAT7* gene*,* a membrane-bound O-acyltransferase family member 7 with novel compound heterozygous (CHetz) variants c.181_193del (p.Phe61ProfsTer12) and c.519C>A (p.Tyr246Ter) as strong candidates. *MBOAT7,* encodes lysophosphatidylinositol acyltransferase 1 (LPIAT1), which is the only member of its family to primarily transfer arachidonic acid from arachidonoyl-CoA to lysophosphatidylinositol (LPI) [168]. LPI is a bioactive lipid that is an important second messenger for metabolic and glucose signaling, necessary for differentiation, cell growth and motility [169, 170]. *MBOAT7* has been associated with seizures and intellectual disability (ID) [170-172]. Because our patient's presentation is different from what has been previously described, we hypothesized that our patient's phenotype represents an expansion to the *MBOAT7*-related disease spectrum.

Patient 2 was initially assessed with having a KSS-like MRI and phenotype. KSS is a rare sporadic disorder caused by large *de novo* mtDNA deletions [71]. After reviewing our patient's raw files (BAM), a large mtDNA deletion was identified (m.7950_15705). The

breakpoints m.7950 to 15705 were verified by LR-PCR and Sanger Sequencing. This breakpoint m.7950-15705 or 7,752bp deletion, has not been previously reported on MitoBreak (http://mitobreak.portugene.com/). This case is a good example of how powerful MRI pattern recognition is for diagnosing leukodystrophies and some genetically determined leukoencephalopathies. Further functional studies could be performed to confirm the pathogenicity of the deletion found. Unfortunately, no tissue is available to perform such studies in this patient.

Patient 3 was found to have a strong candidate in *WD repeat domain 62* (*WDR62)* with the following CHetz variants: c.2470C>T (p.Pro824Ser) and c.4487T>A (p.Leu1496Gln). Pathogenic variants in *WDR62* cause microcephaly, pachygyria, cortical thickening, lissencephaly and schizencephaly [173]. Recent investigations into WDR62 function revealed that decreased expression of WDR62 in glial cells is a major factor contributing to a microcephaly phenotype [174]. The depletion of WDR62 in addition to the cofactor Aurora kinase A (AURKA) was reported to worsen the phenotype [174]. Our patient presented with microcephaly along with other facial features which could probably be due to both variants being missense, compared to the more damaging INDEL alterations. Further testing into WDR62 expression or AURKA would be helpful in identifying functional consequences of patient variants.

Patient 4 was identified with the disease-causing variant c.677G>A; p.Arg226His in *PDHA1*, a gene located on the X chromosome. The first described enzymatic defects were described by Shin *et al.* in 2017 [166]. *PDHA1* encodes pyruvate dehydrogenase E1 alpha 1 subunit, part of the mitochondrial multi-protein pyruvate dehydrogenase complex (PDC), which functions to convert pyruvate into acetyl-CoA [166]. Our patient

presented with a slight lactic acid imbalance, which is compatible with the lactic acidosis seen in patients with PDC deficiency [166].

Finally, we report novel disease related variants in the *AARS* gene in patient 6. Diseases have been associated with mutations in the majority of both cytoplasmic and mitochondrial *ARS* genes (Table 1.2). There are 9/16 mitochondrial *ARS* and 6/14 cytoplasmic *ARS* related to diseases associated with a genetically determined leukoencephalopathy. Patient 6 was suspected to have an *AARS2* related disease based on MRI pattern, however, no variants were found in this gene. Instead, we found variants in the *AARS* gene: c.295G>A (p.E99K) and c.778A>G (p.T260A). These variants were found to be in highly conserved residues and were not previously reported to be disease causing (Table 4.1). Although the patient's phenotype did not match the AD Charcot Marie Tooth axonal type 2N or the AR Early Infantile Epileptic Encephalopathy-29 phenotypes, research has shown overlapping MRI track patterns between cytoplasmic and mitochondrial *ARS.* For example, pathogenic variants were found in *DARS,* which was not originally associated to disease, and these variants were identified in patients with a similar MRI pattern as *DARS2*-related leukodystrophy [105, 116]. As such, we propose that our patient's variants are pathogenic, they present a novel *AARS* phenotype, and expand the spectrum of AR *AARS*-disorders. Strengthening our initial hypothesis, our collaborators identified two other late-onset AR *AARS* cases (Table 4.4) [175].

WES is an ideal choice for novel molecular identifications when analyzing an etiologically heterogeneous group of patients compared to panel-based approaches or WGS. Panel-based sequencing is more effective than WES when identifying potential novel variants in a few select set of genes but is limited to those selected genes [176]. In

the study by Sun *et al.,* pitting WES against WGS for identifying variants in nine patients with heterogeneous neurological disorders, using a 500 gene panel (500GP), WES was able to correctly identify all variants that were found with WGS, with a deeper set average coverage, and negligible false positive variant calling [147]. WES has a high quality of variant calling, i.e. ~99% coverage and is clearly more cost-efficient [147]. While WES does not cover the entire genome, i.e. only the exons and intron-exon boundaries, it does allow for a more efficient analysis due to the smaller number of variants identified, i.e. 20,000 variants compared to the ~4 million variants identified with WGS [147, 177, 178].

The rate of identification between WES and WGS is also very similar, as both maintain about a 25% to 50% identification upon first analysis [178-180]. As such the 45% identification rate with our cohort is in-line with current discovery rates. It will be important to reanalyze our unsolved patients, as it has been shown that diagnostic rates increase with reanalysis after 6-12 months [181]. This increase in identification rate is predominately due to new clinical/genetic evidence, improvements in bioinformatic pipelines as well as new published disease-causing genes [179, 181]. For WES and WGS, a 7-21% increase in identification is shown following reanalysis [178, 179]. Whereas in the case of reanalysis of patients with intellectual disability, predominately caused by *de novo* variants, WGS has shown a 42% identification rate in previously undiagnosed cases by WES [178]. One of the limitations of WES is the inability to identify large deletions, duplications, chromosomal rearrangements, and long repeat extensions that could be caught by WGS. As such those in the cohort not identified by a reanalysis of WES should undergo WGS to further bolster potential identifications of pathogenic variants.

In our cohort, we did not have patients with potentially pathogenic variants in the same genes. This can be explained by the fact that our cohort was clinically heterogeneous and the fact that multiple genes encode proteins important for mitochondrial homeostasis. As we are now almost 10 years after the initial use of NGS for novel gene identification in Mendelian diseases, most of the causal genes for "relatively common" rare diseases have been identified and we are now dealing with ultra-rare diseases. It is therefore often difficult to identify the second, third, etc. patients to have a large enough cohort to have a strong argument regarding the pathogenesis of the identified variants. This highlights the importance of using online tools such as GeneMatcher to identify patients with similar phenotypes and variants in the same genes. This is what we have done for the patient with variants in *MBOAT7.* It allowed us to initiate a collaboration with another group on the disease spectrum of *MBOAT7*-related disorders.

The ACMG guidelines, together with the recent publication of Sherloc, have been essential in helping clinical and research laboratories analyzing NGS data [145]. In order to look at all the potential candidates in our cohort of suspected mitochondrial leukoencephalopathies, in addition to using them for the curation of nDNA variants, we used tools such as: MitoMap, MitoWheel, and MitoBreak for the analysis of mtDNA variants (Table 4.4). Of the 11 patients, only one (patient 3) had a mtDNA variant as a candidate: m.10144A>G, in the *MT-ND3* gene, encoding for a core component of CI in the ETC. The variant has not been reported in MitoMap and the patient phenotype, including the MRI pattern were compatible with a deficit in complex I. The variant was confirmed to be heteroplasmic in the patient, an indication of being potentially pathogenic, but it is a non-conserved location making it a less suitable candidate. Unfortunately, no

cells (e.g. fibroblasts) or samples (e.g. urine) were available to us to confirm or infer the complex I deficiency. Therefore, we cannot conclude that the variant in *MT-ND3* is pathogenic in this patient.

5.1.2 Functional Testing of Pathogenic Variants in the AARS Gene Causing a Novel Phenotype

Our patient had 2 novel variants in the *AARS* gene. First, we validated the variants, and assessed for segregation. Variant c.295G>A was inherited maternally, but c.778A>G appeared to be *de novo*, as it did not show in the paternal gDNA sequencing (Figure 4.1A). Sanger sequencing of ≥10 clones/individual after cDNA and cloning confirmed that both variants were *in trans* in the patient, with c.295G>A being maternally inherited, and c.778A>G being *de novo*.

There is an increasing number of neurological diseases associated with pathogenic variants in genes encoding cytoplasmic and mitochondrial tRNA synthetases. Due to the variety of functions the tRNA synthetases carry out, different potential mechanisms could underlie disease pathogenesis, such as (1) insufficient ARS protein production, (2) for ARS part of the MSC complex, abnormal assembly of the complex, (3) abnormal canonical functions, i.e. aminoacylation and (4) abnormal non-canonical functions. In the case of AARS-related disorders (CMT2N and EIEE-29, AD and AR, respectively), it has been shown that reduced aminoacylation activity accounts for at least part of the disease pathogenesis [106, 109, 182, 183]. It remains unclear why a given variant leads to CMT2N vs. EIEE-29. One hypothesis is that the AD CMT2N variants lead to a gain of function and the AR EIEE-29 variants are hypomorphic, yet, this is not the case. Indeed, all the AD AARS variants, with the exception of the variant E337K, have

been shown to be hypomorphic (Figure 5.1). It is therefore clear that other pathophysiological mechanisms are involved.

Currently, there are 23 disease causing variants published for *AARS* diseases, excluding the variants found in our patient and our collaborator's patient (Table 5.1) [14-16, 38, 106-108, 115]. 10 out of the 23 variants have been functionally tested (Figure 5.1), and out of the 5 implicated for AR-*AARS,* only p.F958S has not been assessed. A patient reported with a truncating and a missense *AARS* variants was shown to have a significant decrease in protein levels (<20%) compared to controls and the heterozygous mother of the patient also showed decreased protein levels (by approximately 50%) [182]. This illustrates that one functioning allele is sufficient for normal function, which is not surprising as this is what we typically observe for AR disorders caused by mutations in housekeeping genes.

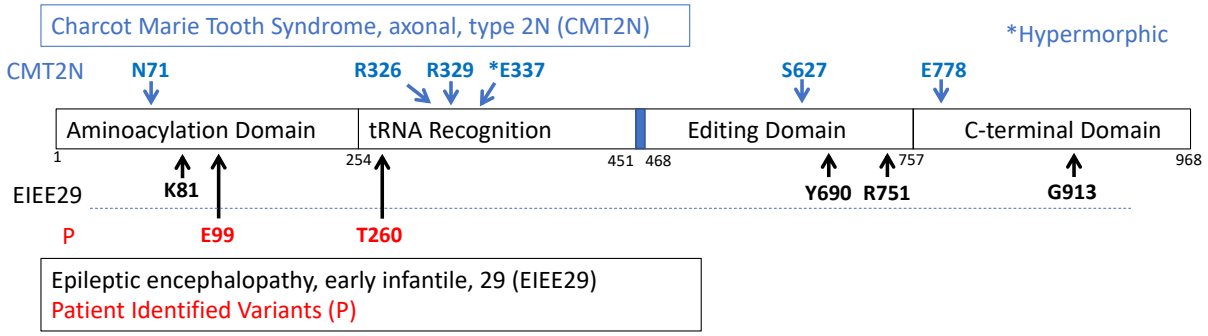


Figure 5.1 Locations of functionally validated *AARS* pathogenic variants. AD Charcot Marie Tooth Syndrome, axonal, type 2N (CMT2N) (blue). AR epileptic encephalopathy, early infantile, 29 (EIEE29) (black). All the variants are hypomorphic with the exception of that variant at position p.E337 which is hypermorphic. Our patient's variants identified at positions p.E99 and p.T260 are located within the aminoacylation and tRNA recognition domains, respectively (red).

Table 5.1: Known pathogenic variants in *AARS* and functional testing performed.

Chromosomal Position	Coding	Protein	OMIM	References	Functional Testing
AD AARS					
chr16:70277087	c.212A>G	p.N71S	Charcot-Marie-Tooth disease, type 2/Motor neuropathy	Dohrn (2017) J Neurochem 143: 507 PubMed: 28902413	None
chr16:70277088	c.211A>T	p.N71Y	Charcot-Marie-Tooth disease, type 2	Lin (2011) PLoS One 6: e29393 PubMed: 22206013 ; McLaughlin (2012) Hum Mutat 33: 244 PubMed: 22009580	AminoA, Editing, Yeast complementation, Localization
chr16:70276995	c.304G>C	p.G102R	Myeloneuropathy, dominant	Motley (2015) Neurology 84: 2040 PubMed: 25904691	None
chr16:70276971	c.328T>C	p.F110L	Charcot-Marie-Tooth disease, type 2	Karakaya (2018) Hum Mutat 39: 1284 PubMed: 29858556	None
chr16:70271949	c.503C>T	p.P168L	Peripheral neuropathy	Laššuthová (2016) Orphanet J Rare Dis 11: 118 PubMed: 27549087	None/VUS
chr16:70271927	c.525C>G	p.F175L	Charcot-Marie-Tooth disease, type 2	Bacquet (2018) BMJ Open 8: e021632 PubMed: 30373780	None
chr16:70302269	c.976C>T	p.R326W	Charcot-Marie-Tooth disease, type 2	Westerman (2018) Hum Mol Genet 27:23 PubMed: 30124830	Yeast complementation
chr16:70268356	c.986G>A	p.R329H	Charcot-Marie-Tooth disease, axonal	Latour (2010) Am J Hum Genet 86: 77 PubMed: 20045102; McLaughlin (2012) Hum Mutat 33: 244 PubMed: 22009580	AminoA, Editing, Yeast complementation, Localization
chr16:70302236	c.1009G>A	p.E337K	Charcot-Marie-Tooth disease, type 2	Westerman (2018) Hum Mol Genet 27:23 PubMed: 30124830	AminoA, Yeast complementation
chr16:70268323	c.1019A>G	p.N340S	Myopathy, distal	Sevy (2016) J Neurol Neurosurg Psychiatry 87: 340 PubMed: 25783436	None/VUS
chr16:70265040	c.1410C>G	p.I470M	Sensory-motor polyneuropathy	Wang (2016) Neurology 86: 1762 PubMed: 27164712	None/VUS
chr16:70259149	c.1823C>T	p.T608M	Peripheral neuropathy	Schabhüttl (2014) J Neurol 261: 970 PubMed: 24627108	None
chr16:70292995	c.1880C>T	p.S627L	Charcot-Marie-Tooth disease, type 2	Westerman (2018) Hum Mol Genet 27:23 PubMed: 30124830	Yeast complementation
chr16:70258168	c.2042G>T	p.G681V	Charcot-Marie-Tooth disease, type 2	Liu (2015) PLoS One 10: e0133636 PubMed: 26274329	None/VUS
chr16:70258147	c.2063A>G	p.E688G	Neuropathy	Bansagi (2015) J Neurol 262: 1899 PubMed: 26032230	None
chr16:70255829	c.2185C>T	p.R729W	Charcot-Marie-Tooth disease, type 2	Lupo (2016) J Mol Diagn 18: 225 PubMed: 26752306	None
chr16:70254688	c.2333A>C	p.E778A	Charcot-Marie-Tooth disease, type 2	McLaughlin (2012) Hum Mutat 33: 244 PubMed: 22009580; Datt (2014) BMC Genomics 15: 1063 PubMed: 25476837	AminoA, Editing, Localization
chr16:70253312	c.2677G>A	p.D893N	Hereditary motor neuropathy	Zhao (2012) Neurology 78: 1644 PubMed:22573628; Datt (2014) BMC Genomics 15: 1063 PubMed: 25476837	None
AR AARS					
chr16:70277057	c.242A>C	p.K81T	Epileptic encephalopathy with persistent myelination defect	Simons (2015) Am J Hum Genet 96: 675 PubMed: 25817015	AminoA, Editing, yeast complementation
chr16:70255763	c.2251A>G	p.R751G	Epileptic encephalopathy with persistent myelination defect	Simons (2015) Am J Hum Genet 96: 675 PubMed: 25817015	AminoA, Editing, yeast complementation
chr16:70252890	c.2738G>A	p.G913D	Microcephaly, hypomyelination and epileptic encephalopathy	Nakayama (2017) Hum Mutat 38: 1348 PubMed: 28493438	Westren, AminoA, Editing
chr16:70258142-70258143	c.2067dupC		Microcephaly, hypomyelination and epileptic encephalopathy	Nakayama (2017) Hum Mutat 38: 1348 PubMed: 28493438	Westren, AminoA, Editing
chr16:70252755	c.2873T>C	p.F958S	Epileptic encephalopathy, early infantile, 29	Karakaya (2018) Hum Mutat 39: 1284 PubMed: 29858556	None

Abbreviations/explanation: AminoA, aminoacylation. Editing, ability to edit incorrect amino acid from an incorrectly loaded t-RNA. Yeast, complementation assay utilizing yeast and the AARS homolog. VUS, Variant of uncertain signification due to lack of functional studies. AD, Autosomal dominant. AR, Autosomal recessive.

To characterize the potential pathogenic defects in our patient, we decided to assess mRNA and protein levels as well as the localization of AARS within the cells. Our collaborators (Dr. Gajja Salomons, Amsterdam) assessed the aminoacylation activity in the patient's fibroblasts, compared to age and sex-matched healthy control.

We started our study by assessing the mRNA levels of *AARS*. We observed a significant increase in the mRNA levels of AARS in our patient's fibroblasts. This was expected as our patient did not carry truncating variants and hence the transcript will not be targeted for non-sense mediated decay. Moreover, the observed increase in mRNA could be due to a feedback loop in which reduced proteins levels (as seen in the Western blot analysis) could trigger an activation in gene transcription. We have also demonstrated a 70% decrease in protein levels of AARS in our patient's fibroblasts compared to age- and sex-matched control. While unexpected, since our patient does not carry truncating mutations, it has been previously reported that in EPRS, non-truncating variants have led to diminished protein levels [112]. Moreover, the variants could lead to defects in protein folding, for example, which might lead to protein degradation.

Since our Western blot analysis showed significantly decreased AARS levels, it is plausible to speculate that AARS protein is targeted for degradation. Therefore, cycloheximide (CHX) experiments could be performed to assess the degradation of AARS protein in our patient. CHX interferes with translation, and therefore can be used to evaluate the half-life of AARS proteins, which has not been investigated previously. Furthermore, by treating the control and patient fibroblasts with MG132 and chloroquine, inhibitors of the proteasome and lysosome respectively, we could further identify through which pathway AARS is being degraded. These experiments would give insight into how our patient's missense variants have led to decreased protein.

Few studies have focused on investigating the localization of ARS proteins within the cell. Pathogenic variants in *ARS* genes for tyrosine (*YARS*) and glycine (*GARS*) have been shown to result in YARS and GARS mislocalization, respectively [91, 103, 114]. GARS is normally present in cytoplasmic granules, however, pathogenic variants in the gene have been shown to lead to a diffuse distribution [114]. Similarly, YARS is found to be localized in growth cones, yet this localization was reduced with *YARS* pathogenic variants [103]. AARS is known to be diffusely expressed in the cytoplasm, which is consistent with our results obtained in the control fibroblasts [184]. One of the few localization studies on human *AARS* was performed with plasmids expressing the p.N71Y variant, located within β-sheets surrounding the AARS catalytic core. Neuroblastoma cells expressing the p.N71Y variant, which causes CMT2N, showed abnormal puncta throughout the cytoplasm compared to fluorescently tagged WT AARS [110]. Interestingly, we observe the same localization defect in our patient fibroblasts (Figure 4.4). Our variant p.E99K is situated within one of the parallel β-sheets next to that of p.N71 and due to structural proximity to the p.N71, we hypothesized that our variant could also be mislocalized. Indeed, in our patient's fibroblasts, AARS showed an abnormal puncta staining indicative of potential protein aggregate formation along with a non-homogenous distribution in the cells (Figure 4.4). We hypothesize that our variant might lead to conformational changes in the protein structure that would lead to its aggregation. The lack of model organisms along with the limited human studies on pathogenic ARS variants make interpretation of pathophysiology of these variants rather difficult. We also speculate that future functional experiments might shed more light on pathogenicity when, for example, using cells that are more vulnerable to ARS mutations, i.e. neurons and glia. Therefore, using induced pluripotent stem cells from patient

fibroblasts to carry out such functional assays could allow us to investigate whether these variants might lead to more detrimental effects in these cell types.

Aminoacylation of AARS in our patient fibroblasts was observed at 16% of the regular activity of control AARS. The threshold of AcL activity below which disease occurs is not known. However, due to the observed dramatic reduction in aminoacylation we can confidently say that this accounts for part of the disease pathogenesis. Since our patient's fibroblasts display a significantly impaired aminoacylation activity, it is plausible to speculate that this would affect global protein synthesis and in turn lead to reduced protein translation, including that of AARS providing an explanation to the reduced AARS protein levels in our patient despite the non-truncating variants.

All cytoplasmic ARS enzymes have the ability to aminoacylate but only half can edit and seven have a dedicated editing domain [93]. AARS has a highly conserved editing domain, yet, two missense variants in this domain, p.S627L & p.R751G, do not alter editing [106, 109]. The frameshift-causing variant p.Tyr690Leufs*3 significantly reduces editing to less than 70%, which is similar to that of a known mouse model for p.A734E [182]. The p.A734E mouse model, or "sticky" mouse, named after appearance of rough, unkempt hair, leads to a two-fold increase in the ARS for serine mischarging tRNAs with alanine instead [185]. Another variant, p.C723A, as well as the p.A734E variant, showed an accumulation of cellular aggregates of abnormal proteins in cardiomyocytes and Purkinje cells [185, 186]. Both of our patient's variants are not in the editing domain, and do not harbor frameshift alterations that could impact the function of the editing domain, it is highly unlikely our patient has variants impacting editing.

As for our patient, with non-HLD, we predict that key proteins important for myelin homeostasis are not produced in sufficient amounts to maintain myelin and this leads to

progressive myelin breakdown. As it is seen in *AARS2*-related disease, specific tracts seem to be more vulnerable. As previously mentioned, we do not understand this tract-specificity seen in ARS and ARS2 and why cytoplasmic ARS lead to similar tract specificity compared to their mitochondrial counterparts.

5.2 Concluding Remarks

In this thesis, WES was used to identify the molecular etiology of well-phenotyped patients with suspected mitochondrial leukoencephalopathies. Five cases had candidate genes identified while no good candidate was identified for the remaining 6 patients. In one patient, we identified and validated two novel variants in the *AARS* gene. Moreover, we carried out functional testing to demonstrate that these variants lead to changes in *AARS* mRNA and protein levels along with changes in its cellular localization. These aspects represent potential disease-causing mechanisms that merit further investigations to exactly pinpoint molecular causes.

References

1. van der Knaap, M.S. and M. Bugiani, *Leukodystrophies: a proposed classification system based on pathological changes and pathogenetic mechanisms.* Acta Neuropathologica, 2017. **134**(3): p. 351-382.
2. Vanderver, A., et al., *Case definition and classification of leukodystrophies and leukoencephalopathies.* Mol Genet Metab, 2015. **114**(4): p. 494-500.
3. Vanderver, A., et al., *Leukodystrophy Overview*, in *GeneReviews*, R.A. Pagon, M. Adam, and H. Ardinger, Editors. 2014, University of Washington: Seattle (WA).
4. Schiffmann, R. and M.S. van der Knaap, *Invited article: an MRI-based approach to the diagnosis of white matter disorders.* Neurology, 2009. **72**(8): p. 750-759.
5. Steenweg, M.E., et al., *Magnetic resonance imaging pattern recognition in hypomyelinating disorders.* Brain, 2010. **133**(10): p. 2971-2982.
6. Ashrafi, M.R. and A.R. Tavasoli, *Childhood leukodystrophies: A literature review of updates on new definitions, classification, diagnostic approach and management.* Brain and Development, 2017. **39**(5): p. 369-385.
7. Bonkowsky, J.L., et al., *The burden of inherited leukodystrophies in children.* Neurology, 2010. **75**(8): p. 718-725.
8. Nahhas, N., et al., *Pelizaeus-Merzbacher-Like Disease 1*, in *GeneReviews*, M. Adam, H. Ardinger, and R. Pagon, Editors. 2019, University of Washington: Seattle, WA.
9. Wolf, N.I., et al., *Clinical spectrum of 4H leukodystrophy caused by POLR3A and POLR3B mutations.* Neurology, 2014. **83**(21): p. 1898-1905.
10. Bernard, G., et al., *Mutations of POLR3A Encoding a Catalytic Subunit of RNA Polymerase Pol III Cause a Recessive Hypomyelinating Leukodystrophy.* The American Journal of Human Genetics, 2011. **89**(3): p. 415-423.
11. Tétreault, M., et al., *Recessive Mutations in POLR3B, Encoding the Second Largest Subunit of Pol III, Cause a Rare Hypomyelinating Leukodystrophy.* The American Journal of Human Genetics, 2011. **89**(5): p. 652-655.
12. Dorboz, I., et al., *Mutation in POLR3K causes hypomyelinating leukodystrophy and abnormal ribosomal RNA regulation.* Neurol Genet, 2018. **4**(6): p. e289.
13. Thiffault, I., et al., *Recessive mutations in POLR1C cause a leukodystrophy by impairing biogenesis of RNA polymerase III.* Nat Commun, 2015. **6**: p. 7623-7631.
14. Mendes, M.I., et al., *Bi-allelic Mutations in EPRS, Encoding the Glutamyl-Prolyl-Aminoacyl-tRNA Synthetase, Cause a Hypomyelinating Leukodystrophy.* The American Journal of Human Genetics, 2018.
15. Wolf, N.I., et al., *Mutations in RARS cause hypomyelination.* Annals of Neurology, 2014. **76**(1): p. 134-139.
16. Taft, R.J., et al., *Mutations in DARS Cause Hypomyelination with Brain Stem and Spinal Cord Involvement and Leg Spasticity.* Am.J.Hum.Genet., 2013. **92**(5): p. 774-780.
17. Kevelam, S.H., et al., *Altered PLP1 splicing causes hypomyelination of early myelinating structures.* Ann Clin Transl Neurol, 2015. **2**(6): p. 648-61.

18. Simons, C., et al., *A De Novo Mutation in the beta-Tubulin Gene TUBB4A Results in the Leukoencephalopathy Hypomyelination with Atrophy of the Basal Ganglia and Cerebellum.* Am.J.Hum.Genet., 2013. **92**(5): p. 767-773.
19. Hamilton, E.M.C., et al., *UFM1 founder mutation in the Roma population causes recessive variant of H-ABC.* Neurology, 2017. **89**(17): p. 1821-1828.
20. Zara, F., et al., *Deficiency of hyccin, a newly identified membrane protein, causes hypomyelination and congenital cataract.* Nat.Genet., 2006. **38**(10): p. 1111-1113.
21. Miyake, N., et al., *X-linked hypomyelination with spondylometaphyseal dysplasia (H-SMD) associated with mutations in AIFM1.* Neurogenetics, 2017. **18**(4): p. 185-194.
22. Paznekas, W.A., et al., *Connexin 43 (GJA1) mutations cause the pleiotropic phenotype of oculodentodigital dysplasia.* Am J Hum.Genet., 2003. **72**(2): p. 408-418.
23. Hudson, L.D., et al., *Mutation of the proteolipid protein gene PLP in a human X chromosome-linked myelin disorder.* Proc Natl Acad Sci U S A, 1989. **86**(20): p. 8128-31.
24. Uhlenberg, B., et al., *Mutations in the gene encoding gap junction protein alpha 12 (connexin 46.6) cause Pelizaeus-Merzbacher-like disease.* Am.J.Hum.Genet., 2004. **75**(2): p. 251-260.
25. Verheijen, F.W., et al., *A new gene, encoding an anion transporter, is mutated in sialic acid storage diseases.* Nat Genet, 1999. **23**(4): p. 462-5.
26. Bondurand, N., et al., *A molecular analysis of the yemenite deaf-blind hypopigmentation syndrome: SOX10 dysfunction causes different neurocristopathies.* Hum Mol Genet, 1999. **8**(9): p. 1785-9.
27. Chelban, V., et al., *Mutations in NKX6-2 Cause Progressive Spastic Ataxia and Hypomyelination.* Am J Hum Genet, 2017. **100**(6): p. 969-977.
28. Simons, C., et al., *A recurrent de novo mutation in TMEM106B causes hypomyelinating leukodystrophy.* Brain, 2017. **140**(12): p. 3105-3111.
29. Gomez-Ospina, N., *Arylsulfatase A Deficiency*, in *GeneReviews*, R. Pagon, M. Adam, and H. Ardinger, Editors. 2006 [updated 14 Dec 2017], University of Washington: Seattle (WA).
30. Orsini, J., et al., *Krabbe Disease*, in *GeneReviews*, R. Pagon, M. Adam, and H. Ardinger, Editors. 2000 [updated 11 October 2018], University of Washington: Seattle, WA.
31. Raymond, G., A. Moser, and A. Fatemi, *X-Linked Adrenoleukodystrophy*, in *GeneReviews*, R. Pagon, M. Adam, and H. Ardinger, Editors. 1999 [updated 15 february 2018], University of Washington: Seattle, WA.
32. van der Knaap, M.S., et al., *Mutations in each of the five subunits of translation initiation factor eIF2B can cause leukoencephalopathy with vanishing white matter.* Ann Neurol, 2002. **51**(2): p. 264-70.
33. Leegwater, P.A., et al., *Subunits of the translation initiation factor eIF2B are mutant in leukoencephalopathy with vanishing white matter.* Nat Genet, 2001. **29**(4): p. 383-8.
34. Pestova, T.V., et al., *Molecular mechanisms of translation initiation in eukaryotes.* Proc Natl Acad Sci U S A, 2001. **98**(13): p. 7029-36.

35. Kevelam, S.H., et al., *Update on Leukodystrophies: A Historical Perspective and Adapted Definition.* Neuropediatrics, 2016. **47**(6): p. 349-354.
36. Goldstein, A. and M.J. Falk, *Mitochondrial DNA Deletion Syndromes*, in *GeneReviews*, M. Adam, H. Ardinger, and R. Pagon, Editors. 2019, University of Washington: Seattle, WA.
37. Bugiani, M., et al., *Clinical and molecular findings in children with complex I deficiency.* Biochim Biophys Acta, 2004. **1659**(2-3): p. 136-47.
38. Steenweg, M.E., et al., *Leukoencephalopathy with thalamus and brainstem involvement and high lactate 'LTBL' caused by EARS2 mutations.* Brain, 2012. **135**(Pt 5): p. 1387-94.
39. Scheper, G.C., et al., *Mitochondrial aspartyl-tRNA synthetase deficiency causes leukoencephalopathy with brain stem and spinal cord involvement and lactate elevation.* Nat.Genet., 2007. **39**(4): p. 534-539.
40. van der Bliek, A.M., M.M. Sedensky, and P.G. Morgan, *Cell Biology of the Mitochondrion.* Genetics, 2017. **207**(3): p. 843-871.
41. Robin, E.D. and R. Wong, *Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells.* J Cell Physiol, 1988. **136**(3): p. 507-13.
42. Gilkerson, R., et al., *The mitochondrial nucleoid: integrating mitochondrial DNA into cellular homeostasis.* Cold Spring Harb Perspect Biol, 2013. **5**(5): p. a011080.
43. Wong, L.J., *Diagnostic challenges of mitochondrial DNA disorders.* Mitochondrion, 2007. **7**(1-2): p. 45-52.
44. Ng, Y.S. and D.M. Turnbull, *Mitochondrial disease: genetics and management.* J Neurol, 2016. **263**(1): p. 179-91.
45. Calvo, S.E., K.R. Clauser, and V.K. Mootha, *MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins.* Nucleic Acids Res, 2016. **44**(D1): p. D1251-7.
46. Boczonadi, V. and R. Horvath, *Mitochondria: impaired mitochondrial translation in human disease.* Int J Biochem Cell Biol, 2014. **48**: p. 77-84.
47. Okamoto, K., et al., *The protein import motor of mitochon- dria- a targeted molecular ratchet driving unfolding and translocation.* The European Molecular Biology Organism 2002. **21**(14): p. 3659-3671.
48. Westermann, B., *Mitochondrial fusion and fission in cell life and death.* Nat Rev Mol Cell Biol, 2010. **11**(12): p. 872-84.
49. Cipolat, S., et al., *OPA1 requires mitofusin 1 to promote mitochondrial fusion.* Proc Natl Acad Sci U S A, 2004. **101**(45): p. 15927-32.
50. Yoon, Y., et al., *The Mitochondrial Protein hFis1 Regulates Mitochondrial Fission in Mammalian Cells through an Interaction with the Dynamin-Like Protein DLP1.* Molecular and Cellular Biology, 2003. **23**(15): p. 5409-5420.
51. Otera, H., et al., *Mff is an essential factor for mitochondrial recruitment of Drp1 during mitochondrial fission in mammalian cells.* J Cell Biol, 2010. **191**(6): p. 1141-58.
52. Griparic, L., T. Kanazawa, and A.M. van der Bliek, *Regulation of the mitochondrial dynamin-like protein Opa1 by proteolytic cleavage.* The Journal of Cell Biology, 2007. **178**(5): p. 757-764.

53. Nakamura, N., et al., *MARCH-V is a novel mitofusin 2- and Drp1-binding protein able to change mitochondrial morphology.* EMBO reports, 2006. **7**(10): p. 1019-1022.
54. Braschi, E., R. Zunino, and H.M. McBride, *MAPL is a new mitochondrial SUMO E3 ligase that regulates mitochondrial fission.* EMBO Rep, 2009. **10**(7): p. 748-54.
55. Han, X.J., et al., *CaM kinase I alpha-induced phosphorylation of Drp1 regulates mitochondrial morphology.* J Cell Biol, 2008. **182**(3): p. 573-85.
56. Taguchi, N., et al., *Mitotic phosphorylation of dynamin-related GTPase Drp1 participates in mitochondrial fission.* J Biol Chem, 2007. **282**(15): p. 11521-9.
57. Cereghetti, G.M., et al., *Dephosphorylation by calcineurin regulates translocation of Drp1 to mitochondria.* Proc Natl Acad Sci U S A, 2008. **105**(41): p. 15803-8.
58. Gilkerson, R., et al., *The Mitochondrial Nucleoid: Integrating Mitochondrial DNA into Cellular Homeostasis.* Cold Spring Harbor Perspectives in Biology, 2013. **5**(5): p. a011080.
59. DiMauro, S. and E.A. Schon, *Mitochondrial Respiratory-Chain Diseases.* New England Journal of Medicine, 2003. **348**(26): p. 2656-2668.
60. Mimaki, M., et al., *Understanding mitochondrial complex I assembly in health and disease.* Biochim Biophys Acta, 2012. **1817**(6): p. 851-62.
61. Fleury, C., B. Mignotte, and J.L. Vayssiere, *Mitochondrial reactive oxygen species in cell death signaling.* Biochimie, 2002. **84**(2-3): p. 131-41.
62. Rizzuto, R. and T. Pozzan, *Microdomains of intracellular Ca2+: molecular determinants and functional consequences.* Physiol Rev, 2006. **86**(1): p. 369-408.
63. Diodato, D., D. Ghezzi, and V. Tiranti, *The Mitochondrial Aminoacyl tRNA Synthetases: Genes and Syndromes.* Int J Cell Biol, 2014. **2014**: p. 787956.
64. Khan, N.A., Govindaraj, P., Meena, A. K., & Thangaraj, K., *Mitochondrial disorders: Challenges in diagnosis & treatment.* Indian Journal Of Medical Research, 2015. **141**(1): p. 13-26.
65. Perrier, S., et al., *Recessive mutations in NDUFA2 cause mitochondrial leukoencephalopathy.* Clin Genet, 2018. **93**(2): p. 396-400.
66. Alves, C., et al., *Neuroimaging of Mitochondrial Cytopathies.* Top Magn Reson Imaging, 2018. **27**(4): p. 219-240.
67. Schapira, A.H.V., *Mitochondrial diseases.* The Lancet, 2012. **379**(9828): p. 1825-1834.
68. Bannwarth, S., et al., *Prevalence of rare mitochondrial DNA mutations in mitochondrial disorders.* J Med Genet, 2013. **50**(10): p. 704-14.
69. Jha, M.K. and B.M. Morrison, *Glia-neuron energy metabolism in health and diseases: New insights into the role of nervous system metabolic transporters.* Exp Neurol, 2018. **309**: p. 23-31.
70. Alston, C.L., et al., *The genetics and pathology of mitochondrial disease.* J Pathol, 2017. **241**(2): p. 236-250.
71. Zeviani, M., et al., *Deletions of mitochondrial DNA in Kearns-Sayre syndrome.* Neurology, 1988. **38**(9): p. 1339-1346.
72. Alexander, C., et al., *OPA1, encoding a dynamin-related GTPase, is mutated in autosomal dominant optic atrophy linked to chromosome 3q28.* 2000.

73. Roesch, K., et al., *Human deafness dystonia syndrome is caused by a defect in assembly of the DDP1/TIMM8a–TIMM13 complex.* Human Molecular Genetics, 2002. **11**(5): p. 477-486.
74. Balaban, R.S., S. Nemoto, and T. Finkel, *Mitochondria, Oxidants, and Aging.* Cell, 2005. **120**(4): p. 483-495.
75. Haas, R.H., et al., *The in-depth evaluation of suspected mitochondrial disease.* Mol Genet Metab, 2008. **94**(1): p. 16-37.
76. Chaliotis, A., et al., *The complex evolutionary history of aminoacyl-tRNA synthetases.* Nucleic Acids Res, 2017. **45**(3): p. 1059-1068.
77. Antonellis, A. and E.D. Green, *The role of aminoacyl-tRNA synthetases in genetic diseases.* Annu Rev Genomics Hum Genet, 2008. **9**: p. 87-107.
78. Schimmel, P., *Aminoacyl tRNA synthetases: General scheme of structure-function relationships in the polypeptides and recognition of transfer RNAs.* Annual Review of Biochemistry, 1987. **56**: p. 125-158.
79. Moras, D., *Structural and functional relationships between aminoacyl-tRNA synthetases.* Trends Biochem Sci, 1992. **17**(4): p. 159-64.
80. Schimmel, P., *Development of tRNA synthetases and connection to genetic code and disease.* Protein Sci, 2008. **17**(10): p. 1643-52.
81. Eriani, G., et al., *Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs.* Nature, 1990. **347**(6289): p. 203-6.
82. Ling, J., N. Reynolds, and M. Ibba, *Aminoacyl-tRNA Synthesis and Translational Quality Control.* Annual Review of Microbiology, 2009. **63**(1): p. 61-78.
83. Perona J.J. and G.-S. I., *Synthetic and Editing Mechanisms of Aminoacyl-tRNA Synthetases. In: Kim S. (eds) Aminoacyl-tRNA Synthetases in Biology and Medicine. Topics in Current Chemistry, vol 344. Springer, Dordrecht.* 2013.
84. Elded, E. and P. Schimmel, *Rapid deacylation by isoleucyl transfer ribonucleic acid synthetase of isoleucine-specific transfer ribonucleic acid aminoacylated with valine.* Journal of Biological Chemistry, 1972. **247**: p. 2961-2964.
85. Fersht, A.R., *Editing mechanisms in protein synthesis. Rejection of valine by the isoleucyl-tRNA synthetase.* Biochemistry, 1977. **16**(5): p. 1025-1030.
86. Dias, J., et al., *Small-angle X-ray solution scattering study of the multi-aminoacyl-tRNA synthetase complex reveals an elongated and multi-armed particle.* J Biol Chem, 2013. **288**(33): p. 23979-89.
87. Kang, J., et al., *Heat shock protein 90 mediates protein-protein interactions between human aminoacyl-tRNA synthetases.* J Biol Chem, 2000. **275**(41): p. 31682-8.
88. Quevillon, S., et al., *The p43 Component of the Mammalian Multi-synthetase Complex Is Likely To Be the Precursor of the Endothelial Monocyte-activating Polypeptide II Cytokine.* The Journal of Biological Chemistry, 1997. **272**(51): p. 32573-32579.
89. Kim, J.Y., et al., *p38 is essential for the assembly and stability of macromolecular tRNA synthetase complex: implications for its physiological significance.* Proc Natl Acad Sci U S A, 2002. **99**(12): p. 7912-6.
90. Huntsman, R.J., et al., *Neuroimaging and neurophysiology studies in carriers of cree leukoencephalopathy.* Can.J.Neurol.Sci., 2011. **38**(2): p. 347-348.
91. Fuchs, S.A., et al., *Aminoacyl-tRNA synthetase deficiencies in search of common themes.* Genet Med, 2018.

92. Kyriacou, S.V. and M.P. Deutscher, *An important role for the multienzyme aminoacyl-tRNA synthetase complex in mammalian translation and cell growth.* Mol Cell, 2008. **29**(4): p. 419-27.
93. Storkebaum, E., *Peripheral neuropathy via mutant tRNA synthetases: Inhibition of protein translation provides a possible explanation.* Bioessays, 2016. **38**(9): p. 818-29.
94. Smirnova, E.V., et al., *Noncanonical functions of aminoacyl-tRNA synthetases.* Biochemistry (Mosc), 2012. **77**(1): p. 15-25.
95. Pang, Y.L., K. Poruri, and S.A. Martinis, *tRNA synthetase: tRNA aminoacylation and beyond.* Wiley Interdiscip Rev RNA, 2014. **5**(4): p. 461-80.
96. Guo, M., P. Schimmel, and X.L. Yang, *Functional expansion of human tRNA synthetases achieved by structural inventions.* FEBS Lett, 2010. **584**(2): p. 434-42.
97. Yao, P. and P.L. Fox, *Aminoacyl-tRNA synthetases in medicine and disease.* EMBO Mol Med, 2013. **5**(3): p. 332-43.
98. Martyn, C. and R. Hughes, *Epidemiology of peripheral neuropathy.* Journal of Neurosurgery and Psychiatry, 1997. **62**(4): p. 310-318.
99. Latour, P., et al., *A major determinant for binding and aminoacylation of tRNA(Ala) in cytoplasmic Alanyl-tRNA synthetase is mutated in dominant axonal Charcot-Marie-Tooth disease.* Am J Hum Genet, 2010. **86**(1): p. 77-82.
100. McLaughlin, H.M., et al., *Compound heterozygosity for loss-of-function lysyl-tRNA synthetase mutations in a patient with peripheral neuropathy.* Am J Hum Genet, 2010. **87**(4): p. 560-6.
101. Antonellis, A., et al., *Glycyl tRNA synthetase mutations in Charcot-Marie-Tooth disease type 2D and distal spinal muscular atrophy type V.* Am J Hum Genet, 2003. **72**(5): p. 1293-9.
102. Safka Brozkova, D., et al., *Loss of function mutations in HARS cause a spectrum of inherited peripheral neuropathies.* Brain, 2015. **138**(Pt 8): p. 2161-72.
103. Jordanova, A., et al., *Disrupted function and axonal distribution of mutant tyrosyl-tRNA synthetase in dominant intermediate Charcot-Marie-Tooth neuropathy.* Nat Genet, 2006. **38**(2): p. 197-202.
104. Edvardson, S., et al., *Deleterious mutation in the mitochondrial arginyl-transfer RNA synthetase gene is associated with pontocerebellar hypoplasia.* Am J Hum Genet, 2007. **81**(4): p. 857-62.
105. Scheper, G.C., et al., *Mitochondrial aspartyl-tRNA synthetase deficiency causes leukoencephalopathy with brain stem and spinal cord involvement and lactate elevation.* Nat Genet, 2007. **39**(4): p. 534-9.
106. Simons, C., et al., *Loss-of-function alanyl-tRNA synthetase mutations cause an autosomal-recessive early-onset epileptic encephalopathy with persistent myelination defect.* Am J Hum Genet, 2015. **96**(4): p. 675-81.
107. Dallabona, C., et al., *Novel (ovario) leukodystrophy related to AARS2 mutations.* Neurology, 2014.
108. Miyake, N., et al., *A novel homozygous mutation of DARS2 may cause a severe LBSL variant.* Clin Genet, 2011. **80**(3): p. 293-6.
109. McLaughlin, H.M., et al., *A recurrent loss-of-function alanyl-tRNA synthetase (AARS) mutation in patients with Charcot-Marie-Tooth disease type 2N (CMT2N).* Hum Mutat, 2012. **33**(1): p. 244-53.

110. Tatsumi, Y., et al., *CMT type 2N disease-associated AARS mutant inhibits neurite growth that can be reversed by valproic acid.* Neurosci Res, 2018.
111. Antonellis, A., et al., *Functional Analyses of Glycyl-tRNA Synthetase Mutations Suggest a Key Role for tRNA-Charging Enzymes in Peripheral Axons.* Journal of Neuroscience, 2006. **26**(41): p. 10397-10406.
112. Mendes, M.I., et al., *Bi-allelic Mutations in EPRS , Encoding the Glutamyl-Prolyl-Aminoacyl-tRNA Synthetase, Cause a Hypomyelinating Leukodystrophy.* The American Journal of Human Genetics, 2018.
113. Niehues, S., et al., *Impaired protein translation in Drosophila models for Charcot-Marie-Tooth neuropathy caused by mutant tRNA synthetases.* Nat Commun, 2015. **6**: p. 7520.
114. Motley, W.W., et al., *Charcot-Marie-Tooth-linked mutant GARS is toxic to peripheral neurons independent of wild-type GARS levels.* PLoS Genet, 2011. **7**(12): p. e1002399.
115. Kuo, M.E., et al., *Cysteinyl-tRNA Synthetase Mutations Cause a Multi-System, Recessive Disease That Includes Microcephaly, Developmental Delay, and Brittle Hair and Nails.* The American Journal of Human Genetics, 2019. **104**(3): p. 520-529.
116. Taft, R.J., et al., *Mutations in DARS cause hypomyelination with brain stem and spinal cord involvement and leg spasticity.* Am J Hum Genet, 2013. **92**(5): p. 774-80.
117. McMillan, H.J., et al., *Compound heterozygous mutations in glycyl-tRNA synthetase are a proposed cause of systemic mitochondrial disease.* BMC Medical Genetics, 2014. **15**(1): p. 36.
118. Vester, A., et al., *A loss-of-function variant in the human histidyl-tRNA synthetase (HARS) gene is neurotoxic in vivo.* Hum Mutat, 2013. **34**(1): p. 191-9.
119. Puffenberger, E.G., et al., *Genetic mapping and exome sequencing identify variants associated with five novel diseases.* PLoS One, 2012. **7**(1): p. e28936.
120. Kopajtich, R., et al., *Biallelic IARS Mutations Cause Growth Retardation with Prenatal Onset, Intellectual Disability, Muscular Hypotonia, and Infantile Hepatopathy.* Am J Hum Genet, 2016. **99**(2): p. 414-22.
121. Santos-Cortez, R.L., et al., *Mutations in KARS, encoding lysyl-tRNA synthetase, cause autosomal-recessive nonsyndromic hearing impairment DFNB89.* Am J Hum Genet, 2013. **93**(1): p. 132-40.
122. Casey, J.P., et al., *Identification of a mutation in LARS as a novel cause of infantile hepatopathy.* Mol Genet Metab, 2012. **106**(3): p. 351-8.
123. Gonzalez, M., et al., *Exome sequencing identifies a significant variant in methionyl-tRNA synthetase (MARS) in a family with late-onset CMT2.* J Neurol Neurosurg Psychiatry, 2013. **84**(11): p. 1247-9.
124. Zhang, X., et al., *Mutations in QARS, encoding glutaminyl-tRNA synthetase, cause progressive microcephaly, cerebral-cerebellar atrophy, and intractable seizures.* Am J Hum Genet, 2014. **94**(4): p. 547-58.
125. Wolf, N.I., et al., *Mutations in RARS Cause Hypomyelination.* American Neurological Association, 2014. **76**(1): p. 134-139.
126. Karaca, E., et al., *Genes that Affect Brain Structure and Function Identified by Rare Variant Analyses of Mendelian Neurologic Disease.* Neuron, 2015. **88**(3): p. 499-513.

127. Nowaczyk, M.J.M., et al., *A novel multisystem disease associated with recessive mutations in the tyrosyl-tRNA synthetase (YARS) gene.* American Journal of Medical Genetics Part A, 2017. **173**(1): p. 126-134.
128. Gotz, A., et al., *Exome sequencing identifies mitochondrial alanyl-tRNA synthetase mutations in infantile mitochondrial cardiomyopathy.* Am J Hum Genet, 2011. **88**(5): p. 635-42.
129. Dallabona, C., et al., *Novel (ovario) leukodystrophy related to AARS2 mutations.* Neurology, 2014. **82**(23): p. 2063-2071.
130. Coughlin, C.R., 2nd, et al., *Mutations in the mitochondrial cysteinyl-tRNA synthase gene, CARS2, lead to a severe epileptic encephalopathy and complex movement disorder.* J Med Genet, 2015. **52**(8): p. 532-40.
131. Elo, J.M., et al., *Mitochondrial phenylalanyl-tRNA synthetase mutations underlie fatal infantile Alpers encephalopathy.* Hum Mol Genet, 2012. **21**(20): p. 4521-9.
132. Pierce, S.B., et al., *Mutations in mitochondrial histidyl tRNA synthetase HARS2 cause ovarian dysgenesis and sensorineural hearing loss of Perrault syndrome.* Proc Natl Acad Sci U S A, 2011. **108**(16): p. 6543-8.
133. Schwartzentruber, J., et al., *Mutation in the nuclear-encoded mitochondrial isoleucyl-tRNA synthetase IARS2 in patients with cataracts, growth hormone deficiency with short stature, partial sensorineural deafness, and peripheral neuropathy or with Leigh syndrome.* Hum Mutat, 2014. **35**(11): p. 1285-9.
134. Pierce, S.B., et al., *Mutations in LARS2, encoding mitochondrial leucyl-tRNA synthetase, lead to premature ovarian failure and hearing loss in Perrault syndrome.* Am J Hum Genet, 2013. **92**(4): p. 614-20.
135. Hadchouel, A., et al., *Biallelic Mutations of Methionyl-tRNA Synthetase Cause a Specific Type of Pulmonary Alveolar Proteinosis Prevalent on Reunion Island.* Am J Hum Genet, 2015. **96**(5): p. 826-31.
136. Bayat, V., et al., *Mutations in the mitochondrial methionyl-tRNA synthetase cause a neurodegenerative phenotype in flies and a recessive ataxia (ARSAL) in humans.* PLoS Biol, 2012. **10**(3): p. e1001288.
137. Vanlander, A.V., et al., *Two siblings with homozygous pathogenic splice-site variant in mitochondrial asparaginyl-tRNA synthetase (NARS2).* Hum Mutat, 2015. **36**(2): p. 222-31.
138. Sofou, K., et al., *Whole exome sequencing reveals mutations in NARS2 and PARS2, encoding the mitochondrial asparaginyl-tRNA synthetase and prolyl-tRNA synthetase, in patients with Alpers syndrome.* Mol Genet Genomic Med, 2015. **3**(1): p. 59-68.
139. Belostotsky, R., et al., *Mutations in the mitochondrial seryl-tRNA synthetase cause hyperuricemia, pulmonary hypertension, renal failure in infancy and alkalosis, HUPRA syndrome.* Am J Hum Genet, 2011. **88**(2): p. 193-200.
140. Diodato, D., et al., *VARS2 and TARS2 mutations in patients with mitochondrial encephalomyopathies.* Hum Mutat, 2014. **35**(8): p. 983-9.
141. Sasarman, F., et al., *A novel mutation in YARS2 causes myopathy with lactic acidosis and sideroblastic anemia.* Hum Mutat, 2012. **33**(8): p. 1201-6.
142. Parikh, S., et al., *A clinical approach to the diagnosis of patients with leukodystrophies and genetic leukoencephelopathies.* Mol Genet Metab, 2015. **114**(4): p. 501-515.

143. Heather, J.M. and B. Chain, *The sequence of sequencers: The history of sequencing DNA.* Genomics, 2016. **107**(1): p. 1-8.
144. Hamilton, A., et al., *Concordance between whole-exome sequencing and clinical Sanger sequencing: implications for patient care.* Molecular Genetics & Genomic Medicine, 2016. **4**(5): p. 504-512.
145. Nykamp, K., et al., *Sherloc: a comprehensive refinement of the ACMG-AMP variant classification criteria.* Genet Med, 2017. **19**(10): p. 1105-1117.
146. Ewans, L.J., et al., *Whole-exome sequencing reanalysis at 12 months boosts diagnosis and is cost-effective when applied early in Mendelian disorders.* Genet Med, 2018.
147. Sun, Y., et al., *Next-Generation Diagnostics: Gene Panel, Exome, or Whole Genome?* Human Mutation, 2015. **36**(6): p. 648-655.
148. Vanderver, A., et al., *Whole exome sequencing in patients with white matter abnormalities.* Ann Neurol, 2016. **79**(6): p. 1031-7.
149. Richards, 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, 2015. **17**(5): p. 405-24.
150. Vanderver, A., et al., *Whole Exome Sequencing in Patients with White Matter Abnormalities.* American Academy of Neurology, 2016. **79**(1): p. 1031-1037.
151. Tetreault, M., et al., *Recessive Mutations in POLR3B, Encoding the Second Largest Subunit of Pol III, Cause a Rare Hypomyelinating Leukodystrophy.* Am J Hum.Genet., 2011. **89**(5): p. 652-5.
152. Wong, L.J., *Mitochondrial syndromes with leukoencephalopathies.* Semin Neurol, 2012. **32**(1): p. 55-61.
153. Bell, C.J., et al., *Carrier testing for severe childhood recessive diseases by next-generation sequencing.* Sci Transl Med, 2011. **3**(65): p. 65ra4.
154. Quail, M.A., et al., *Optimal enzymes for amplifying sequencing libraries.* Nat Methods, 2011. **9**(1): p. 10-1.
155. DePristo, M.A., et al., *A framework for variation discovery and genotyping using next-generation DNA sequencing data.* Nature Genetics, 2011. **43**(5): p. 491-498.
156. McKenna, A., et al., *The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data.* Genome Res, 2010. **20**(9): p. 1297-303.
157. Li, H., et al., *The Sequence Alignment/Map format and SAMtools.* Bioinformatics, 2009. **25**(16): p. 2078-9.
158. Wang, K., M. Li, and H. Hakonarson, *ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data.* Nucleic Acids Res., 2010. **38**(16): p. e164.
159. Fu, W., et al., *Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants.* Nature, 2013. **493**(7431): p. 216-20.
160. Karczewski, K.J., et al., *Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes.* bioRxiv, 2019: p. 531210.
161. Lek, M., et al., *Analysis of protein-coding genetic variation in 60,706 humans.* Nature, 2016. **536**(7616): p. 285-91.

162. Alazami, A.M., et al., *Accelerating novel candidate gene discovery in neurogenetic disorders via whole-exome sequencing of prescreened multiplex consanguineous families.* Cell Rep, 2015. **10**(2): p. 148-61.
163. Sobreira, N., et al., *GeneMatcher: a matching tool for connecting investigators with an interest in the same gene.* Hum Mutat, 2015. **36**(10): p. 928-30.
164. Ovchinnikov, I.V., et al., *Whole Human Mitochondrial DNA Sequencing*, in *Forensic DNA Typing Protocols*, W. Goodwin, Editor. 2016, Springer New York: New York, NY. p. 157-171.
165. Jia, H., et al., *Long-range PCR in next-generation sequencing: comparison of six enzymes and evaluation on the MiSeq sequencer.* Sci Rep, 2014. **4**: p. 5737.
166. Shin, H.K., et al., *Enzymatic testing sensitivity, variability and practical diagnostic algorithm for pyruvate dehydrogenase complex (PDC) deficiency.* Mol Genet Metab, 2017. **122**(3): p. 61-66.
167. Wolf, N.I.M., PhD; Toro, Camilo MD; Kister, Ilya MD; Latif, Kartikasalwah Abd MD; Leventer, Richard MD; Pizzino, Amy MGC; Simons, Cas PhD; Abbink, Truus E.M. PhD; Taft, Ryan J. PhD; van der Knaap, Marjo S. MD, PhD; Vanderver, Adeline MD, *DARS-associated leukoencephalopathy can mimic a steroid-responsive neuroinflammatory disorder.* Neurology, 2015. **84**(3): p. 226-230.
168. Lee, H.-C., et al., *Caenorhabditis elegans mboa-7, a member of the MBOAT family, is required for selective incorporation of polyunsaturated fatty acids into phosphatidylinositol.* Molecular biology of the cell, 2008. **19**(3): p. 1174-1184.
169. Arifin, S.A. and M. Falasca, *Lysophosphatidylinositol Signalling and Metabolic Diseases.* Metabolites, 2016. **6**(1).
170. Johansen, A., et al., *Mutations in MBOAT7, Encoding Lysophosphatidylinositol Acyltransferase I, Lead to Intellectual Disability Accompanied by Epilepsy and Autistic Features.* Am J Hum Genet, 2016. **99**(4): p. 912-916.
171. Hu, H., et al., *Genetics of intellectual disability in consanguineous families.* Mol Psychiatry, 2018.
172. Santos-Cortez, R.L.P., et al., *Novel candidate genes and variants underlying autosomal recessive neurodevelopmental disorders with intellectual disability.* Hum Genet, 2018. **137**(9): p. 735-752.
173. Bilguvar, K., et al., *Whole-exome sequencing identifies recessive WDR62 mutations in severe brain malformations.* Nature, 2010. **467**(7312): p. 207-10.
174. Lim, N.R., et al., *Glial-Specific Functions of Microcephaly Protein WDR62 and Interaction with the Mitotic Kinase AURKA Are Essential for Drosophila Brain Growth.* Stem Cell Reports, 2017. **9**(1): p. 32-41.
175. Tang, Y., et al., *AARS2 leukoencephalopathy: A new variant of mitochondrial encephalomyopathy.* Mol Genet Genomic Med, 2019. **7**(4): p. e00582.
176. Mak, T.S.H., et al., *Coverage and diagnostic yield of Whole Exome Sequencing for the Evaluation of Cases with Dilated and Hypertrophic Cardiomyopathy.* Sci Rep, 2018. **8**(1): p. 10846.
177. de Ligt, J., et al., *Diagnostic exome sequencing in persons with severe intellectual disability.* N Engl J Med, 2012. **367**(20): p. 1921-9.
178. Gilissen, C., et al., *Genome sequencing identifies major causes of severe intellectual disability.* Nature, 2014. **511**(7509): p. 344-7.

179. Jalkh, N., et al., *The added value of WES reanalysis in the field of genetic diagnosis: lessons learned from 200 exomes in the Lebanese population.* BMC Med Genomics, 2019. **12**(1): p. 11.
180. Trujillano, D., et al., *Clinical exome sequencing: results from 2819 samples reflecting 1000 families.* Eur J Hum Genet, 2017. **25**(2): p. 176-182.
181. Ewans, L.J., et al., *Whole-exome sequencing reanalysis at 12 months boosts diagnosis and is cost-effective when applied early in Mendelian disorders.* Genet Med, 2018. **20**(12): p. 1564-1574.
182. Nakayama, T., et al., *Deficient activity of alanyl-tRNA synthetase underlies an autosomal recessive syndrome of progressive microcephaly, hypomyelination, and epileptic encephalopathy.* Hum Mutat, 2017. **38**(10): p. 1348-1354.
183. Weterman, M.A.J., et al., *Hypermorphic and hypomorphic AARS alleles in patients with CMT2N expand clinical and molecular heterogeneities.* Hum Mol Genet, 2018. **27**(23): p. 4036-4050.
184. Chihade, J.W., et al., *Origin of mitochondria in relation to evolutionary history of eukaryotic alanyl-tRNA synthetase.* Proc Natl Acad Sci U S A, 2000. **97**(22): p. 12153-12157.
185. Lee, J.W., et al., *Editing-defective tRNA synthetase causes protein misfolding and neurodegeneration.* Nature, 2006. **443**(7107): p. 50-5.
186. Liu, Y., et al., *Deficiencies in tRNA synthetase editing activity cause cardioproteinopathy.* Proc Natl Acad Sci U S A, 2014. **111**(49): p. 17570-5.