Defining the Genetic and Molecular Mechanisms of Disease-Associated Mutations in Aminoacyl-tRNA Synthetase Genes

by

Rebecca Meyer-Schuman

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Human Genetics)
in the University of Michigan
2021

Doctoral Committee:

Professor Anthony Antonellis, Chair
Professor David Ginsburg
Assistant Professor Kristin Koutmou
Professor John Moran
Professor Thomas Wilson

Rebecca Meyer Schuman

rebmeyer@umich.edu

ORCID iD: 0000-0002-2004-3505

Dedication

To all of the teachers and mentors who have believed in me, inspired me, and opened my eyes to the beauty of biology.

Acknowledgements

I am fortunate to have decided to pursue a doctoral degree with the Department of Human Genetics at the University of Michigan. Over the past six years with this community, I have grown in confidence, curiosity, and passion for learning, experimenting, and teaching. I have been empowered not only with an exceptional education, but with mentors and friends who will shape the contours of my life as a scientist in the years to come.

To Molly Martin, for making it easy to be a student. Thank you for calmly and efficiently untangling any logistical or bureaucratic headache I brought to your doorstep.

To Miriam Meisler, Jacob Kitzman, Guy Lenk, and all of the other wonderful members of our Thursday group meetings—thank you for sharing your wisdom, expertise, and enthusiasm with me. This community has been a bottomless well of support and scientific inspiration, and I am grateful to have had a place in it. To Miriam, Guy, and Young Park—thank you for your help and guidance in developing and characterizing the *Tars1* mouse lines. I will always be chasing the momentary rush of being convinced I see a phenotype, and dashing to find Guy or Young to tell them about it.

To my thesis committee members, Kristin Koutmou, David Ginsburg, John Moran, and Tom Wilson—thank you for the hours you have spent pushing me to think more critically about my science. Your support, encouragement, and guidance has been instrumental in helping me reach this moment. I am extremely grateful to Tom for his insightful suggestions on yeast plasmids and selection systems, which made Chapter 4 of this dissertation possible. Thank you Kristin, for sharing your wealth of knowledge on protein translation with me, and for lending me your expertise and resources to attempt exciting new experiments. I have learned so much from our collaboration, and from Tyler. As always, I am grateful to John and his generosity with his resources, his time, and his advice, as well as to John Moldovan for his help with co-

immunoprecipitations. Finally, I cannot thank David enough for the unwavering support he has shown me over the past six years, and his enthusiasm for my successes. He has helped me feel valued as a student and a scientist, and I will always be thankful for his mentorship.

To Stephanie Oprescu, Chetna Gopinath, and Liz Fogarty—thank you for training me, guiding me, and caring for me through the early years of graduate school. You set the bar for my colleagues and friends very high. To Molly Kuo, Jennifer Pierluissi, and Megan Forrest, who vaulted right over that bar—I am so lucky to have worked and learned alongside you. You have made me a better scientist and a better friend. To Sheila Marte, Christina del Greco, and Allison Cale—it has been a joy to work with you and to teach you. The Antonellis lab is fortunate to have you, and I can't wait to see what you do next.

To Tony. I will always, always be grateful that you saw potential in me. I have learned more from you than I can put into words. Thank you for everything you have done for me over the past six years. If I can become even half the scientist and mentor that you are, I will consider myself very lucky indeed.

To my parents, for their continuous love and support, and for always asking me what I was working on.

To Ari. I couldn't have done it without you.

Table of Contents

List of Tables

List of Figures

List of Appendices

Abstract

Protein synthesis is an essential process that allows the nucleic acid code to be translated in the cell. For protein translation to occur, the ribosome must be supplied with tRNA molecules charged with amino acids. The enzymes responsible for ligating tRNA and amino acids are the aminoacyl-tRNA synthetases (ARS), encoded by a family of 37 genes. Variants in all 37 ARS genes can cause dominant and/or recessive human genetic diseases. The recessive diseases are severe, multi-system disorders that primarily affect the central nervous system, as well as the muscles, lung, and liver. The dominant diseases are axonal peripheral neuropathies commonly classified as Charcot-Marie-Tooth disease. The full clinical and genetic spectrum for both dominant and recessive ARS-related disease has yet to be determined. Additionally, the mechanism of these diseases is poorly understood. This dissertation seeks to expand the locus, allelic, and phenotypic heterogeneity of ARS-mediated diseases, and to define the mechanism of ARS-mediated disease, through 1) characterizing newly identified patient alleles; 2) developing a pipeline of model organisms to predict novel ARS disease candidates and define their associated dominant and recessive phenotypes, and 3) testing pathogenic alanyl-tRNA synthetase (*AARS1*) alleles for a dominant-negative effect in a yeast model.

Here, we characterized 15 variants across glycyl-(*GARS1*), histidyl-(*HARS1*), methionyl-(*MARS1*), asparginyl-(*NARS1*), and threonyl-tRNA synthetase (*TARS1*) identified in patients with dominant peripheral neuropathy or multisystem recessive diseases. Through synthesizing genetic and functional evidence, we expanded the allelic spectrum of *GARS1*- and *HARS1*- mediated dominant neuropathy, and the allelic and phenotypic spectrum of *MARS1*- and *TARS1*-mediated recessive disease. We also identified *NARS1* as a candidate gene for dominant peripheral neuropathy. To complement these efforts, we developed a predictive pipeline using the defined phenotypes of pathogenic ARS alleles in yeast, *C. elegans*, and mouse. We used this pipeline to design deleterious mutations in *TARS1* and

assess them for a dominant peripheral neuropathy or multi-system recessive phenotypes. Through studies in yeast and worm, we identified a hypomorphic *TARS1* allele, R433H. When tested in mouse, *in trans* with a null *TARS1* allele, R433H causes a recessive phenotype of lung failure, growth restriction, and hair defects. This model will be an asset to determine how reduced *TARS1* function differentially impacts mammalian tissues, and can inform clinical efforts to identify and treat patients with bi-allelic *TARS1* mutations.

Finally, we directly tested the hypothesis that dominant ARS variants are dominant-negative alleles. We focused on two variants in alanyl-tRNA synthetase (*AARS1*) with strong genetic evidence for pathogenicity, R329H and G102R *AARS1*. These variants reduce gene function in a yeast complementation assay, indicating that they are loss-of-function alleles. However, they repress yeast growth when co-expressed with wild-type *AARS1*, indicating that they are also dominantly toxic. To determine if this dominant toxicity was due to dimerization with wild-type AARS1, we designed a dimer domain mutation that impaired dimerization, and placed it *in cis* with R329H and G102R. This double-mutant rescued yeast growth, showing that dimerization is required for toxicity and that R329H and G102R are dominant-negative alleles. We also assessed three additional *AARS1* variants, and found that they also are dominant-negative alleles in this assay.

In sum, this work significantly contributes to defining the known genetic and phenotypic spectrum of ARS-mediated diseases, to expanding the role of model organisms in identifying candidate pathogenic ARS variants, and to defining the mechanism of dominant ARS-mediated peripheral neuropathy.

Chapter 1
Aminoacyl-tRNA Synthetases in Genetic Disease

1.1 Protein Translation and Disease

A cell's ability to divide, function, and differentiate depends on the flow of information from the nucleotide code into proteins that perform tasks required to sustain life and confer cellular identity. Gene sequences are transcribed from deoxyribonucleic acid (DNA) into pre-messenger ribonucleic acid (pre-mRNA), which is then processed into messenger RNA (mRNA) with a 5' cap and a 3' polyA tail, and shuttled out of the nucleus to the ribosome.[1] When the ribosome has assembled on the mRNA, transfer RNAs (tRNAs) charged with amino acids arrive to pair with their cognate sequences in the mRNA and donate their amino acid to a growing polypeptide chain.[1] The peptide chain is then ultimately released from the ribosome and folds to become a protein.[1] This process depends on the availability of tRNAs that are charged with the appropriate amino acid. The ligation of tRNA and amino acid is catalyzed by a family of enzymes, the aminoacyl-tRNA synthetases (ARSs). ARS enzymes, and the genetic disorders caused by ARS mutations, are the subject of this dissertation.

Rebecca Meyer-Schuman is the sole contributor to this chapter. This chapter is adapted from "Emerging mechanisms of aminoacyl-tRNA synthetase mutations in recessive and dominant human disease," published in *Human Molecular Genetics* (Volume 26, Issue R2, pages R114-R127, October 1 2017, License number 5114320712101) and "Evidence for a dominant-negative mechanism in HARS1-mediated peripheral neuropathy," published in *The FEBS Journal* (Volume 288, Issue 1, pages 91-94, September 17 2020, License number 5114321014941). Part of Figure 1.1 is published in *The Journal of Clinical Investigation* (Volume 129, Issue 12, pages 5568-5583, December 2 2019, License number 1135456-2).

1.1.1 Key stages in protein translation

Eukaryotic protein translation is initiated by the GTP-bound form of eukaryotic initiation factor 2 (eIF2) bringing the methionine-charged initiator $tRNA^{Met}$ to the 40S ribosomal subunit.[2] This complex then binds the 5' cap of mRNA and, assisted by helicases that unwind the mRNA's secondary structure, scans in the 3' direction until it finds an AUG codon.[2] When the AUG is located, the eIF2-GTP is hydrolyzed to eIF2-GDP and is released for recycling.[3] Then, the 60S ribosomal subunit joins the initiation complex to form the 80S subunit, with the charged initiator $tRNA^{Met}$ already positioned in the peptidyl site (P site) of the ribosome.[4] To begin the process of elongation, the elongation factor eEF1A, bound to GTP, brings a tRNA molecule charged with an amino acid to the acceptor site (A site) of the ribosome.[5] The anticodon of the tRNA recognizes the corresponding mRNA codon, and eEF1A-GTP hydrolysis enables the tRNA to properly fit into the ribosome.[5] Then, the ribosome catalyzes the formation of a peptide bond between the new amino acid and the methionine in the P site.[5] This triggers the transfer of the deacylated $tRNA^{Met}$ to the exit site (E site), and the transfer of the new tRNA into the P site. The process repeats with a new charged tRNA molecule moving into the A site. As the next tRNA in the P site moves into the E site, the $tRNA^{Met}$ in the E site is released.[5] This process continues until the ribosome reaches one of three termination codons: UAA, UAG, or UGA. Then, termination factors bind to the ribosome and promote the hydrolysis of the peptidyl-tRNA. Ribosomes are then recycled for future iterations of protein translation.[5]

1.1.2 Translation regulation

Protein translation is a dynamic process that is responsive to environmental cues and changes in intracellular metabolites.[6] For example, the cell can sense limited amino acid availability, which triggers a repression of global protein translation and upregulation of amino acid biosynthetic pathways to restore cellular homeostasis. This process starts with the kinase GCN2, which has a tRNA binding domain derived from histidyl-tRNA synthetase that recognizes accumulating uncharged tRNA in the cell.[3] When GCN2 binds uncharged tRNA, this triggers the kinase domain of GCN2 to phosphorylate the α subunit of eIF2.[3] eIF2α phosphorylation prevents eIF2B from recycling the GDP bound to eIF2 into GTP.[3] This prevents eIF2 delivering charged initiator $tRNA^{Met}$ to the ribosome, inhibiting translation initiation. However, although the dearth of GTP-

bound eIF2 represses translation globally, it *increases* specific translation of ATF4 by allowing the ribosome to bypass an inhibitory upstream open reading frame (ORF) and translate the *ATF4* coding sequence.[7] ATF4 then localizes to the nucleus where it increases the transcription of genes that help the cell respond to stress,[3] including genes involved in amino acid biosynthesis or transport.[8]

1.1.3 Genetic defects in protein translation cause disease

Mutations in genes that encode components of the translation machinery can cause a variety of genetic diseases. Defects in ribosomal proteins cause Diamond Blackfan anemia, which manifests as anemia, skeletal abnormalities, short stature, and cardiac and genitourinary malformations.[9] Mutations can also arise in the factors that process ribosomal RNA, causing bone marrow failure, cirrhosis, microcephaly, and leukoencephalopathy.[9] Defects in translation initiation or elongation factors cause similar central nervous system defects; mutations in the eEF1 complex, which delivers charged tRNA to the ribosome, cause microcephaly, epilepsy, autism, and intellectual disability.[10] Defects in the eIF2B complex (which normally helps recycle eIF2 to re-initiate protein translation) causes Vanishing White Matter disease, a leukoencephalopathy that can include ovarian failure in women.[11] Protein translation diseases can also be caused by mutations in tRNAs, or the enzymes that modify them.[12] Defects in proteins that process cytoplasmic tRNA can cause intellectual disability, developmental delay, and microcephaly, while defects in proteins that process mitochondrial tRNA cause anemia, encephalopathy, cardiomyopathies, and lactic acidosis.[12] Overall, the phenotypes caused by mutations in the protein translation machinery indicate that the central nervous system, the liver, the heart, and the hematopoietic system are particularly sensitive to disrupted cytoplasmic or mitochondrial protein translation.

The subject of this dissertation is another component of the protein translation machinery, the aminoacyl-tRNA synthetase (ARS) enzymes. This chapter will discuss the basic biology of ARS enzymes, as well as the dominant and recessive diseases caused by defects in ARS genes. It will also summarize the outstanding questions in the field pertaining to disease heterogeneity and disease mechanism. This thesis aims to address these questions.

1.1.4 An introduction to aminoacyl-tRNA synthetases

In order for the genetic code to be faithfully executed, each tRNA that recognizes a mRNA codon must be charged with the appropriate amino acid. The aminoacyl-tRNA synthetase (ARS) enzymes are responsible for charging these tRNAs with the required amino acids. They perform this pairing in a two-step reaction: first, the ARS enzyme binds the amino acid and ATP, then hydrolyzes the ATP to form an aminoacyl-adenylate, releasing a pyrophosphate molecule.[13] Then, the ARS anticodon binding domain recognizes and binds the appropriate tRNA, and the amino acid is transferred to the 5' end of the tRNA, releasing the charged tRNA and AMP.[13]

ARS enzymes must perform this reaction for each of the 20 amino acids that comprise the building blocks of proteins. Additionally, this reaction must occur both in the cytoplasm and the mitochondria to meet the needs of protein translation in both compartments. To fulfill these requirements, there are 37 members of the ARS family, enough to service each amino acid in each cellular compartment. Each ARS can charge one of the twenty amino acids, with two major exceptions: 1) there are two ARS proteins, FARSA and FARSB, that come together in a multi-unit structure to charge tRNA with phenylalanine, and 2) a single cytoplasmic ARS enzyme, EPRS1, charges both glutamic acid and proline.[13] All ARS genes are encoded in the nuclear genome, but upon translation, 17 ARS proteins localize to the mitochondria and 18 localize to cytoplasm. Two ARS can function in both compartments: lysyl-tRNA synthetase (*KARS1*), which undergoes alternative splicing to include or exclude the mitochondrial targeting sequence (MTS)[14]; and glycyl-tRNA synthetase (*GARS1*), which encodes a single mRNA with an upstream translation start site that includes the MTS in the open reading frame, and a downstream translation start site that excludes the MTS.[15] There is no glutaminyl-tRNA synthetase for mitochondria; instead, EARS2 improperly recognizes $tRNA^{Gln}$, charging it with glutamic acid.[16] Then, the glutamic acid is trans-aminated by glutamyl-tRNA amidotransferase into glutamine.[16]

The nomenclature for the human ARS family dictates that the gene or enzyme name begin with the one letter code of the amino acid substrate followed by "ARS". This is then followed by a

"1" if the enzyme functions in the cytoplasm or a "2" if it functions in the mitochondria. For example, the gene encoding the cytoplasmic threonyl-tRNA synthetase is *TARS1*, and the gene encoding the mitochondrial threonyl-tRNA synthetase is *TARS2*.

Each of the 37 ARS enzymes can be grouped into one of two classes, distinguished by their approach to binding tRNA (Table 1.1). Class I enzymes utilize a Rossmann nucleotide binding fold, which binds the minor groove side of the tRNA acceptor stem and ligates the amino acid to the 2' hydroxyl group on the terminal tRNA adenosine.[17,18] Class II enzymes utilize a core of anti-parallel β sheets, which binds the major groove side of the tRNA acceptor stem and ligates the amino acid to the 3' hydroxyl group on the terminal adenosine.[17,18] Class II enzymes primarily function as multimers, whereas Class I enzymes are usually, but not exclusively, monomers[19] (Table 1.1).

Interestingly, nine of the cytoplasmic ARS are known to congregate in a larger multi-synthetase complex (MSC) (Table 1.1), along with three non-ARS scaffolding proteins (AIMPs) that stabilize the complex and promote tRNA binding.[20] Although the function of the MSC is poorly understood, association with the MSC is thought to be required for the tRNA charging function of its members.[21] One hypothesis proposes that the MSC functions as a docking system for its ARS members; while localized to the MSC, they perform canonical tRNA charging, but upon dissociation they pursue non-canonical roles in signaling pathways[22] (for example, LARS1 can act as a leucine sensor in the mTORC1 pathway[23]). The MSC may also coordinate other aspects of protein translation: one of the scaffolding proteins (AIMP3) has been shown to be critical for delivery of charged initiator $tRNA^{Met}$ to eIF2 for translation initiation.[24]

In addition to LARS1, several other cytoplasmic ARS enzymes contribute to cellular pathways unrelated to tRNA charging. Some of these non-canonical functions employ the full-length ARS protein, but others involve isoforms that arise from alternative splicing or proteolytic fragmentation.[25,26] There is substantial evidence that ARS proteins and their fragments play a role in immune responses. Anti-synthetase Syndrome, an autoimmune disease marked by interstitial lung disease, arthritis, myositis, fever, and decreased blood flow to fingers is

Table 1.1. ARS localization and structural groups.

Gene name	Gene Symbol	Enzyme Cellular Compartment	Enzyme Class[28]	Monomer or Dimer	Participant in the multisynthetase complex (MSC)[29]
alanyl-tRNA synthetase 1	*AARS1*	Cytoplasm	II	Dimer[30]	No
alanyl-tRNA synthetase 2	*AARS2*	Mitochondria	II	Dimer[31]	No
cysteinyl-tRNA synthetase 1	*CARS1*	Cytoplasm	I	Dimer[32,33]	No
cysteinyl-tRNA synthetase 2	*CARS2*	Mitochondria	I	ND; monomer in *E. coli*[34]	No
aspartyl-tRNA synthetase 1	*DARS1*	Cytoplasm	II	Dimer[35]	Yes
aspartyl-tRNA synthetase 2	*DARS2*	Mitochondria	II	Dimer[36]	No
glutamyl-prolyl-tRNA synthetase	*EPRS*	Cytoplasm	I/II[37]	Dimer[29]	Yes
glutamyl-tRNA synthetase 2	*EARS2*	Mitochondria	I	Monomer[38]	No
phenylalanyl-tRNA synthetase alpha subunit	*FARSA*	Cytoplasm	II	Tetramer[39]	No
phenylalanyl-tRNA synthetase beta subunit	*FARSB*	Cytoplasm	II	Tetramer[39]	No
phenylalanyl-tRNA synthetase 2	*FARS2*	Mitochondria	II	Monomer[38]	No
glycyl-tRNA synthetase 1	*GARS1*	Cytoplasm and mitochondria	II	Dimer[40]	No
histidyl-tRNA synthtetase 1	*HARS1*	Cytoplasm	II	Dimer[41]	No
histidyl-tRNA synthetase 2	*HARS2*	Mitochondria	II	Dimer[42]	No
isoleucyl-tRNA synthetase 1	*IARS1*	Cytoplasm	I	Monomer[29]	Yes
isoleucyl-tRNA synthetase 2	*IARS2*	Mitochondria	I	ND; monomer in *T. thermophiles*[43]	No
lysyl-tRNA synthetase 1	*KARS1*	Cytoplasm and mitochondria	II[44]	Dimer/Tetramer[45]	Yes
leucyl-tRNA synthetase 1	*LARS1*	Cytoplasm	I	Monomer[29]	Yes
leucyl-tRNA synthetase 2	*LARS2*	Mitochondria	I	ND; monomer in *T. thermophilus*[46]	No
methionyl-tRNA synthetase 1	*MARS1*	Cytoplasm	I	Monomer[47]	Yes

methionyl-tRNA synthetase 2	*MARS2*	Mitochondria	I	ND; monomer in *E. coli*[48]	No
asparaginyl-tRNA synthetase 1	*NARS1*	Cytoplasm	II	Dimer[49]	No
asparaginyl-tRNA synthetase 2	*NARS2*	Mitochondria	II	Dimer[50]	No
prolyl-tRNA synthetase 2	*PARS2*	Mitochondria	II	Dimer[51]	No
glutaminyl-tRNA synthetase 1	*QARS1*	Cytoplasm	I	Monomer[52]	Yes
arginyl-tRNA synthetase 1	*RARS1*	Cytoplasm	I	Monomer[53]	Yes
arginyl-tRNA synthetase 2	*RARS2*	Mitochondria	I	Monomer[38]	No
seryl-tRNA synthetase 1	*SARS1*	Cytoplasm	II	Dimer[54]	No
seryl-tRNA synthetase 2	*SARS2*	Mitochondria	II	Dimer[55]	No
threonyl-tRNA synthetase 1	*TARS1*	Cytoplasm	II	Dimer[56]	No
threonyl-tRNA synthetase 2	*TARS2*	Mitochondria	II	Dimer[57]	No
valyl-tRNA synthetase 1	*VARS1*	Cytoplasm	I	Monomer[58]	No
valyl-tRNA synthetase 2	*VARS2*	Mitochondria	I	ND; monomer in *T. Thermophilus*[59]	No
tryptophanyl-tRNA synthetase 1	*WARS1*	Cytoplasm	I	Dimer[60]	No
tryptophanyl-tRNA synthetase 2	*WARS2*	Mitochondria	I	ND; dimer in *B. stearothermophilus*[61]	No
tyrosyl-tRNA synthetase 1	*YARS1*	Cytoplasm	I	Dimer[62]	No
tyrosyl-tRNA synthetase 2	*YARS2*	Mitochondria	I	Dimer[36]	No

ND: “No Data”

characterized by the presence of autoantibodies against ARS enzymes.[27] These antibodies most commonly recognize HARS1, but have also been observed to recognize TARS1, AARS1, IARS1, GARS1, NARS1, FARS1, KARS1, QARS1, and YARS1.[27] Additionally, multiple ARS proteins respond to cytokine signaling. For instance, the cytokine IFNγ triggers the alternative splicing of WARS1 into a truncated "mini-WARS" protein that is secreted to the extracellular space and acts as an anti-angiogenic factor.[63,26] IFNγ also triggers the secretion of full-length WARS1 to the extracellular space, where it binds the macrophage receptor TLR4 and participates in innate immunity against infection.[64] Similarly, under apoptotic conditions[65] or upon treatment with the cytokine tumor necrosis factor alpha (TNF-alpha)[63] YARS1 is secreted into the extracellular space where it is cleaved into two fragments.[65] The N-terminal fragment, termed "mini-YARS," can bind to vascular endothelial cells and trigger angiogenic signaling pathways.[63] The C-terminal fragment has cytokine properties, and triggers the release of TNF-alpha from macrophages.[65,66] Finally, in myeloid cells, IFNγ triggers the release of EPRS from the multi-synthetase complex.[67] EPRS then participates in another multi-protein complex known as the GAIT complex, which binds to the 3' UTR of inflammatory mRNAs and suppresses their translation.[67] Overall, these studies implicate aminoacyl tRNA synthetase enzymes in a complex network of immune system regulation.

1.2 ARS mutations in disease

It is unsurprising that ARS enzymes are ubiquitously expressed and essential to cellular life, given the critical role that they perform in protein translation. Although complete loss of any ARS enzyme is incompatible with life, bi-allelic variants that severely impair ARS function can cause a spectrum of recessive disorders, many of which have phenotypic overlap with other translation-related diseases discussed above in Section 1.1.3. Additionally, mono-allelic variants in 5 of the 37 ARS genes have been implicated in a tissue-restricted phenotype, dominant axonal peripheral neuropathy. This section will provide an overview of both recessive and dominant ARS-mediated disease.

1.2.1 Recessive ARS-mediated disease genotypes

All recessive ARS-mediated disorders are, by definition, caused by bi-allelic mutations. Typically, patients with these diseases are homozygous for missense mutations, compound heterozygous for missense mutations, or compound heterozygous for one missense mutation and one null allele.[68] Homozygosity or compound heterozygosity for ARS null alleles is lethal due to the essential nature of ARS enzymes; this has been experimentally demonstrated with a gene trap insertion in *Gars1* that ablates *Gars1* mRNA expression and is homozygous lethal.[69] The genotypes identified in patients with recessive ARS-related disease strongly suggest a loss-of-function mechanism for disease pathogenesis.

1.2.2 Clinical heterogeneity in ARS-mediated recessive phenotypes

Recessive ARS-mediated disease comprises a wide range of clinical manifestations (Table 1.2). Pathogenic variants in ARS genes encoding a mitochondrial enzyme tend to cause phenotypes in tissues with a high metabolic demand, particularly in the central nervous system. Leukoencephalopathies,[70-76] myopathies,[77-81] and liver disease[82,83] are all common features of mitochondrial ARS disease phenotypes. Additionally, lactic acidosis,[75,77,84-88] epilepsy,[70,71,74,82,86,89,90] developmental delay and intellectual disability,[73,89,91-93] ovarian failure,[42,70,94-96] and sensorineural hearing loss[72,91,94,97-101] are frequently observed in patients with mitochondrial ARS mutations. Variants in some mitochondrial ARS can cause a range of diverse phenotypes, possibly related to the degree of enzyme impairment. For example, variants in *AARS2* can cause both fatal infantile cardiomyopathy and adult onset leukodystrophy, with or without ovarian failure (Table 1.2). Through structural modeling, Euro et al. found that patients diagnosed with cardiomyopathy had two alleles that were predicted to severely reduce function, whereas leukodystrophy patients had one allele predicted to severely reduce function and one predicted to only moderately reduce function.[31] Even in cases where individuals have identical genotypes, the severity of protein translation defects can determine the severity of the phenotype. In the case of siblings with identical pathogenic *RARS2* mutations, the sibling with the greater reduction in mitochondrial OXPHOS protein complex levels presented with lactic acidosis and neurological symptoms, whereas the sibling with a milder OXPHOS reduction had lactic acidosis but no neurological symptoms.[86]

Table 1.2. Phenotypes associated with ARS variants.

Gene	Locus	Location of Protein Function	Mode of Inheritance	Disease Phenotype(s)
AARS1	16q22	Cytoplasm	Autosomal Recessive	Early-onset epileptic encephalopathy with myelination defect[102]
				Microcephaly with hypomyelination, epileptic encephalopathy, and spasticity[103]
				Microcephaly, developmental delay, acute liver failure[104]
				Trichothiodystrophy including neurodevelopmental delay and microcephaly[105]
			Autosomal Dominant	Charcot-Marie-Tooth disease type 2N[106-111]
				Distal hereditary motor neuropathy[112]
AARS2	6p21.1	Mitochondria	Autosomal Recessive	Leukoencephalopathy with or without ovarian failure[70,94,95,113-120]
				Cardiomyopathy[78,121]
				Optic atrophy and retinopathy[122]
				Ataxia, vision loss, and cognitive impairment without leukodystropy[123]
				Ovarian failure with no reported neurological symptoms[124]
				Primary pulmonary hypoplasia without evidence of cardiomyopathy[125]
CARS1	11p15.4	Cytoplasm	Autosomal Recessive	Microcephaly, developmental delay, brittle hair and nails[126]
CARS2	13q34	Mitochondria	Autosomal Recessive	Epileptic encephalopathy[71,127]
				Progressive myoclonic epilepsy[90]
DARS1	2q21.3	Cytoplasm	Autosomal Recessive	Hypomyelination with brain stem and spinal cord involvement and leg spasticity[128-130]
DARS2	1q25.1	Mitochondria	Autosomal Recessive	Leukoencephalopathy with brain stem and spinal cord involvement and lactate elevation[75,76,131-137]
EPRS	1q41	Cytoplasm	Autosomal Recessive	Hypomyelinating leukodystrophy[138]

EARS2	16p12.2	Mitochondria	Autosomal Recessive	Leukoencephalopathy with thalamus and brainstem involvement and high lactate [76,139-141]
				Neonatal lactic acidosis, recurrent hypoglycemia, agenesis of corpus callosum[84]
				Multiple respiratory chain complex defects[121]
				Seizures, liver dysfunction (no thalamus or brain stem involvement)[142]
				Late onset leukoencephalopathy[143]
FARSA	19p13.13	Cytoplasm	Autosomal Recessive	Multisystem disease including growth delay, hypotonia, brain calcifications, liver dysfunction[144]
FARSB	2q36.1	Cytoplasm	Autosomal Recessive	Multisystem disease including growth restriction, failure to thrive, developmental delay, interstitial pulmonary disease, cirrhosis and portal hypertension, brain calcifications[145-147]
FARS2	6p25.1	Mitochondria	Autosomal Recessive	Early onset epileptic encephalopathy with lactic acidosis, liver dysfunction, developmental delay, and premature death[74,148-150]
				Juvenile onset epilepsy[83,151,152]
				Spastic paraplegia[153-155]
				Spastic tetraparesis, developmental delay, and 11onoclinic epilepsy[156]
GARS1	7p15	Mitochondria and Cytoplasm	Autosomal Recessive	Systemic mitochondrial disease[157,158]
				Cardiomyopathy[121]
				Severe multisystem disorder[159]
			Autosomal Dominant	Dominant peripheral neuropathy with or without sensory involvement (CMT disease type 2D, dSMA-v, dHMN)[160-171]
HARS1	5q31.3	Cytoplasm	Autosomal Recessive	Multisystem ataxic syndrome[172]
				Usher syndrome[173]
			Autosomal Dominant	Dominant peripheral neuropathy with or without sensory involvement (CMT, dHMN)[174-176]
HARS2	5q31.3	Mitochondria	Autosomal Recessive	Perrault syndrome[42,177-180]
IARS1	9q22.31	Cytoplasm	Autosomal Recessive	Growth restriction, neonatal cholestasis, muscular hypotonia, intellectual disability, infantile hepatopathy[181-184]
IARS2	1q41	Mitochondria	Autosomal Recessive	Cataracts, growth hormone deficiency, sensory neuropathy, sensorineural hearing loss, skeletal dysplasia syndrome; Leigh syndrome[97-99,185]

KARS1	16q23.1	Mitochondria and Cytoplasm	Autosomal Recessive	Neurological disorder including leukodystrophy, sensorineural hearing loss, microcephaly, cerebral calcifications, epilepsy, vision loss nystagmus[72,100,186-194]
				Recessive intermediate Charcot-Marie-Tooth disease type B, dysmorphic features, developmental delay, self-abusive behavior, vestibular Schwannoma[195]
				Hypertrophic cardiomyopathy and combined mitochondrial respiratory chain defect[81]
LARS1	5q32	Cytoplasm	Autosomal Recessive	Infantile liver failure, intrauterine growth restriction, neurodevelopmental delay, microcytic anemia, recurrent infections, hypotonia, hypoalbuminemia[196–201]
LARS2	3p21.31	Mitochondria	Autosomal Recessive	Perrault syndrome[96,202-209]
				Hydrops, lactic acidosis, sideroblastic anemia (HLASA)[85]
				HLASA with sensorineural hearing loss and developmental delay[101]
				Reversible myopathy, lactic acidosis, and developmental delay[101]
MARS1	12q13.3	Cytoplasm	Autosomal Recessive	Multisystem disease including interstitial lung and liver disease, pulmonary alveolar proteinosis, anemia, failure to thrive, developmental delay [210–215]
				Trichothiodystrophy including ataxia, dysmorphic features, follicular keratosis, ichthyosis, and intellectual disability (no reported lung or liver involvement)[105]
			Autosomal Dominant	Peripheral neuropathy (Charcot-Marie-Tooth disease type 2U)[216–219]
MARS2	2q33.1	Cytoplasm	Autosomal Recessive	Developmental delay, sensorineural hearing loss[91]
				Autosomal recessive spastic ataxia with leukoencephalopathy[220]
NARS1	18q21.31	Cytoplasm	Autosomal Recessive	Microcephaly, psychomotor developmental delay, seizures, dysmorphisms, ataxia, peripheral neuropathy[221,222]
			Autosomal Dominant	Severe global developmental delay, intellectual disabilities, dysmorphia, seizures, spasticity, peripheral neuropathy, ataxia[222]
NARS2	11q14.1	Mitochondria	Autosomal Recessive	Alpers syndrome[223,224]
				Developmental delay, intellectual disability, epilepsy, myopathy[89,92]
				Nonsyndromic deafness[50]
				Reversible COX deficiency[225]
				Lethal epileptic encephalopathy with global brain atrophy[226]
				Leigh syndrome[50,227]
PARS2	3p21.31	Mitochondria	Autosomal Recessive	Alpers syndrome[223]
				Infantile-onset developmental delay, epilepsy, hypotonia, ataxia[89,228,229]
				Infantile spasms, microcephaly, facial dysmorphy, cardiomyopathy and multiorgan failure[51]

QARS1	3p21.31	Cytoplasm and Mitochondria	Autosomal Recessive	Intractable seizures, progressive microcephaly, cerebral-cerebellar atrophy, hypomyelination, developmental delay[230–232]
				Severe linear growth retardation, poor weight gain, microcephaly, cutaneous syndactyly of the toes, intellectual disability, and characteristic facial features[233]
RARS1	5q34	Cytoplasm	Autosomal Recessive	Hypomyelination, motor delay, ataxia, spasticity, intellectual disability with a broad range of severity[234–237]
RARS2	6q16.1	Mitochondria	Autosomal Recessive	Pontocerebellar hypoplasia[238–244]
				Early onset epileptic encephalopathy[245,246]
				Lactic acidosis with or without neurological symptoms (microcephaly, seizures, developmental delay)[86]
				Dysmorphic features, epileptic spasms, optic atrophy, severe hypotonia[247]
				Intellectual disability[93]
SARS1	1p13.3	Cytoplasm	Autosomal Recessive	Intellectual disability, ataxia, microcephaly, speech impairment, aggressive behavior[248]
SARS2	19q13.2	Mitochondria	Autosomal Recessive	Hyperuricemia, pulmonary hypertension, renal failure, and alkalosis[249–251]
				Spastic paresis[252]
TARS1	5p13.3	Cytoplasm	Autosomal Recessive	Trichothiodystrophy including ichthyosis, recurrent respiratory infection, developmental delay, and follicular keratosis[253]
TARS2	1q21.2	Mitochondria	Autosomal Recessive	Hypotonia, severe developmental delay, epilepsy[254,255]
VARS1	6p21.33	Cytoplasm	Autosomal Recessive	Severe developmental delay, microcephaly, seizures, cerebellar atrophy[256–259]
VARS2	6p21.33	Mitochondria	Autosomal Recessive	Multisystem disorder including microcephaly, epilepsy, ataxia, lactic acidosis, failure to thrive, with or without cardiomyopathy or pulmonary hypertension[80,254,260-263]
				Multiple respiratory chain complex defects[121]

WARS1	14q32.2	Cytoplasm	Autosomal Dominant	Distal hereditary motor neuropathy[264,265]
WARS2	1p12	Mitochondria	Autosomal Recessive	Intellectual disability, ataxia, Parkinsonism, microcephaly, lactic acidosis, cerebral atrophy, developmental delay, intellectual disability, and/or epilepsy[73,87,248,266-270]
YARS1	1p35.1	Cytoplasm	Autosomal Recessive	Severe multisystem disorder including failure to thrive, sensorineural hearing loss, brain dysmyelination, nystagmus, liver disease, pulmonary disease, anemia, hypotonia, and/or developmental delay[271–274]
			Autosomal Dominant	Dominant intermediate Charcot-Marie-Tooth disease type C[275,276]
YARS2	12p11.21	Mitochondria	Autosomal Recessive	Myopathy, lactic acidosis, sideroblastic anemia, cardiomyopathy, respiratory insufficiency[77,79,88,277–281]
				Multiple respiratory chain complex defects[121]

A similar dynamic may explain the extreme clinical heterogeneity of *SARS2*-mediated disease, although the relationship between genotype and phenotype is less clear. Homozygosity or compound heterozygosity for missense mutations that decrease the levels of aminoacylated mitochondrial $tRNA^{Ser}_{AGY}$ cause hyperuricemia, pulmonary hypertension, infantile renal failure, and alkalosis (termed HUPRA syndrome).[249,251] However, homozygosity for a splice site mutation that significantly reduces SARS2 levels and decreases the levels of aminoacylated mitochondrial $tRNA^{Ser}_{AGY}$ causes childhood-onset spastic paresis, with no apparent kidney dysfunction, uricemia, or alkalosis.[252] A side-by-side comparison of HUPRA variants with the spastic paresis splice variant will be required to understand the relationship between genotype and phenotype, and determine whether there is differential SARS2 impairment and/or differential defects in mitochondrial protein translation.

Mutations in ARS genes encoding cytoplasmic enzymes also cause a spectrum of recessive disorders, which can affect a wider array of tissues but also typically includes a neurological component. The recessive neurological phenotypes associated with cytoplasmic ARSs include hypomyelination,[128,130,138,234] microcephaly,[103,126,186,230,231,248,256] seizures,[102,103,230,231,256] sensorineural hearing loss,[100,173,194,273] and developmental delay.[181,195,210,231,256,271] Mutations in cytoplasmic ARS genes also frequently affect the lung or liver (Table 1.2). Interestingly, mutations in some ARS cause a uniquely severe phenotype in a commonly affected tissue. For example, although mutations in *FARSA*,[144] *FARSB*,[145,147] *IARS1*,[181] *MARS1*,[210] and *YARS1*[272] all cause liver dysfunction as one component of a multisystem disease, mutations in *LARS1* cause a severe, acute form of infantile liver failure.[196–199] Similarly, pulmonary disease is particularly pronounced in individuals with bi-allelic *FARSB*[145–147] and *MARS1* mutations[210,213], including a *MARS1*-specific form of pulmonary alveolar proteinosis.[211]

This relationship between a particular ARS gene and its tissue-predominant pathology is poorly understood. It is possible that these tissues produce critical proteins that have a high requirement for a specific amino acid; for example, the liver may require certain proteins with a high leucine content. Here, if defects in LARS1 reduce the availability of charged $tRNA^{Leu}$, this may preferentially affect the synthesis of leucine-abundant proteins that are critical for proper liver function, such as proteins in the lipid biosynthesis.[196] (Consistent with this hypothesis, liver

biopsies from patients with *LARS1*-mediated liver disease show large fat deposits.[196]) In this model, the degree to which synthesis of these leucine-rich proteins would be impaired would depend on multiple factors: 1) the expression profile of the required $tRNA^{Leu}$, which may vary between tissues[282] and modify the tRNA charging defect; 2) the availability of leucine based on nutrient intake, amino acid transportation, and the requirement for that amino acid in other cellular metabolic pathways; and 3) the cumulative need for charged $tRNA^{Leu}$ across the other polysomes in the cell. Significant work is required to evaluate this hypothesis and determine whether it contributes to the phenotypic heterogeneity of ARS-mediated recessive disease.

1.2.3 Dominant ARS-mediated disease genotypes

Variants in five ARS genes have been confidently implicated in dominant human disease, which manifests as dominant axonal peripheral neuropathy (discussed in Section 1.2.4). These are: glycyl-(*GARS1*),[160] tyrosyl-(*YARS1*),[275] alanyl-(*AARS1*),[106] histidyl-(*HARS1*),[174] and tryptophanyl-tRNA synthetase (*WARS1*).[264] In addition, a small number of variants in methionyl-tRNA synthetase (*MARS1*) have been reported in patients with dominant peripheral neuropathy.[216–219] However, there is currently insufficient genetic evidence to conclude that mutations in *MARS1* cause dominant peripheral neuropathy, as none of the reported variants segregate with disease in a large, multi-generational family. In general, the allelic spectrum of dominant pathogenic variants comprises missense mutations, with the exception of two in-frame deletions, one in *GARS1*[170] and one in *YARS1*.[275]

It is striking that only five ARS genes have been confidently implicated in dominant peripheral neuropathy to date, whereas nearly all have been implicated in recessive disorders. This may indicate that there is something unique about these five ARS genes that makes neurons susceptible to mutations in them. Alternately, if variants in additional ARS genes continue to be identified in patients with dominant peripheral neuropathy, this will more convincingly point to a generalizable defect in tRNA charging. Identifying patients with dominant peripheral neuropathy harboring variants in other ARS genes (and characterizing these variants to define their role in disease) will provide additional support for the hypothesis that defects in tRNA charging are a common mechanism of disease. These goals are discussed further in Chapter 3.

It is worth noting that patients with recessive *GARS1-, YARS1-, HARS1-,* or *AARS1-* mediated disease are frequently compound heterozygous or homozygous for missense alleles, but no parents carrying these missense alleles have been reported to have a peripheral neuropathy. This may indicate a separation of function between dominant missense alleles and recessive missense alleles, or a differential effect on protein stability. However, this interpretation is complicated by the fact that most parents of children with recessive ARS-mediated disease are not carefully evaluated by a neurologist. Additionally, dominant ARS-mediated peripheral neuropathy can be mild, late-onset, or incompletely penetrant[163], which might make it difficult to detect a phenotype. The studies presented in Chapter 4 of this dissertation provide a platform to begin distinguishing recessive missense alleles from dominant missense alleles. These topics are also discussed further in Chapter 5.

1.2.4 Dominant ARS variants cause axonal peripheral neuropathies

Dominant ARS variants cause dominant axonal peripheral neuropathies, frequently classified as Charcot-Marie-Tooth (CMT) disease Type 2. CMT disease is a genetically and clinically diverse group of peripheral neuropathies that is estimated to affect between 1 in 1,200 and 1 in 2,500 individuals.[283,284] CMT disease is characterized by decreased sensory and/or motor nerve function in the distal extremities.[285] This leads to sensory loss and muscle atrophy, which often begins in the feet and peroneal musculature. Later, this atrophy may reach the calves and hands, and even later, the forearms.[285] One unique aspect of *GARS1*-mediated CMT disease is that it presents as an upper-limb predominant phenotype, beginning with weakness and muscle atrophy in the hands.[286] Only about half of affected individuals develop lower limb symptoms, which can vary in severity.[286] The reasons for this upper limb predominance are poorly understood. One consideration is that GARS1 is the only CMT-associated ARS enzyme that functions in both the cytoplasm and mitochondria[13]; it is possible that its role in mitochondrial protein translation contributes to the severity of the upper limb phenotype.

CMT is divided into two groups. CMT Type 1 is a demyelinating form of peripheral neuropathy; it is caused by a primary defect in the Schwann cells that form the myelin sheath around

peripheral neurons.[287] These defects lead to decreased nerve conduction velocities.[287] Mutations that cause CMT Type 1 are frequently found in genes involved in myelin production, such as *PMP22* or *MPZ*.[287] CMT Type 2 is caused by a primary defect in the axon, and is identified by decreased amplitudes of compound muscle action potentials or sensory nerve action potentials, rather than decreased velocities.[288] Mutations that cause CMT Type 2 can affect genes particularly important for neuronal function, such as ion channels and axonal transport factors, but can also affect housekeeping genes, such as *MFN2*.[287,289] *MFN2* is a ubiquitously expressed, essential gene that encodes a protein with a role in mitochondrial fusion.[290] CMT-associated alleles in *MFN2* have been shown to decrease mitochondrial fusion (classified as loss-of-function alleles) or induce mitochondrial aggregation (classified as gain-of-function alleles).[290] Although *MFN2* is ubiquitously expressed, neurons are thought to be particularly sensitive to defects in *MFN2,* as mitochondrial must localize across the axon to provide ATP production to dendrites and synapses.[290]

Similarly, dominant mutations in the ubiquitously expressed *RAB7* gene have been implicated in CMT Type 2.[291] RAB7 is a GTP-ase that regulates the dynamics of late endosomes and phagosomes, including their fusion with lysosomes[291]; CMT-associated mutations are thought to partially reduce RAB7 function.[292] Although *RAB7* is ubiquitously expressed, peripheral neurons may be particularly sensitive to partial loss of RAB7 function, since endosomes are required to carry neuronal signaling molecules across long axons.[293] Interestingly, a recent study by Cioni et al. demonstrated that ribosomes, mRNAs, RNA-binding proteins, and mitochondria frequently associate with endosomes in the axons of *Xenopus* retinal ganglia cells.[294] The authors also found that 35% of RAB7-marked endosomes were adjacent to mitochondria. Of these mitochondria-adjacent endosomes, 80% were associated with RNA granules, and 76% were associated with *de novo* protein translation, as measured by puromycin incorporation.[294] The authors demonstrated that two proteins important for mitochondrial function—Lamin B2 (which is required for mitochondrial integrity and axonal survival) and VDAC2 (which exchanges solutes across the outer mitochondrial membrane)—were translated at these late endosomes, indicating that these endosomes serve as a platform for the local translation of mitochondrial proteins.[294] Expression of CMT-associated *RAB7* mutations lead to decreased local translation of both Lamin B2 and VDAC, correlating with disrupted mitochondrial morphology and trafficking.[294] This study

highlights the interrelatedness of axonal trafficking, mitochondrial integrity, and local protein translation, as well as their importance for axonal health.

In addition to CMT Type 2, dominant ARS variants cause a closely related form of peripheral neuropathy, distal hereditary motor neuropathy (dHMN). dHMN results in a similar lower limb weakness and atrophy of the peroneal musculature, but primarily affects the motor neurons, compared to both sensory and motor neurons in CMT disease.[295] Similar to CMT disease, mutations that cause dHMN can affect genes specific to neuronal function, such as ion channels and axonal transporters, but also genes involved in basic cellular function such as RNA metabolism and DNA integrity.[295] In the case of *GARS1,* this motor neuron-predominant neuropathy is termed distal spinal muscular atrophy type V (dSMA-V), and, like *GARS1*-mediated CMT, predominantly affects the upper limbs.

Some ARS mutations, such as E71G and D500N *GARS1*, can cause both dSMA-V and CMT within the same family.[160,164,296] Considering that all affected family members carry the same genetic variant, it is unknown what dictates the degree of sensory neuron involvement that leads to the two distinct classifications. *De novo GARS1* variants can also cause a severe, infantile-onset form of dSMA-V. Individuals who are heterozygous for these *GARS1* variants have delayed or regressing motor milestones, severe muscle wasting, respiratory distress, and poor feeding abilities.[162,165,170,171] It is unclear what distinguishes these variants from the *GARS1* variants that cause a later onset, milder phenotype. One potential explanation is that these variants cause a unique neomorphic interaction that is toxic to motor neurons. These hypotheses are considered further in Section 1.4.4.

1.3 The molecular mechanisms of ARS-related genetic disease

1.3.1 Assays to evaluate impaired ARS function

One approach to defining pathogenic ARS alleles for both dominant and recessive disease is to determine if the variant impacts protein function. For variants with strong genetic evidence supporting their pathogenicity, these studies provide insight into the mechanism of disease. As the mechanism of disease is further defined, these approaches can also be used to build a case for

or against the pathogenicity of a newly characterized variant of uncertain significance. There are three commonly used assays for determining whether a variant impairs ARS function: 1) *in vitro* aminoacylation assays with recombinant proteins, 2) *in vitro* aminoacylation assays with patient cell lysates, and 3) *in vivo* yeast complementation assays.

Human ARS proteins can be expressed in bacteria, purified, and tested in an *in vitro* aminoacylation assay that measures its steady-state enzymatic activity.[297] Aminoacyl-tRNA synthetases charge tRNA with amino acids in a two-step reaction: in the first step, the enzyme binds ATP and the amino acid, resulting in the release of pyrophosphate. This step can be measured by treating the reaction with a colorimetric reagent that can quantify the amount of free phosphates released.[298] Then, the enzyme binds to the appropriate tRNA, ligating it to the amino acid and releasing AMP. The efficacy of the entire reaction can be assessed by incubating purified ARS enzyme with tRNA, ATP, and radiolabeled amino acid.[297] Aliquots of the mixture are then spotted on filter paper across time points. The tRNA is precipitated from the filter paper and washed to remove any unincorporated amino acid. Then, radioactivity of the amino acid that remains ligated to the tRNA is measured as the output of the reaction.[297]

Interpreting data from these kinetic assays can be challenging for partial loss-of-function variants. If these assays are performed with an excess of substrate (ATP, tRNA, and/or amino acid), this can mask the reduced function of a hypomorphic ARS. For example, the P234KY GARS1 protein was initially reported as fully functional allele when tested with saturating concentrations of glycine and tRNA.[69] When re-analyzed under Michaelis-Menten conditions, the mutation was found to significantly decrease enzymatic activity.[170] Furthermore, it is difficult to know what substrate concentrations best reflect substrate availability *in vivo*, or what degree of impaired function is required to have a downstream effect in the relevant tissue.

One way to assess ARS function in the context of an affected cell is a steady-state aminoacylation assay performed with patient cell lysates. In these assays, whole cell lysate is incubated in a reaction buffer that contains additional ATP, yeast total tRNA, and stable-isotope amino acid. The tRNA is then precipitated and washed before ammonia is added to release the stable-isotope amino acids from the tRNAs, which are then quantified with LC-MS/MS. This

assay has successfully been used to identify tRNA charging defects in patients cells with bi-allelic *CARS1*[126], *AARS1*,[105] *MARS1*,[105] *VARS1*,[257] *EPRS*,[138] and *NARS1*[221,222] variants. However, because this data reflects the total tRNA charging defect in the cell, it is not informative about the degree to which each of the two alleles contribute to the loss-of-function effect. Additionally, these reactions are performed with fibroblasts or lymphoblast cells that do not have a detectable phenotype, and so may not reflect the defective charging in the affected patient tissues. This assay has not been employed for cells from patients with dominant peripheral neuropathy, and is less likely to detect a significant defect in these cells, due to the later-onset and tissue-specific nature of ARS-mediated dominant neuropathy.

An alternate *in vivo* approach for assessing the impact of an ARS variant on gene function is a yeast complementation assay. This is a relatively straightforward, inexpensive, and quick approach that is well-suited to testing patient variants of uncertain significance. The assay relies on the fact that yeast need ARS function to survive and grow; as such, yeast growth serves as a proxy for ARS gene function. In these assays, the endogenous yeast ARS gene is deleted—temporary viability is maintained with the yeast ARS gene expressed from its endogenous promoter, on a *URA3*-bearing plasmid. The human ARS ortholog, expressed from a strong ubiquitous promoter on a high-copy number plasmid, can be transformed into yeast cells, and selected for using the *LEU2* auxotrophic marker expressed from the plasmid backbone. Then, yeast cells are plated on media containing the drug 5-fluoroorotic acid (5-FOA). Ura3 converts 5-FOA into a toxic compound, causing cells that express *URA3* from the maintenance vector to die, and selecting for the cells that have spontaneously lost the maintenance vector during cell division.[299] The only source of ARS in these surviving yeast is the exogenous human ARS, and the function of this gene determines whether the yeast will grow or not. In this assay, the growth of yeast expressing patient mutations are compared to the growth of yeast expressing wild-type ARS. An empty vector serves as a negative control, ensuring that 5-FOA selection is complete.

Yeast complementation assay are an important *in vivo* counterpart to *in vitro* aminoacylation assays, as they demonstrate the degree of ARS impairment required to impact a living cell. Yeast complementation assays almost always correspond to results from *in vitro* aminoacylation assays; variants that significantly impair *in vitro* aminoacylation also do not support yeast

growth. For example, the re-analyzed P234KY *GARS1* variant (discussed above) does not support yeast growth, consistent with its loss of function *in vitro*[170] (Figure 1.1). Similarly, ΔETAQ *GARS1*, which is severely enzymatically impaired *in vitro*[170] does not support yeast growth (Figure 1.1). Occasionally, yeast complementation data is discordant with enzymatic assay, specifically when human variants are modeled in the yeast ortholog. For example, G240R GARS1 significantly impairs enzyme function *in vitro*[300], but does not reduce yeast growth when modeled in *GRS1*, the yeast ortholog of *GARS1*.[301] However, when G240R is modeled in the human *GARS1* open reading frame, it significantly impairs yeast growth compared to wild-type *GARS1* (Figure 1.1), consistent with the *in vitro* enzymatic data for the human protein. This highlights the discrepancies that can arise between functional assays with the yeast gene and functional assays with the human gene, and the importance of testing variants in the human ARS gene when possible.

One limitation of the yeast complementation assay is the incongruence between a yeast cell and the human tissues affected in ARS-related disease, particularly neurons, which are affected in nearly all patients with dominant or recessive ARS-related phenotypes. On one hand, mutations that may be deleterious in the relevant human context may not affect yeast growth, particularly when over-expressed in rapidly dividing cells growing in nutrient-rich media. This is especially relevant to mutations that may only moderately reduce enzymatic function; although this reduction may be tolerable to yeast, it might be pathogenic in a human patient if it is *in trans* with a null mutation. On the other hand, reduced function of the human gene in a yeast cell may not reflect the effect of pathogenic variants in affected human tissues due to differences in the human transcriptome and proteome, and the tissue-specific availabilities of ARS substrates.

1.3.2 Reduced enzymatic function causes recessive ARS diseases

Bi-allelic variants identified in patients with ARS-mediated disease are typically missense mutations, splice site mutations, or insertions or deletions that lead to a frameshift and a premature stop codon. Premature stop codons generate null alleles, reducing total ARS protein levels.[145,159,103] Similarly, splice site mutations cause aberrant inclusion of introns[242] or exclusion of exons, shifting the frame of protein translation and introducing premature stop codons that

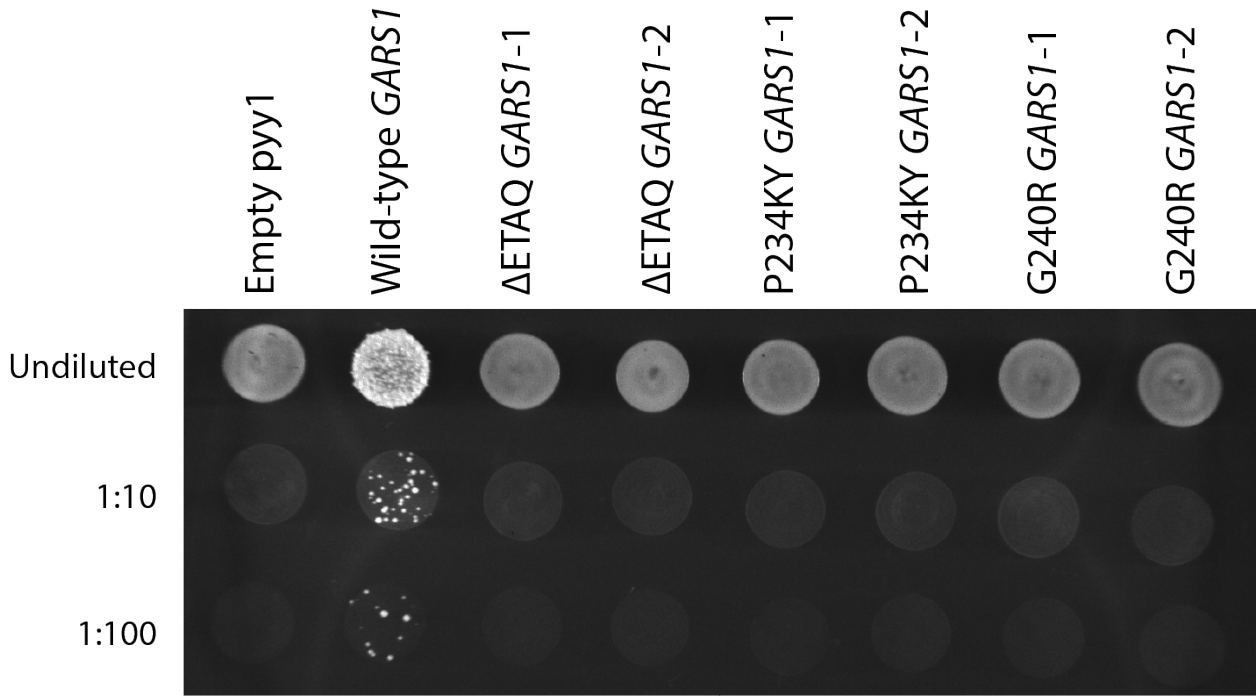


Figure 1.1. Pathogenic *GARS1* variants do not support yeast growth.

Representative images of yeast strains lacking endogenous *GRS1*, transformed with an empty vector ("Empty pyy1"), a vector expressing the wild-type human *GARS1* open reading frame, or vectors expressing the mutant human *GARS1* open reading frame. Two independent colonies were tested for each *GARS1* mutation (from left to right, ΔETAQ, P234KY, and G240R). Each colony was plated undiluted or in 1:10 serial dilutions on media containing 5-FOA. Yeast were grown at 30°C.

generate null alleles.[147] Recessive missense mutations either impair protein stability, decrease enzymatic function, or both. In many cases, the primary defect of the missense mutation appears to be through decreased protein levels. This is detected through immunostaining for the affected ARS protein in patient cell lines, and comparing protein levels to those of heterozygous parents or wild-type controls. For example, two patients who are compound heterozygous for different *AARS1* missense variants have 15% or 30% of wild-type AARS1 levels.[105] Similarly, individuals who are homozygous for a missense *MARS1* variant, or homozygous for a missense *TARS1* variant have 30% of wild-type MARS1 levels and less than 20% of wild-type TARS1 levels, respectively.[105,253] Perhaps most dramatically, an individual who is compound heterozygous for a missense mutation and a frameshift mutation in *FARSB* showed an approximate 97% reduction in FARSB levels[145], indicating that the missense allele produced only a small amount of FARSB protein.

In these cases, any existing protein must have some enzymatic function, or the genotype would be incompatible with life. To discern how much function is retained, it is necessary to test the allele in isolation, either using *in vitro* aminoacylation assays or *in vivo* yeast complementation assays. These studies can be useful for estimating the total decrease in ARS activity in a patient cell, which can inform efforts to understand the relationship between genotype and phenotype. For example, two individuals could each have an 80% reduction in ARS protein. However, if one individual has almost no functional enzyme in that remaining 20%, their phenotype is likely to be more severe than an individual with a fully functional 20%.

In some cases, the primary defect does appear to be in protein abundance, not enzymatic function. For example, a recent study evaluated patients with a multi-system ataxic syndrome and bi-allelic *HARS1* variants. One patient was compound heterozygous for D206Y *HARS1* and V244Cfs*6 *HARS1*.[172] V244Cfs*6 *HARS1* was not expressed at the mRNA level, indicating that it was unstable or subjected to nonsense-mediated decay.[172] This allele would be expected to decrease HARS1 levels by 50%. However, total HARS1 protein levels were significantly below 50% of wild-type levels, indicating that D206Y also impairs protein production or stability. When D206Y was tested in a yeast complementation assay, it did not lead to a decrease in yeast growth, indicating that it did not significantly impair gene function when over-expressed in

yeast.[172] This suggests that in human cells, D206Y is deleterious primarily through decreasing protein levels, and not through significantly decreasing protein function.

In other cases, a missense allele can decrease both protein levels and protein function. Here, an informative case is found in two siblings presenting with microcephaly, hypomyelination, and epileptic encephalopathy, both of whom were compound heterozygous for Y690Lfs*3 *AARS1* and G913D *AARS1*.[103] Immunostaining of patient and parent-derived lymphoblastoid cells showed that Y690Lfs*3 decreased AARS1 protein levels by ~50% in a heterozygous parent and did not result in a truncated AARS1 band (of note, these studies were performed with an antibody to the N terminus that would recognize a truncated protein).[103] Compared to wild-type controls, the compound heterozygous patient had less than 20% of wild-type AARS1 protein levels—50% of this decrease was likely due to Y690Lfs*3, and the other 30% or more due to the G931D allele.[103] To discern how much activity G931D retained, *in vitro* aminoacylation assays were performed. G931D decreased enzymatic activity by ~70% compared to WT AARS1.[103] This indicates that, in addition to G931D decreasing protein abundance, any remaining G931D has decreased enzymatic activity.

Lastly, it is possible that the primary defect of a missense variant is in enzymatic activity, rather than decreased protein levels. To date, this has only been effectively demonstrated for one variant, R310Q *GARS1*. This variant was identified *in trans* with a frameshift variant, G831Ifs*6 *GARS1*, in a patient with a severe multisystem disorder.[159] Careful analysis of protein levels from the patient and from unaffected wild-type controls identified a 50% reduction of GARS1 protein, consistent with G831Ifs*6 ablating protein expression and the remaining GARS1 protein representing the R310Q allele.[159] *In vitro* measurements of initial charging velocity for R310Q GARS1 found that it decreases enzymatic activity by greater than 99% compared to wild-type GARS1.[159] It is likely that R310Q retains more function in patient tissues, as such a significant decrease would be incompatible with life. Regardless, this demonstrates that R310Q primarily decreases *GARS1* function by impairing enzymatic activity, not by decreasing protein levels.

From these studies and others, it is clear that recessive ARS-mediated disease is due to a loss-of-function mechanism. This is also supported by work on *IARS*,[182] *QARS*,[230] and *KARS1*,[194] in

zebrafish, and models of *HARS2*[42] and *LARS2*[96] in *C. elegans* demonstrating that knocking down the ARS gene recapitulates key elements of the recessive disorder. However, there is still work to be done to define the downstream molecular consequences of reduced ARS activity. On this front, studies of mutations in mitochondrial ARS genes lead the way.

Investigations of recessive mitochondrial ARS disorders have demonstrated a relationship between decreased mitochondrial ARS activity, reduced protein translation, and impaired mitochondrial function. Studies of bi-allelic partial-loss-of-function variants in *VARS2*,[254] *TARS2*,[254] and *NARS2*[50] have examined corresponding changes in aminoacylated tRNA levels. Total RNA was extracted from patient cells under acidic condition to preserve the ester bond between tRNA and amino acid, separated on a urea polyacrylamide gel, transferred to a membrane, and hybridized with radiolabeled probes against the respective tRNAs.[50,254] Then, the amount of high molecular weight aminoacylated tRNA was compared to the amount of low molecular weight uncharged tRNA. Defects in VARS2 and TARS2 corresponded to a 50% and 60% reduction in aminoacylated $tRNA^{Val}$ or $tRNA^{Thr}$, respectively;[254] defects in NARS2 corresponded to a similar depletion of aminoacylated $tRNA^{Asn}$.[50] To investigate how decreased abundance of charged tRNA corresponded to the function of the mitochondrial respiratory chain (MRC) complexes, each member of the MRC complexes was analyzed using spectrophotometric assays that analyze the appearance or disappearance of MRC substrates and products.[50,254] Muscle samples from all patients demonstrated reduced function of at least one Complex, with the exception of Complex II, which is the only complex comprising proteins translated in the cytoplasm.[50,254] All samples also showed a decrease in oxygen consumption, consistent with mitochondrial defects.[50,254] Similar studies performed for other mitochondrial ARS variants show either decreased abundance of MRC complex proteins (for patients with variants in *FARS2*,[82] *LARS2*,[85] or *AARS2*[302]), decreased activity of MRC complexes (for patients with variants in *EARS2*[141] or *RARS2*[244]), or both (for patients with variants in *YARS2*[303] and *MARS2*[91]). Comprehensively, these studies demonstrate that reduced mitochondrial ARS function leads to reduced levels of aminoacylated tRNA, reduced abundance and/or activity of mitochondrial ARS proteins, and impaired mitochondrial function.

In sum, these studies provide evidence that recessive ARS alleles cause a partial reduction of ARS activity. This could ultimately result in impaired protein translation by one of two mechanisms (Figure 1.2). First, impaired translation could be a direct result of uncharged tRNA binding to GCN2 through its tRNA recognition domain, which subsequently triggers GCN2 to phosphorylate eIF2α.[3] Phosphorylated eIF2α then blocks the recycling of GDP-eIF2 to GTP-eIF2, preventing it from delivering initiator $tRNA^{Met}$ to the ribosome and causing global shut-down in protein synthesis.[3] Second, a dearth of charged tRNAs for a given amino acid could cause the ribosome to stall at these amino acid codons, reducing protein expression from the transcript.[304] Interestingly, ribosome stalling has also been shown to activate GCN2;[305] this suggests that both mechanisms of decreased protein synthesis may co-exist in cells with reduced ARS activity.

1.3.3 Proposed mechanisms of dominant ARS disease

In contrast to the recessive phenotypes, the molecular mechanism of ARS-related dominant axonal neuropathy is less clear. The fact that mutations in five genes encoding an aminoacyl-tRNA synthetases (*GARS1*, *YARS1*, *AARS1*, *HARS1*, and *WARS1*) cause a similar dominant phenotype points to a common disease mechanism. In support of this, over-expression of neuropathy-associated *GARS1* and *YARS1* mutants in a *Drosophila* model cause a strikingly similar phenotype.[306] Although there is a growing body of work defining various non-canonical functions for the above five ARS enzymes (Table 1.3), none are common to all five enzymes, nor do they relate to neuron (or axon) function. Thus, it is currently challenging to investigate if a loss of some non-canonical function is responsible for ARS-related neuropathy. Also, considering that most neuropathy-associated ARS mutations impair rather than enhance enzyme function,[307] it is unlikely that a gain of canonical function is responsible for disease. As a result, there are currently two mechanisms being explored: impaired ARS activity and toxic gain-of-function effects (Figure 1.3). It should be emphasized that these pathogenic mechanisms may not be mutually exclusive—for example, impaired tRNA charging may be a prerequisite for a gain-of-function effect, or the two molecular consequences may work in concert to modulate phenotypic severity.

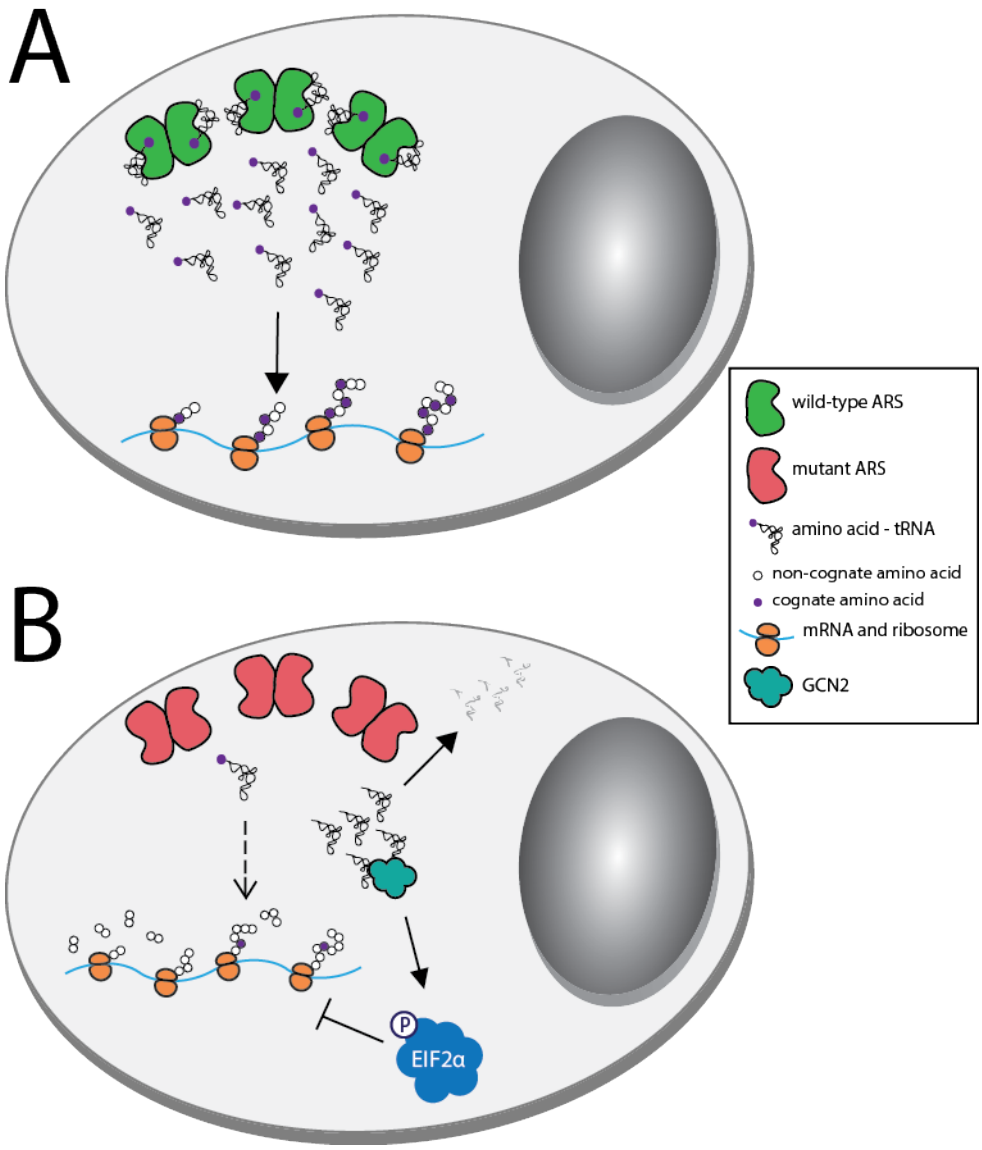


Figure 1.2. Potential mechanisms of ARS-related recessive disease.

(A) Two wild-type ARS alleles supply cells with the requisite charged tRNA for protein translation. **(B)** Two loss-of-function ARS alleles severely reduce the amount of charged tRNA available for translation, which impairs protein production. Uncharged tRNA is either degraded or binds to GCN2, which phosphorylates eIF2α and inhibits global translation. In both panels, dimeric enzymes functioning in the cytoplasm are shown for simplicity; however, please note that some ARS enzymes act as monomers and that some effects apply to mitochondrial translation.

Table 1.3. Non-canonical functions of dominant disease-associated ARS proteins.

ARS	Species	Non-canonical function
AARS1	*Homo sapiens*	C-terminal splice variant binds DNA[30]
GARS1	*Bos taurus*	Localizes to the nucleus and activates NFκB1 and mTOR gene expression[308]
	Homo sapiens	Secreted in response to damage in mesenchymal stem cells, promotes differentiation and migration[309]
	Homo sapiens	Chaperone in neddylation pathway[310]
	Homo sapiens	Tumorigenesis defense[311]
	Saccharomyces cerevisiae	mRNA 3' end formation[312]
HARS1	*Danio rerio*	Angiogenesis regulation[313]
	Saccharomyces cerevisiae	Autoregulatory repression of *HARS* mRNA translation in response to low tRNA levels[314]
	Homo sapiens	Epitope for autoantibodies in inflammatory myositis[315–317]
WARS1	*Homo sapiens*	In response to IFN-γ stimulation, is a cellular entry factor for *Enterovirus*[318]
	Homo sapiens	In response to IFN-γ stimulation, increases tryptophan uptake into cells[319]
	Homo sapiens	Upon secretion from cells, participates in the antiviral innate immune response[64,320]
	Homo sapiens	Mini-WARS1 inhibits angiogenesis[65,321–324]
	Homo sapiens	In response to IFN-γ stimulation, facilitates p53 activation[325]
YARS1	*Homo sapiens*	Fragmented YARS1 stimulates megakaryopoiesis and platelet production[326]
	Homo sapiens	Locates to the nucleus and protects against DNA damage[327–329]
	Homo sapiens	Mini-YARS1 promotes angiogenesis[63,322–324,330]

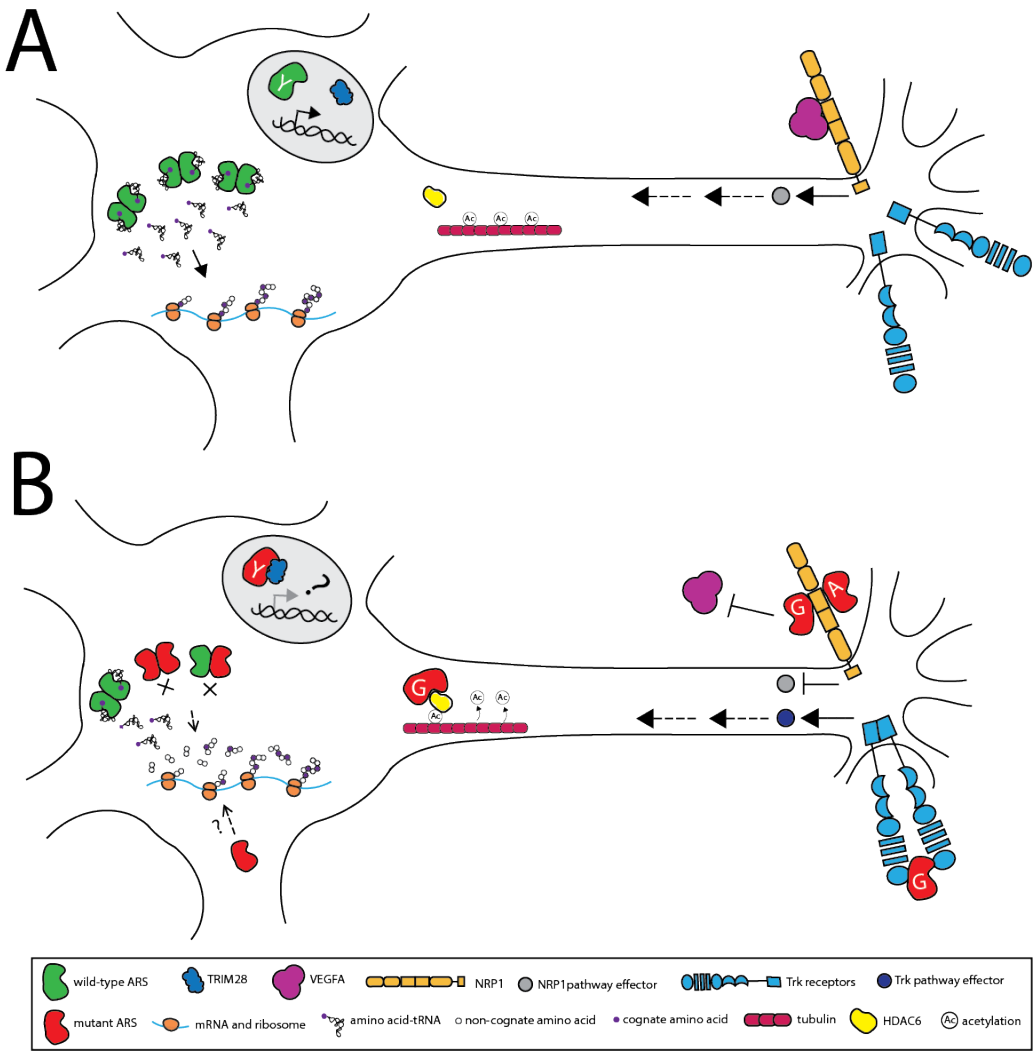


Figure 1.3. Potential mechanisms of ARS-related dominant axonal neuropathy.

Neurons are illustrated with the cell body on the left and the axon extending to the right. A wild-type neuron **(A)** has functional ARS activity (green dimers) facilitating protein translation. There is appropriate NRP1 (orange) and Trk signaling (blue). YARS1 translocates to the nucleus upon oxidative stress and binds TRIM28 (blue), changing the regulation of DNA damage response genes. HDAC6 (yellow) de-acetylates tubulin under homeostatic conditions. Proposed mechanisms of ARS-mediated peripheral neuropathy are represented in **(B)**; neuronal function may be compromised by impaired protein translation due to an unknown function of mutant ARS (red subunits) and/or a depletion in available charged tRNA from a reduction of aminoacylation activity. For peripheral neuropathy related to *GARS1*, mutant GARS1 may interfere with NRP1 signaling by preventing VEGFA (magenta) from binding to NRP1. In developing sensory neurons, mutant GARS1 may also act as a ligand for Trk receptors, aberrantly activating Trk signaling. Mutant GARS1 could also bind to HDAC6, increasing its de-acetylation activity and leading to hypo-acetylated tubulin. For peripheral neuropathy related to *YARS1*, increased mutant YARS1 binding to TRIM28 (blue) may alter regulation of DNA damage response genes or genes important for neuronal function.

1.3.4 Impaired ARS function in dominant axonal neuropathy

Data from *in vitro* aminoacylation and yeast complementation assays indicate that almost all neuropathy-associated ARS mutations cause deficits in tRNA charging.[110,170,171,176,307] Additionally, data from three animal models suggest that the mutant proteins are sub-functional. First, in *Drosophila* projection neurons, the morphological defects caused by neuron-specific homozygosity for a null *gars1* allele were fully rescued by a wild-type human *GARS1* transgene, partially rescued by the neuropathy-associated E71G allele, and not rescued by the neuropathy-associated L129P allele.[331] Second, whereas over-expressing wild-type *gars1* rescues the neuromuscular phenotype of zebrafish homozygous for a loss-of-function allele, over-expressing the neuropathy-associated G526R allele showed no rescue.[332]. Finally, mice heterozygous for P234KY or C201R *Gars1* display a dominant neuropathy, but homozygosity for these mutations is lethal.[333] In sum, there is an abundance of data showing that neuropathy-associated ARS missense mutations have a deleterious effect on gene function, indicating that this is a component of disease pathogenesis.

In contrast, three lines of evidence argue against a simple loss-of-function effect (*i.e.*, haploinsufficiency) as the underlying mechanism of ARS-related neuropathy. First, mice heterozygous for a *Gars1* null allele have a wild-type phenotype.[69] Second, none of the individuals with ARS-mediated dominant peripheral neuropathy are heterozygous for a null allele. However, recessive individuals are frequently compound heterozygous for a missense and a null allele. In these cases, parents carrying the null allele are not reported to have peripheral neuropathy symptoms, although few have been clinically evaluated. Additionally, data from the Genome Aggregation Database (gnomAD) indicate that *AARS1, HARS1, GARS1, YARS1,* and *WARS1* are tolerant of protein truncating variants in the heterozygous state.[334] In this database, the probability that a gene is haploinsufficient is calculated by comparing the number of protein-truncating variants that would be expected in the gene based on its sequence content, size, and methylation to the number of protein-truncating variants seen in the general population (excluding individuals with pediatric diseases and their first degree relatives).[334,335] This probability is communicated as "pLI" scores, or "probability of being loss-of-function intolerant," where the pLI score for a haploinsufficient gene is 1. The pLI scores for the five ARS genes implicated in dominant peripheral neuropathy are: 0 (*AARS1*), 0 (*HARS1*), 0

(*YARS1*), 0.31 (*GARS1*), and 0.3 (*WARS1*).[334] Therefore, heterozygosity for a null allele in any of these five genes does not cause a phenotype that is subject to purifying selection. Since gnomAD only excludes individuals with known pediatric disease, we do not know if these heterozygous null individuals in the database have a peripheral neuropathy. However, cumulatively, these data strongly argue against haploinsufficiency as the primary mechanism for a penetrant, ARS-related neuropathy.

One possible explanation for the role of loss-of-function missense mutations in dominant neuropathy is a dominant-negative effect. Dominant-negative mutations, or anti-morphs, are loss-of-function mutations that also inhibit the function of the wild-type gene product.[336] The catalog of Mendelian diseases is replete with pathogenic dominant-negative alleles in genes encoding transcription factors such as p53,[337] receptors like G-protein subunits[338], ion channels[339,340], and structural proteins like keratins.[341] Dominant-negative mutations in enzymes are rarer, likely because it takes a large change in enzymatic concentration or function to impact metabolic flux.[342] As such, most loss-of-function mutations in enzymes are recessive.[342] However, dominant-negative mutations have been described in *PLKR* [343] (encoding pyruvate kinase, a critical metabolic enzyme), *PSEN1*[344] (which encodes a member of the gamma secretase complex that cleaves amyloid precursor protein), *UBC12*[345] (a neddylation conjugating enzyme), and *GALT*[346] (an enzyme involved in galactose metabolism).

Prerequisites for a dominant-negative effect include: (**1**) the mutant protein should be stably expressed; (**2**) the mutant protein should have reduced or ablated function; and (**3**) the affected protein should normally dimerize (or oligomerize) and mutant subunits should retain the ability to interact with wild-type subunits. Indeed, AARS1, YARS1, GARS1, HARS1, and WARS1 all charge tRNA in the cytoplasm as dimers (Table 1.1); if an inactive mutant subunit dimerizes with a wild-type subunit, it could result in a significant reduction in tRNA charging compared to the haploinsufficient state. This would shift the burden of tRNA charging onto the reduced population (*i.e.*, 25%) of wild-type/wild-type dimers.

Multiple lines of evidence support a dominant-negative effect: **(1)** data from primary patient cell lines indicate that pathogenic ARS alleles do not reduce the total amount of protein detectable

via Western blots[170,301,333,347,348]; **(2)** *in vitro* and *in vivo* functional assays demonstrate that the vast majority of pathogenic ARS variants impair enzyme function,[110,170,171,176,264,307,349] and **(3)** ultracentrifugation experiments or co-immunoprecipitation experiments have shown that mutant ARS retain the ability to dimerize.[176,264,275,300,350] However, no studies have demonstrated that mutant human ARS is dominantly toxic when co-expressed with wild-type human ARS, nor that this toxicity is dependent on the hetero-dimerization between wild-type subunits and mutant subunits.

Several studies have generated data that is consistent with a dominant-negative effect. The first study found that yeast cells expressing one wild-type and one mutant copy of yeast tyrosyl-tRNA synthetase showed reduced growth compare to yeast cells expressing only the wild-type enzyme.[275] The second study identified a loss-of-function zebrafish *gars1* mutation, T209K, that ablates dimerization.[332] Zebrafish that are homozygous for this allele show a severe neuromuscular defect, and zebrafish heterozygous for this allele have no phenotype.[332] When T209K was over-expressed in either *gars1*$^{T209K/+}$ or *gars1*$^{+/+}$ zebrafish, the fish had no phenotype, indicating that a dimer-reducing loss-of-function *gars1* mutation is not dominantly toxic.[332] However, over-expression of G526R *gars1*, which dimerizes[332] but is non-functional[301], caused neuromuscular junction defects. Notably, over-expression of T209K in *cis* with G526R improved the neuromuscular junction phenotype, suggesting that dimerization is required for the toxicity of G526R *gars1*.[332] The third study investigated H257R *WARS1*, which decreases enzyme activity *in vitro* but does not impact dimerization.[264] To measure the potential downstream impact on protein synthesis, cultured cells were co-transfected with a construct to express wild-type or H257R *WARS1* (or an empty vector) and a plasmid expressing β-Gal or luciferase. β-Gal or luciferase activity was interpreted as a read-out for translation of the respective enzyme, and protein synthesis as a whole. Whereas wild-type WARS1 increased reporter activity above that of the empty vector, H257R WARS1 decreased reporter activity below that of empty vector.[264] There are significant limitations to this approach, including an inability to control for a consistent copy number of each vector across a population of transfected cells, and the reliance on the enzymatic activity of two reporter proteins as an indication of global protein synthesis. However, the authors concluded that these data demonstrated that H257R WARS1 exhibited a dominant-negative effect.

Lastly, a study of *HARS1* variants provides significant circumstantial evidence of a dominant-negative mechanism of neuronal toxicity. Mullen et al., examine the variant R137Q *HARS1*[175], as well as V155G *HARS1* and Y330C *HARS1*, both of which are characterized in Section 2.3.2. The authors over-expressed these alleles in PC12 cells, which can be differentiated to generate axon-like projections. Over-expression of all three *HARS1* alleles showed an increase in EIF2α phosphorylation, a marker of the integrated stress response and an indication of accumulating uncharged tRNA,[351] consistent with significantly reduced HARS1 function.[350] This was accompanied by an approximate 20% reduction in global protein synthesis, as measured by OP-Puromycin incorporation.[350] This also corresponded to a modest decrease in the length of the longest neurite in each cell.[350] These findings suggest that the ability to form or maintain long neuronal processes, such as the long axons of the peripheral nerve, is dependent on protein synthesis. Critically, these three phenotypes—increased EIF2α phosphorylation, decreased protein synthesis, and shortened length of the longest neurite—were recapitulated when the cells were treated with histidinol, a small-molecule inhibitor of HARS1.[350] This directly demonstrates that pharmacological impairment of HARS1 phenocopies the toxicity of dominant *HARS1* mutations, consistent with a loss-of-function effect achieved through a dominant-negative mechanism.

As a complementary approach, Mullen et al. injected wild-type zebrafish embryos with V155G or Y330C human *HARS1* mRNA. By 48 hours post fertilization, these zebrafish neurons showed improper guidance.[350] Neuronal processes in fish expressing mutant HARS1 were also shorter than those of fish expressing wild-type human HARS1 protein.[350] Unsurprisingly for such severe morphological defects, the fish also displayed motor deficits in behavioral assays.[350] This phenotype was replicated by treating the fish with the protein synthesis inhibitor cyclohexamide, which also shortened the length of the neuronal processes in the dorsal root ganglia.[350] This demonstrates that chemically inhibiting protein synthesis will mimic the phenotype of *HARS1* mutations, supporting the hypothesis that reduced protein synthesis is part of the ARS-associated neuropathy disease mechanism.

It is important to consider how impaired ARS function would specifically affect peripheral nerve axons. Of note, Mullen et al. found that over-expressing patient *HARS1* mutations in PC12 cells did not affect the size of the cell body or the number of neurite projections the cells produced, but did decrease the length of the longest neurite; in other words, cells could not grow or sustain projections beyond a certain distance from the soma.[350] Similarly, expression of these mutations in zebrafish caused a decreased axon length in dorsal root ganglia cells.[350] It is possible that maintaining the health of long axons, such as the long axons of the peripheral nervous system, is particularly difficult with defects in housekeeping functions, such as protein translation. Indeed, mutations in other ubiquitously expressed genes (*e.g.*, *MFN2* and *RAB7*) have been implicated in axonal neuropathy[289,352], as discussed in Section 1.2.4.

However, a dominant-negative effect may not apply to all neuropathy-associated ARS mutations. One major piece of evidence points away from this as a unifying mechanism; in theory, if ARS alleles are dominant-negatives, then loss of ARS function should drive the pathology, which should then be rescued by over-expression of wild-type ARS. However, in two mouse models of *Gars1*-mediated peripheral neuropathy, *Gars1*$^{P234KY/+}$ and *Gars1*$^{C201R/+}$, over-expressing wild-type human GARS1 was not sufficient to rescue the phenotype[333]. There are significant caveats to these two mouse models; neither represent a human disease allele, and both cause early-onset phenotypes that are similar to the mouse model of ΔETAQ *GARS1,* a mutation which causes an early-onset, severe SMA-like phenotype in humans.[170] In particular, *Gars1*$^{P234KY/+}$ causes premature lethality that is not seen in individuals with dominant peripheral neuropathy.[69] As such, it remains to be seen whether these findings can be generalized to other ARS alleles that cause a later-onset, milder phenotype in human. As future work investigates a dominant-negative mechanism, it will be important to determine if it can be alleviated by supplying additional wild-type enzyme to the cell. To the best of our knowledge, there are only a small number of studies showing that over-expression of a wild-type allele rescues a dominant-negative effect; moreover, these studies have been performed *in vitro*, and for dominant-negative alleles in structural proteins (Type VII Collagen[353]) or receptors (Follicle Stimulating Hormone Receptor[354]). It remains to be seen whether a dominant negative mutation in an enzyme can be rescued by over-expression of the wild-type protein.

1.3.5 Proposed gain-of-function mechanisms in dominant axonal neuropathy

In contrast to a loss-of-function mechanism of disease, a second possibility is that neuropathy-associated ARS mutations cause the encoded enzymes to gain a novel, dominantly toxic function that specifically affects the peripheral nervous system. This may not be mutually exclusive with a loss-of-function effect, as a mutation may simultaneously impair enzymatic activity and facilitate novel binding partners. There is evidence that *YARS1*,[355] *GARS1*,[356] *HARS1*,[357] and AARS1[348] mutations change the conformation of the enzyme and expose internal protein residues, which are posited to be new binding interfaces for aberrant protein interactions.

One aberrant interaction that has been explored for mutations in both *GARS1* and *AARS1* is neomorphic binding to neuropillin-1 (*NRP1*). Neuropilin-1 is a transmembrane protein that participates in the development of the nervous system and cardiovascular system.[358] It acts as a receptor for semaphorin axon guidance factors, as well as a receptor for vascular endothelial growth factor (VEGF)[358], which is critical for both neuron development and angiogenesis.[359] A 2015 study that aimed to identify novel binding partners of mutant GARS1 found that immunoprecipitation of NRP1 co-immunoprecipitated a small amount of wild-type GARS1, and a significantly increased amount of mutant GARS1 (here, the high-confidence variants L129P, G240R, and E71G were tested, along with the mouse spontaneous mutation P234KY).[360] By systematically deleting different NRP1 domains and assessing whether NRP1 could still co-immunoprecipitate GARS1, the authors mapped the GARS1 binding location to the b1 domain, which is the binding site of VEGF-A_{165}.[360] The authors then showed that increasing concentrations of VEGF-A_{165} could decrease the amount of mutant GARS1 bound to NRP1 *in vitro*, and vice versa.[360] Additionally, to explain what would be an extracellular interaction for a cytoplasmic protein, the authors demonstrated that GARS1 is present in exosomes of NSC34 motor neuron-like cells by enriching for exosomes in the cell media and immunostaining for GARS1.[360] This finding was supported by treating cells with an exosome-pathway inhibitor, which decreased the amount of extracellular GARS1 detected, and by treating cells with an activator of microvesicle release, which increased the amount of GARS1 detected.[360]

The authors followed this work by investigating the interaction between NRP1 and mutant GARS1 in *Gars1*$^{P234KY/+}$ mice (as noted above, this mouse does not model a human disease

allele, and exhibits a severe early-onset neuromuscular phenotype with pre-mature death between 5 and 8 weeks of age[69], which does not recapitulate patient phenotypes).[360] The authors found that *Gars1*$^{P234KY/+}$ mice had defects in facial neuron migration, similar to those seen in mouse models of *Nrp1* and *Vegf* mutants.[360] They also found that *Gars1*$^{P234KY/+}$ mice who were also heterozygous for a *Nrp1* null allele had a significantly more severe phenotype, indicating that *Nrp1* is a genetic modifier of the mouse pathology.[360] Finally, the authors showed that treating *Gars1*$^{P234KY/+}$ with VEGF-A_{165}—but no other trophic factors—improved motor function. From these studies, the authors posit that the mechanism of GARS1-mediated peripheral neuropathy is due to a gain-of-function interaction with NRP1.

To demonstrate that this interaction is not limited to GARS1, these authors have performed similar *in vitro* studies demonstrating an interaction between mutant AARS1 alleles and NRP1.[348] Here, both NRP1 and AARS1 were ectopically expressed in NSC34 motor neuron-like cells. Immunoprecipitation of NRP1 co-immunoprecipitated three mutant forms of AARS1 (N71Y, G102R, and R329H).[348] However, NRP1 did not co-immunoprecipitate wild-type AARS1, or any AARS1 produced from three mutations in the editing domain and C-terminal domain.[348] The authors also demonstrated that immunoprecipitation of NRP1 in patient lymphocyte cells co-immunoprecipitated R329H AARS1, but did not co-immunoprecipitate wild-type AARS1 in control lymphocytes.[348] Lastly, they pursue a similar domain mapping strategy as they performed for GARS1, and determine that mutant AARS1 binds to the same b1 domain as GARS1.[348] Here, to explain the interaction between a cytoplasmic protein and an extracellular domain, the authors remove cell culture media from HEK293T cells and immunoblot for AARS1. They do not demonstrate that it is present in exosomes, or control for the possibility of cell lysis.

There are several major gaps in this model of ARS-mediated peripheral neuropathy. One inconsistency is the seemingly neuron-specific interaction between GARS1 and NRP1, despite the fact that GARS1 is ubiquitously expressed, and the signaling pathway of NRP1 and VEGF-A_{165} is not only critical for neuronal development but for cardiovascular development as well. If GARS1 interferes with VEGF-A_{165} binding to NRP1 in other tissues, this interaction should cause cardiovascular defects. However, this is not a known phenotype of any patient with

GARS1-mediated peripheral neuropathy. This has also been assessed in another *Gars1* mouse model, *Gars1*$^{C201R/+}$ (this mutation is also not a patient mutation, but arose from *ENU* mutagenesis), where no cardiovascular defects were identified.[361] The expression pattern of NRP1 is also inconsistent with the human phenotypes, as NRP1 plays a critical role in development,[362] and the ARS-mediated peripheral neuropathy is a degenerative process that usually onsets in adolescence or adulthood. Finally, if impaired NRP1 signaling were to play a role in any of the human neuropathies, it would most likely contribute to the severe, childhood-onset motor neuropathies, like that caused by ΔETAQ *GARS1*.[170] Indeed, a *Gars1*$^{\Delta ETAQ/+}$ mouse model shares a similar phenotypic severity with the *Gars1*$^{P234KY/+}$ mouse discussed above.[170] However, overexpression of ΔETAQ GARS1 in NSC34 cells followed by immunoprecipitation of NRP1 did not detect an interaction between ΔETAQ GARS1 and NRP1.[170] Mass spectrometry of ΔETAQ GARS1 binding partners in motor neuron cells also failed to detect an interaction with NRP1.[170] Based on the inability to replicate this interaction with one of the most toxic *GARS1* mutations, it is unlikely that neomorphic binding to NRP1 is a common mechanism of ARS-mediated dominant peripheral neuropathy.

An alternate gain-of-function mechanism that has been proposed is an aberrant interaction with members of the tropomycin receptor kinase (TRK) family—TrkA, TrkB, and TrkC. Trk proteins are membrane-bound receptors that bind neurotrophic growth factors in a signaling pathway required for the proper development of sensory neurons.[363] This proposed mechanism stems from the observation that sensory neuron fate is disturbed in *Gars1*$^{C201R/+}$ mice, with a prenatal imbalance in subtypes of sensory neurons that leads to changes in sensory behavior upon birth.[364] To test a possible interaction between mutant GARS1 and Trk proteins, wild-type, P234KY, or C201R GARS1 were transfected into NSC34 cells.[364] Then TrkA, TrkB, or TrkC were immunoprecipitated, and co-immunoprecipitated proteins were immunostained for GARS1.[364] All three members of the Trk family co-immunoprecipitated P234KY and C201R GARS1, although not wild-type.[364] To validate this interaction with human *GARS1* mutations, recombinant wild-type, L129P, or G240R GARS1 was added to the media of N2a neuroblastoma cell lines.[364] Both mutations (but not wild-type) caused an increase in ERK1/2 phosphorylation, an integral component of the Trk signaling cascade.[364] These data lead the authors to propose that extracellular mutant GARS1 aberrantly binds and activates Trk receptors, changing sensory

neuron differentiation and/or survival early in development. Although it is intriguing to consider the possibility that developmental defects predispose sensory neurons for later-stage degeneration in ARS-mediated CMT, additional *in vivo* work is required to link these defects to impaired Trk signaling. Additional work is also required to test the specificity of this interaction; if this interaction can be detected between Trk receptors and a benign GARS1 polymorphism, or between Trk receptors and a mutant GARS1 that is strictly linked to motor neuron phenotypes such as ΔETAQ, it would argue against this as a mechanism of sensory neuron impairment in *GARS1*-mediated CMT.

An alternate possibility is that *GARS1* mutations cause inappropriate binding of intracellular proteins, specifically HDAC6, a histone deacetylase that acts on alpha-tubulin in the cytoskeleton.[365] One of several post-translational modifications that regulate tubulin function,[366] alpha-tubulin acetylation promotes axonal transport by increasing kinesin binding to microtubules.[367] In this model, HDAC6 activity decreases alpha-tubulin acetylation, which leads to decreased kinesin binding and axonal transport. Defects in axonal transport are linked to a number of neurological diseases, including peripheral neuropathies.[295]

Studies of the *Gars1*$^{P234KY/+}$ and the *Gars1* $^{C201R/+}$ mouse models have found reduced tubulin acetylation in the dorsal root ganglia and sciatic nerve, accompanied by defects in axonal transport.[368,369] Inhibiting HDAC6 with the small molecule tubastatin A rescues these defects and improves the motor function of both mouse models.[368,369] Two studies perform co-immunoprecipitation experiments *in vitro* to detect an interaction between GARS1 and HDAC6: Benoy et al. identify an interaction between HDAC6 and both wild-type and C102R GARS1, whereas Mo et al. shows that only mutant GARS1, not wild-type, can interact with HDAC6. Mo et al. also find that the cells expressing the three *GARS1* mutants with the strongest HDAC6 interaction (P234KY, S581L, and G598A) also have the lowest amount of acetylated alpha-tubulin. This leads them to propose a mechanism in which mutant GARS1 aberrantly binds to HDAC6 and *increases* its de-acetylation activity, decreasing tubulin acetylation and impairing axonal transport.[370] However, this interaction between GARS1 and HDAC6 does not appear to be specific to pathogenic variants. One of the *GARS1* variants with the strongest effects on HDAC6 binding and alpha-tubulin deacetylation, S581L, has been re-evaluated and found to be

non-pathogenic, since it does not segregate with disease in affected families.[371] Therefore, it is unclear whether this interaction is meaningful for *GARS1*-mediated peripheral neuropathy. It is more likely that the benefits of HDAC6 inhibition are not specific to the genetic or environmental insult, but that improving axonal transport ameliorates the overall neuronal dysfunction in peripheral neuropathy. For example, HDAC6 inhibitors have been shown to improve peripheral neuropathy phenotypes in mice with *Mfn2* mutations,[372] *Hspb1* mutations,[373] and mice with chemotherapy-induced peripheral neuropathy.[374]

Lastly, there is a body of research investigating non-canonical functions of wild-type YARS1 in the nucleus. ARS enzymes can charge tRNA in the nucleus, as a proofreading mechanism to ensure that a tRNA molecule is properly spliced and folded before it is exported to the cytoplasm.[375] However, recent studies focus on a novel role for YARS1 in the nucleus. The nuclear localization signal (NLS) of YARS1 is found in the anticodon binding domain, and is masked by bound $tRNA^{Tyr}$.[328] Under oxidative stress conditions, $tRNA^{Tyr}$ is cleaved and the NLS is exposed, increasing YARS1 nuclear localization.[329] Through a combination of co-immunoprecipitation and mass spectrometry experiments, YARS1 was found to bind nuclear proteins TRIM28 and HDAC1.[329] Both of these proteins are transcriptional co-factors that work together to deacetylate the transcription factor E2F1, repressing its activity.[329] In this proposed model, YARS1 sequesters TRIM28 and HDAC6, preventing them from deacetylating E2F1, and increasing the function of E2F1 to upregulate its target genes, which include DNA damage repair genes.[329] Interestingly, when YARS1 is excluded from the nucleus by mutating the NLS, the expression of these DNA damage repairs decreases, supporting the nuclear role of YARS1 in driving their expression.[329]

Small angle X-ray scattering studies and hydrogen-deuterium exchange assays have defined protein conformation changes associated with three dominant *YARS1* mutations (E196K, G41R, and 153-156ΔVKQV).[355] This corresponds to increased interaction with TRIM28 and HDAC1, and subsequent increased E2F1 acetylation and expression of DNA damage response genes.[376] In this proposed mechanism, *YARS1* mutations *increase* the activity of a non-canonical YARS1 function (i.e., are hypermorphs). However, it is unclear how increased activity of the E2F1 transcription factor or increased expression of DNA repair genes relate to peripheral neuropathy

phenotypes. Indeed, pharmacologically repressing E2F1 activity does not improve the neurotoxicity of E196K *YARS1* in a *Drosophila* model.[376]

One of the most compelling pieces of evidence from these studies is that inhibiting the translocation of E196K YARS1 from the cytoplasm to the nucleus by mutating the NLS does in fact appear to rescue numerous neurological phenotypes in *Drosophila*, including locomotion defects and aberrant neuromuscular junction morphology.[376] There are a number of possible explanations for this observation. The authors, Bervoets et al., propose that this indicates a gain-of-function mechanism, in which mutant YARS1 aberrantly binds to transcription factors in the nucleus and misregulates gene expression. Alternately, it is possible that in this model, YARS1 is interfering with the nuclear proofreading function of drosophila tyrosyl-tRNA synthetase, and that incompletely processed tRNAs are exiting the nucleus and impairing translation in the cytoplasm. This hypothesis is supported by transcriptomic data of misregulated genes in the brains of *Drosophila* over-expressing either wild-type *YARS1* or E196K *YARS1*. Here, many of the differentially expressed genes play a role in stress response, protein misfolding, and ribosome biogenesis, indicating a broader defect in protein translation.[376] These defects may not be sufficient to cause a detectable neurological phenotype in flies expressing wild-type *YARS1,* but may be exacerbated by E196K *YARS1*. It is also important to consider that over-expressing human *YARS1* (wild-type and mutant) in a *Drosophila* model may produce spurious protein-protein interactions or cellular pathology that is not representative of *YARS1* biology in a human peripheral nerve, where it is expressed at endogenous levels.

Although the finding of novel binding interactions with NRP1 receptors, Trk receptors, HDAC6, and nuclear transcription co-factors could yield new insights into the pathogenesis of individual ARS mutations, it would be surprising if any of these mechanisms were shared across neuropathy-associated mutations in different ARS loci. The structures of the five neuropathy-associated ARS enzymes differ significantly; YARS1 and WARS1 are Class I enzymes, whereas GARS1, AARS1, and HARS1 are Class II enzymes (Table 1.1). Furthermore, pathogenic mutations do not localize to a specific domain (Figure 1.4), so it is unlikely that they would all have the capacity to bind to the same proteins. If different mutant ARS enzymes were to aberrantly bind to different proteins that act in a common pathway, it is possible that this may

explain the shared pathogenic effect; however, this has not been shown. Such a common pathway may be related to neuronal signaling or axonal transport, as discussed above, or may be related to protein translation independent of deficits in aminoacylation. Interestingly, the latter possibility could provide an explanation for the translation defects observed in *Drosophila* models of *GARS1* and *YARS1* mutants. When several neuropathy-associated mutations in human *GARS1* (E71G, G240R, and G526R) or *YARS1* (G41R, 153-156delVKQV, and E196K) are over-expressed in *Drosophila* motor or sensory neurons, they reduce protein translation rates and cause muscle denervation and morphological defects.[306] However, this study concluded that mutant GARS1 does not impair the endogenous activity of *Drosophila gars1*, and that the reduced translation rate caused by over-expressing G240R human *GARS1* cannot be rescued by over-expressing wild-type *Drosophila gars1*. These findings are particularly interesting since, if the translation defects in flies over-expressing mutant GARS1 are not a result of mutant GARS1 suppressing the endogenous protein via a dominant-negative effect, it is possible that they are caused by aberrant interactions between GARS1 mutant proteins and other components of the translational machinery.

1.3.6 Future directions for defining a mechanism of dominant ARS disease

There is currently evidence to support multiple proposed mechanisms of ARS-mediated peripheral neuropathy; however, additional research is needed to determine if either mechanism applies to the majority of neuropathy-associated ARS mutations and loci. For the loss-of-function model, it will be critical to determine if dimerization is required for pathogenicity. This question can be addressed by designing ways to decrease the dimerization of a pathogenic ARS protein and testing for phenotypic rescue in a relevant model organism (see Chapter 4), or by increasing dimerization and testing for an exacerbation of the phenotype. It will also be important to determine which dominant ARS mutations cause phenotypes that can be rescued by over-expression of the wild-type allele, which would suggest an overall loss of ARS function is central to the pathology.

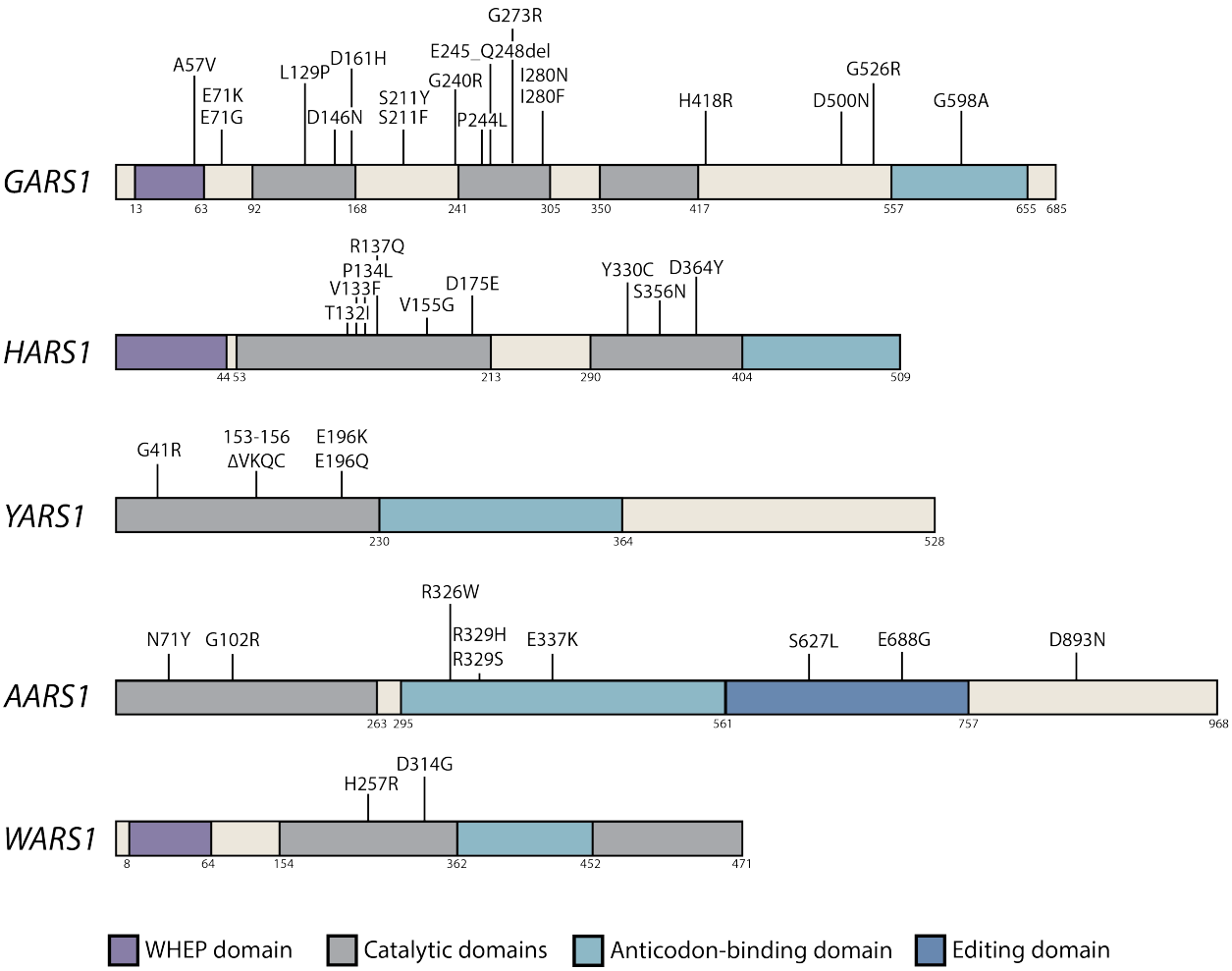


Figure 1.4. Cartoon map of ARS variants associated with dominant peripheral neuropathy.

Cartoon illustrations of the protein domains of *GARS1, HARS1, YARS1, AARS1,* and *WARS1* are shown. The WHEP domain is shown in purple, the catalytic domains in gray, the anticodon-binding domain in light blue, and the editing domain for *AARS1* in dark blue. Patient variants with evidence for pathogenicity are shown using vertical lines to indicate where they map on the linear protein domains.

One argument for a dominant-negative mechanism is that only homodimeric ARS have been implicated in disease to date. It will be important to determine if mutations in monomeric enzymes, such as MARS1, can cause dominant axonal neuropathy, since a dominant-negative mechanism would not be possible for monomeric enzymes. Identifying *MARS1* variants that segregate with disease in large pedigrees or demonstrating that *MARS1* variants cause dominant peripheral neuropathy in animal models would be important steps toward resolving this issue.

For the gain-of-function model, it will be important to show that any novel protein-protein interactions—whether with NRP1, TRIM28, Trk receptors, HDAC6, or other proteins—are specific to mutations associated with neuropathy, and that these interactions do not occur with nonpathogenic protein variants. Additionally, showing that multiple mutant ARS enzymes can participate in the same aberrant interaction, or different aberrant interactions that lead to the same cellular effect, will add weight to this model. Finally, demonstrating that mutations in other components of these pathways also lead to peripheral neuropathy, which has not been shown yet, would be a strong confirmation of this mechanism.

After refining the loss- and gain-of-function models, the next step will be to determine if there is any interplay between the two mechanisms that may affect phenotypic outcome. For example, some mutations, like G598A[162] and ΔETAQ *GARS1*[170] are linked to an early-onset, severe spinal muscular atrophy, which may be due to the compound effects of loss-of-function and gain-of-function mechanisms.

1.4 Conclusions

ARS genes are emerging as a significant cause of rare inherited diseases, including recessive mitochondrial disorders, recessive multisystem disorders, and dominant axonal neuropathies. All of the 37 human ARS enzymes have been implicated in a genetic disease phenotype. However, the full phenotypic and allelic spectrum of these disorders is undefined. In particular, it is unknown if additional ARS genes can cause dominant peripheral neuropathy, or what the mechanism of dominant disease is for *AARS1, GARS1, HARS1, WARS1,* and *YARS1* mutations. In this dissertation, I will present and functionally characterize ARS alleles that have been newly

identified in patients with multi-system recessive disorders or dominant peripheral neuropathies (Chapter 2). These efforts will broaden our understanding of the allelic spectrum of these diseases, as well as shed light on how these enzymes function and tolerate variation. I will also present a model organism-based prediction pipeline to implicate additional ARS in dominant and recessive phenotypes, which I apply to threonyl-tRNA synthetase (*TARS1*). These studies resulted in the identification of a hypomorphic *TARS1* allele and characterization of recessive *TARS1* phenotypes in worm and mouse (Chapter 3). Finally, I will address the mechanism of dominant ARS-mediated disease by showing that dominant pathogenic variants in alanyl-tRNA synthetase (*AARS1*) exert a dominant-negative when co-expressed with wild-type *AARS1* in yeast (Chapter 4). This work aims to contribute to the known clinical, locus, and allelic spectrum ARS-mediated disease; expand the role of model organism research in predicting pathogenic ARS alleles and investigating related phenotypes; and advance efforts to test a dominant-negative mechanism of ARS-mediated peripheral neuropathy.

Chapter 2
Expanding the Locus, Allelic, and Phenotypic Heterogeneity of ARS-Mediated Disease

2.1 Introduction

All 37 members of the ARS gene family have been implicated in either dominant or recessive human genetic diseases. Bi-allelic variants in 36 ARS genes have been found in patients with multisystem recessive phenotypes, which frequently includes central nervous system pathology.[68,377] These patients are either homozygous for hypomorphic missense alleles, compound heterozygous for two hypomorphic alleles, or compound heterozygous for one hypomorph and one null; complete loss of ARS function is not compatible with life.[68] Based on the patient genotypes and functional studies of these variants, a loss-of-function mechanism is responsible for these recessive disorders.

Mono-allelic variants in 5 ARS genes have been confidently implicated in dominant phenotypes, and all cause similar dominant axonal peripheral neuropathies. Here, these variants comprise missense mutations or small in-frame deletions. While these variants impair ARS function *in vitro* or *in vivo*,[68] haploinsufficiency for these genes does not appear to cause peripheral neuropathy (as discussed in Section 1.3.4). A pathogenic mechanism encompassing both loss of enzymatic function and dominant toxicity is currently under investigation (Chapter 4).

For recessive ARS-mediated disorders, there is broad clinical and allelic heterogeneity that remains to be fully defined. For dominant ARS-mediated disease, the full spectrum of ARS loci and variants that can cause dominant peripheral neuropathy is not known. The studies presented in this Chapter contribute to resolving these questions by evaluating previously uncharacterized

alleles in five ARS genes for a role in recessive or dominant disease based on genetic and functional evidence.

One of the strongest forms of evidence for the pathogenicity of a variant is co-segregation with disease in a large pedigree, with known genotypes of affected and unaffected individuals.[378] The first pathogenic variants in aminoacyl-tRNA synthetases were identified through this approach, using co-segregation of disease with chromosomal markers (linkage analysis) in large pedigrees, followed by targeted gene sequencing to identify the precise genetic lesion.[106,160,275,379] However, as next-generation sequencing has become a more accessible diagnostic tool, more aminoacyl-tRNA synthetase variants have been found in small families or as *de novo* events. The number of ARS genes known to cause disease has also grown, from just 7 in 2010,[75,79,106,160,195,241,275] to all 37 members of the gene family in 2021.[144,221,222,377] One major challenge facing the ARS research and clinical community is the interpretation of novel patient variants as benign or pathogenic, particularly in the absence of sufficient genetic data.

There are multiple forms of genetic evidence that can provide robust evidence for pathogenicity. If a variant co-segregates with disease in a small pedigree comprising only a few genotyped individuals, it is possible that this co-segregation is due to chance. However, if the variant co-segregates with disease in a large, multi-generational pedigree with many genotyped individuals, it becomes highly unlikely that such segregation is due to chance,[378] and provides compelling evidence that the variant is responsible for the phenotype. For example, the variant R329H *AARS1* was initially found in a large French family with 23 affected individuals, 17 of whom were genotyped (6 unaffected individuals were also genotyped).[106] This data provides strong genetic evidence that R329H is a pathogenic allele.

Another form of genetic evidence for pathogenicity is variant enrichment in affected populations, compared to unaffected control populations. Overall, ARS-mediated disease is too rare to have identified many affected individuals with the same variant. Again, R329H *AARS1* is an illustrative example. Since it was initially reported by Latour et al. in 2010, it has been identified in an additional 9 families on at least 3 independent haplotypes, bringing the total published number of individuals with R329H *AARS1*-mediated CMT to at least 46 (this conservative

Table 2.1. R329H *AARS1* is a recurrent pathogenic allele.

Family or Cohort Study	Number of affected, genotyped individuals	Ethnicity/Geographic Location	Reference
Family	17	French	Latour et al. 2010
Family	3	French	Latour et al. 2010
Family	8	Australian	McLaughlin et al. 2012
Family	3	UK	Bansagi et al. 2015
Family	1	UK	Bansagi et al. 2015
Family	2	UK	Bansagi et al. 2015
Family	2	UK	Bansagi et al. 2015
Family	2	Ireland	Bansagi et al. 2015
Cohort Study	1	Mediterranean	Bacquet et al. 2018
Cohort Study	1	Mediterranean	Bacquet et al. 2018
Family	3	Canary Islands	Lousa et al. 2019
Family	2	Korean	Lee et al. 2020
Family	3	United States	This study

estimate considers that two individuals identified in a French cohort study may have been previously identified by Latour et al.) (Table 2.1). It is also necessary to consider the frequency of R329H *AARS1* in unaffected populations.

One accessible tool for this work is gnomAD, a database of variants that have been identified in large-scale sequencing projects, along with the variant's allelic frequency. This data excludes individuals with severe pediatric disease and their first-degree relatives, providing a glimpse of allelic variation in all other individuals.[334] However, it does not exclude individuals with adult-onset diseases such as ARS-mediated peripheral neuropathy, making it an imperfect control population for these analyses. Despite these limitations, the fact that R329H is not present in gnomAD indicates that it is an extremely rare allele. This suggests that its identification in 11 families with Charcot-Marie-Tooth disease is unlikely to be due to chance.

A final piece of genetic information that can be used to build an argument for or against pathogenicity is conservation analysis. Aminoacyl-tRNA synthetases are found in all unicellular and multi-cellular organisms. If an amino acid is invariant across multiple different biologically diverse organisms, it is presumed to be important to enzyme function; any amino acid changes that are deleterious would have undergone purifying selection. As an example, the R329 residue in *AARS1* is conserved between humans and bacteria,[107] suggesting that this arginine is important for enzyme function, and changing it to a different amino acid may be deleterious.

Investigating the conservation and population frequency of an ARS variant are important initial steps in assessing the evidence for its pathogenicity. However, in the absence of substantial genetic evidence (co-segregation in families and/or recurrence in multiple affected individuals), it is still difficult to determine its contribution to disease. Here, it can be useful to assess the impact of the variant on gene function, either through *in vitro* aminoacylation assays or *in vivo* models. If a variant has a similar effect as a high-confidence, known pathogenic variant, this lends support to an argument for pathogenicity.

One way to determine if an ARS variant impacts enzyme function is to express recombinant human ARS in bacteria, purify it, and perform an *in vitro* aminoacylation assay. Here, wild-type

or mutant ARS enzyme is incubated with its substrates: ATP, tRNA and (radiolabeled) amino acid.[297] The first step of the reaction—in which the ARS enzyme binds ATP and the amino acid, and releases pyrophosphate—can be measured by addition of a colorimetric reagent that recognizing free pyrophosphate.[298] To investigate the complete aminoacylation reaction, the amount of charged tRNA can be measured; at various time points within the first 5 minutes of the reaction, aliquots of the mixture are spotted on filter paper soaked with trichloroacetic acid (TCA).[297] After the tRNA precipitates are washed and dried, the amount of radiolabeled amino acid that been ligated to the tRNA is measured with a scintillation counter.[297] These assays should be performed according to Michaelis-Menten conditions, and data should be fit to the Michaelis-Menten equation, in order to accurately detect defects in aminoacylation.[297]

Aminoacylation assays for high-confidence pathogenic ARS alleles demonstrate that these variants reduce enzymatic function. For example, R329H *AARS1* shows 1/50th of the enzymatic activity of wild-type *AARS1*.[107] Other high-confidence variants in *GARS1* and *YARS1*, all of which were identified through linkage analysis,[160,296,380,275,379,381] show similar defects. L129P and G240R *GARS1* both significantly reduce enzymatic activity,[300,371] as does a high-confidence *YARS1* variant, G41R.[275,382] In the case of another high-confidence *YARS1* variant, E196K, there is conflicting evidence between reduced activity in a pyrophosphate release assay[275] and no effect on enzyme activity in a steady state aminoacylation assay.[382] E196K does impair gene function in yeast complementation assays[275,383]; these discrepancies highlight the importance of using multiple functional assays to build a consensus on the variant's effect.

Yeast complementation assays test the ability of an ARS variant to support yeast growth in the absence of the endogenous yeast ARS gene. Yeast viability is maintained by expressing the wild-type yeast gene from a maintenance vector bearing a *URA3* auxotrophic marker. Then, the pathogenic ARS variant is transformed into the strain (either modeled in the yeast gene or the human open reading frame). Yeast are plated on media containing 5-fluoroorotic acid (5FOA), which selects for spontaneous loss of the maintenance vector,[299] and yeast growth is supported solely by the function of the wild-type or mutant ARS under investigation. An ARS variant that reduces yeast growth compared to the wild-type allele is defined as a loss-of-function allele. This assay has been validated with high-confidence pathogenic dominant alleles. For example, R329H

AARS1 causes a loss-of-function in yeast complementation assays, either when modeled in the yeast ortholog *ALA1*[107] or human *AARS1* (Section 4.3.1). The pathogenic *YARS1* alleles G41R and E196K also reduce yeast growth in complementation assays when modeled in either the yeast ortholog *TYS1*[275] or human *YARS1*,[383] as does the pathogenic *GARS1* allele L129P in the yeast ortholog *GRS1*[301], and the pathogenic *GARS1* allele G240R in the human gene (Figure 1.1).

In vitro aminoacylation assays and yeast complementation assays can also be employed to investigate recessive ARS variants. Here, the *KARS1* variant R505H is a useful illustration. This variant has been identified in six individuals with leukoencephalopathy and hearing loss, either in the homozygous state[191] or *in trans* to other *KARS1* missense alleles (P533S[72,187] or A526V[190]. Although it has not been identified in any large pedigrees, this degree of recurrence in a small patient population with a consistent phenotype provides strong evidence of pathogenicity. R505H *KARS1* is found in gnomAD at an allele count of 1 in 251,414,[334] indicating that it is a rare allele. Conservation analysis shows that it is invariant between humans, fruit flies, and yeast.[72] Functional assessment using *in vitro* aminoacylation assays demonstrate that it has 1/20th the activity of wild-type KARS1, and when modeled in the yeast ortholog *KRS1*, it causes a mild reduction of yeast growth[72]. Both *in vitro* and *in vivo* evidence is consistent with R505H *KARS1* partially reducing *KARS1* function as a hypomorph. Because R505H *KARS1* has been found in a homozygous individual with leukoencephalopathy,[191] it is unsurprising that it retains partial function, since complete loss of function would be incompatible with life. Consistent with this, it has been identified *in trans* with P533S in four individuals,[72,187] a variant with 1/200th of wild-type activity in aminoacylation assays.[72] It is likely that the partial enzymatic function of R505H KARS1 significantly contributes to the viability of these patients.

There are limitations to *in vitro* aminoacylation assays and yeast complementation assays. Although they provide clues about a variant's effect on gene or enzyme function, neither assay is a model of the human tissue affected in ARS-related diseases, and neither can provide evidence of the variant's impact in that biological context. This is especially true for hypomorphic alleles, where the deleterious effects of the variant may be modulated by the cellular environment, such as amino acid availability, tRNA levels, and protein translation demands.

A major limitation of all experimental approaches that detect a loss-of-function effect is that, while impaired enzyme function explains the recessive disease mechanism, it is only one component of the dominant mechanism. In addition to reducing enzyme function, dominant ARS mutations are also toxic, particularly to long neurons. As discussed in Section 1.3.4 and 1.3.5, this may be related to the decreased enzyme function through a dominant-negative mechanism, or could be through a neomorphic gain-of-function mechanism (or some combination of the two). It is currently unclear what differentiates a toxic loss-of-function mutation from a non-toxic loss-of-function mutation. Testing patient variants for a dominant toxic effect in neurons currently requires over-expressing ARS variants in flies,[384] worms,[174] or fish[110,350] then assessing motor behavior or neuron morphology in these organisms. Alternately, mouse models can be generated and evaluated for peripheral neuropathy phenotypes[170], which is even more expensive and time consuming. An inexpensive and rapid assay in an easily manipulated model, such as yeast, is required to effectively test large numbers of patient variants for dominant toxicity. This is discussed further in Chapters 4 and 5.

All data from functional assays must be evaluated in the context of the genetic evidence for pathogenicity, and ideally in the context of other functional assays. Although each approach has limitations, validating these assays with high-confidence pathogenic ARS variants provides a framework with which to investigate newly identified patient variants of uncertain significance. In this Chapter, I will present six studies that describe previously unreported ARS variants identified in patients, evaluate their impact on gene function, and assess their contribution to disease pathology. These studies span both dominant and recessive diseases, as well as variants in both established and candidate ARS disease genes. Such efforts contribute to building a catalog of pathogenic ARS variants, which helps define the full genetic and phenotypic spectrum of these diseases. They also contribute to defining disease mechanisms and understanding how mutations in the translation machinery affect cellular health and tissue function.

Parts of this chapter are adapted from: "A recurrent *GARS* mutation causes distal hereditary motor neuropathy" published in *Journal of the Peripheral Nervous System* (Volume 24, Issue 4, pages 320-323, October 19 2019, License Number 5114850026750); "Substrate interaction defects in histidyl-tRNA synthetase linked to dominant axonal peripheral neuropathy," published

in *Human Mutation* (Volume 39, Issue 3, pages 415-432, December 26 2017, License Number 5114850191607); and "*MARS* variant associated with both recessive interstitial lung and liver disease and dominant Charcot-Marie-Tooth disease," published in *European Journal of Medical Genetics* (Volume 61, Issue 10, pages 616-620, October 2018). For the latter, I retain the right to include this Elsevier article in a thesis or dissertation. Permission was requested for reproduction of data through the Copyright Clearance Center (see license numbers above). Additionally, the Contracts, Rights, and Permissions Coordinator from IOS press granted use of materials from "A novel mutation in *MARS* in a Patient with Charcot-Marie-Tooth Disease, Axonal, Type 2U with Congenital Onset," published in the *Journal of Neuromuscular Diseases* (Volume 6, Issue 3, pages 333-339, July 22 2019). The author performed all the studies in this chapter with the following exceptions: the clinical evaluations, phenotypic classifications, and diagnostic sequencing was performed by clinical collaborators; Jamie Abbott performed the aminoacylation assays for *HARS1* variants (Table 2.2). Christina del Greco assisted with cloning R131H, V372I, R619C, Q639P, and R663Q *TARS1,* and performed the yeast complementation assays for these variants (Figure 2.6).

2.2 Materials and methods

2.2.1 Identification of patient variants

ARS variants were identified in patients using a CMT gene panel (G327R *GARS1* Patient 1*,* Y330C *HARS1*) or whole-exome sequencing (G327R *GARS1* Patient 2, V155G *HARS1*, S356N *HARS1*, A397T *MARS1,* R619C *MARS1,* Y307C *MARS1*, ΔM236 *NARS1,* S461F *NARS1*, C342Y *NARS1,* R663Q *TARS1*, R619C *TARS1,* and Q639P *TARS1*). The *TARS1* variants R131H and V327I were identified by a combination of homozygosity mapping and whole-exome sequencing. For the R619C/Y307C *MARS1* patient, compound heterozygous variants in *TG* and *VPS13C* were also identified. The ΔM236 *NARS1* patient was also heterozygous for a variant in *SQSTM1*. No other additional variants were reported.

2.2.2 ClustalW alignments

Multiple-species alignments were generated using Clustal Omega software.[385] Protein sequences were obtained from the NCBI Protein Database. For glycyl-tRNA synthetase, the accession numbers were: BAA06338.1 (*H. sapiens*), AAH21747.1 (*M. musculus*), CCD72866.1 (*C. elegans*), KZV13199.1 (*S. cerevisiae*), AUP44736.1 (*E. coli*). For histidyl-tRNA synthetase, the accession numbers were: NP_002100.2 (*H. sapiens)*, NP_032240.3 (*M. musculus*), AAB38116.1 (*C. elegans*), QHB12288.1 (*S. cerevisiae*), B1XAY9.1 (*E. coli*). For methionyl-tRNA synthetase, the accession numbers were: NP_004981.2 (*H. sapiens)*, NP_001165053.1 (*M. musculus*), CAA97803.1 (*C. elegans*), CAA97293.1 (*S. cerevisiae*), A7ZNT3.1 (*E. coli*). For asparaginyl-tRNA synthetase, the accession numbers were: NP_004530.1 (*H. sapiens*), NP_001136422.2 (*M. musculus*), CAA95808.1 (*C. elegans*), GHM92552.1 (*S. cerevisiae*), BAL38068.1 (*E. coli*). For threonyl-tRNA synthetase, the accession numbers were: AAH10578.2 (*H. sapiens*), AAH55371.1 (*M. musculus*), CAA93762.1 (*C. elegans*), P04801.2 (*S. cerevisiae*), P0A8M3.1 (*E. coli*).

2.2.3 Generation of ARS expression constructs

For all yeast complementation assays, the human ARS proteins were expressed from a *ADH1* promoter on the gateway-compatible pYY1 construct,[159,386] which harbors a 2-micron origin of replication resulting in a high plasmid copy number per cell and bears a *LEU2* auxotrophic marker. The open reading frames (ORFs) of wild-type human *MARS1, HARS1*, *NARS1*, and *TARS1* were amplified from HeLA or fibroblast cDNA with using primers containing the attB1 and attB2 recombination sequences (primer sequences in Appendix A). The *GARS1* ORF was previously amplified with attB-containing primers from the original pYY1 vector.[159,386] All PCR-generated ORFs were purified with Qiagen Spin Miniprep columns and recombined into pDONR221 using the Gateway cloning technology (Invitrogen) BP reaction. The reaction was transformed into Top10 cells (Invitrogen). Clones were isolated and grown in media containing kanamycin to select for the presence of the kanamycin resistance cassette on pDONR221. Plasmid DNA was purified from these clones using the Qiagen Spin Miniprep kit, then digested with *Bsr*GI to identify clones with successful insertion. These clones were sequence-verified to ensure there were no PCR amplification errors. Then, the QuickChange II XL Site-Directed Mutagenesis Kit (Agilent) was used to introduce patient variants into the appropriate wild-type

ORF (primer sequences in Appendix A). The reaction was transformed into Top10 cells (Invitrogen) and grown in kanamycin-containing LB to select for the presence of the kanamycin-resistance gene on the pDONR221 plasmid. The plasmids were extracted and purified for Sanger sequencing using the Qiagen Miniprep kit. Sequencing ensured that the desired mutation was generated and that there were no additional off-target mutations. Then, the Gateway LR reaction was performed to recombine wild-type and mutant ORFs into the pYY1 vector. The reaction was transformed into Top10 cells, which were grown in ampicillin-containing LB to select for the presence of the ampicillin-resistance gene on the pYY1 vector. The plasmids were then extracted, purified, and digested with *Bsr*GI (New England Biolabs) to identify recombinants.

2.2.4 Yeast complementation assays

Yeast complementation assays were performed using strains with the yeast ARS ortholog deleted (Δ*HTS1*,[175] Δ*MES1*,[216] Δ*GRS1*,[387] Δ*DED81* (Horizon Discovery, Clone ID 20982), and Δ*THS1* (Horizon Discovery, Clone ID 21471). These strains maintain viability by carrying a pRS316 vector that expresses the yeast ARS gene from its endogenous promoter, contains a yeast centromere sequence that results in a low copy number per cell, and bears the auxotrophic marker *URA3*. The pYY1 vector (expressing the human ARS ORF or an empty control) was introduced into the appropriate strain using lithium acetate yeast transformation, performed at 30°C with 200ng of plasmid. Transformed yeast were grown on solid media lacking uracil and leucine to select for cells with both the pRS316 maintenance vector and the pYY1 experimental vector. After 3 days of growth at 30°C, individual colonies were picked into 2mL liquid media lacking uracil and leucine. They were grown for 2 days at 30°C, shaking at 275 rpm until saturated. 1 mL of saturated culture was then centrifuged at 15,000 rpm for 1 minute and re-suspended in 50µl water. Serial 1:10 dilutions were made to 1:100 or 1:1,000, and then 10µl of each diluted mix (including undiluted) was spotted on plates containing 0.1% 5-fluoroorotic acid (5-FOA) (Teknova). Yeast growth was visually assessed after 3 to 5 days. In each experiment, at least two independently generated clones were tested for each mutation.

2.2.5 HARS1 aminoacylation assays

Aminoacylation assays were performed using a modified version of the previously described Uhlenbeck-Wolfson assay.[388,389] Multiple turnover experiments were conducted in buffer comprised of 50 mM HEPES pH 7.5, 150 KCl, 10 mM $MgCl_2$, 5 mM β-ME, 2 U/ml PPase and ^{32}P-labeled $tRNA^{His}$. A fixed concentration of enzyme (5 nM for WT HARS1 and 20 nM for CMT variants) was used with saturating concentrations of two of the three substrates. The saturating concentrations of $tRNA^{His}$, ATP, and histidine were 5 μM, 10 mM, and 5 mM, respectively. The variable concentrations of these substrates were 100 nM to 15 μM, 25 μM to 5 mM, and 1 μM to 5 mM, for $tRNA^{His}$, ATP and histidine, respectively. Experiments with mutant HARS1 typically required a 2-3 fold higher final concentration of the variable substrate than the wild-type HARS1. To calculate the initial rates, the reactions were typically sampled during the first minute. After the reaction was terminated with quenching buffer (400 mM NaOAc [pH 4.5] and 0.1% SDS), the tRNA was digested to single nucleotides with 0.1ug of P1 nuclease (Sigma) for 40-60 minutes at room temperature. Radiolabeled aminoacylated A76 was separated from non-aminoacylated A76 by thin-layer chromatogragraphy on PEI-cellulose plates (Scientific Adsorbents) with a mobile phase of 0.1M ammonium acetate and 5% acetic acid. Radioisotopic imaging on a phosphor screen (Bio-Rad Molecular Imager FX) was used to detect radioactive products. The concentration of aminoacylated $tRNA^{His}$ was quantified by comparing the ratio of the relative amount of aminoacylated A76 to total radiolabeled product.

2.3 Results

2.3.1 A recurrent GARS1 *mutation causes distal hereditary motor neuropathy.*

Two patients with a similar peripheral neuropathy were examined by neurologists at the University of Iowa or the University of Pennsylvania, respectively. Both patients were young women with an adolescent onset of distal weakness in the legs and arms. Clinical electrophysiology showed decreased motor axon (but not sensory axon) conduction velocities, indicating a distal motor neuropathy. Whole-exome or gene panel sequencing was performed, which revealed a c.979G>A variant in both (presumably) unrelated patients that changes glycine at amino acid number 327 to arginine (G327R). In both patients, the variant likely arose *de novo*; the unaffected mother and father of Patient 1 did not carry the mutation, nor did the unaffected

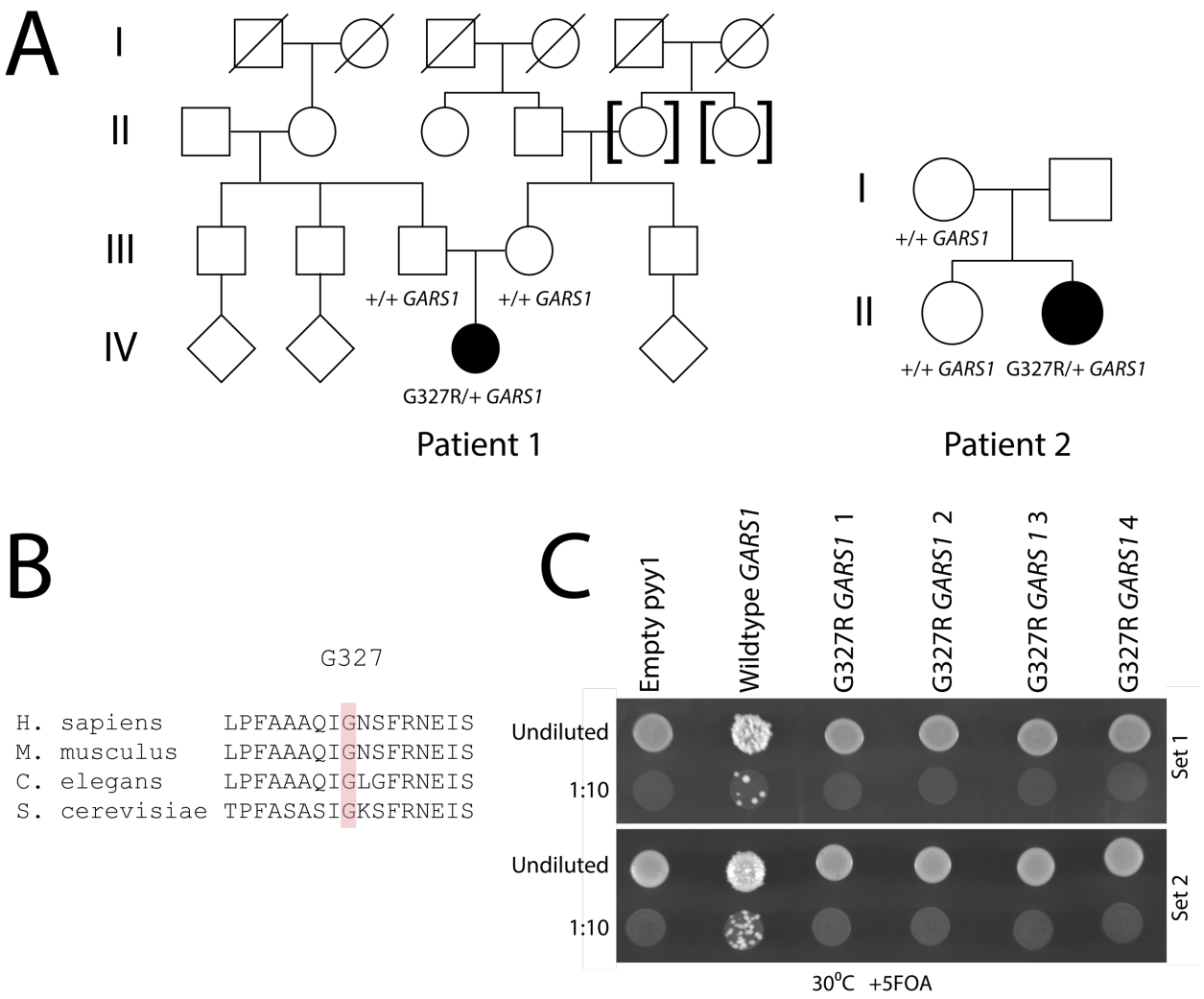


Figure 2.1. Identification and characterization of G327R *GARS1*.

(A) Pedigrees of patients 1 and 2. The individuals in brackets were adopted into the family. **(B)** Conservation analysis of the affected residue. The position of G327 GARS1 is shown with flanking amino acid acids from evolutionarily diverse species. G327 is highlighted in pink. **(C)** Representative images of haploid yeast lacking the endogenous *GRS1* transformed with an empty construct ("Empty pyy1"), wild-type human ΔMTSΔWHEP *GARS1*, or an expression construct with G327R ΔMTSΔWHEP *GARS1*. Four independently generated G327R *GARS1* expression clones were tested (across the top, 1-4) and two independent colonies were evaluated per transformation (Set 1 and Set 2). Each colony was plated undiluted or diluted on media containing 5-FOA. Yeast were grown at 30°C.

Table 2.2. gnomAD allele counts of characterized ARS variants.

Gene	Variant	Allele Count	Allele Number	Allele Frequency	Number of homozygotes
GARS1	G327R	Not present	N/A	N/A	N/A
HARS1	V155G	Not present	N/A	N/A	N/A
	Y330C	Not present	N/A	N/A	N/A
	S356N	6	282,884	2.12E-05	0
MARS1	A397T	Not present	N/A	N/A	N/A
	Y307C	2	251,470	7.95E-06	0
	R618C	3	282,892	1.06E-05	0
NARS1	ΔM236	Not present	N/A	N/A	N/A
	C342Y	Not present	N/A	N/A	N/A
	S461F	Not present	N/A	N/A	N/A
TARS1	R663Q	Not present	N/A	N/A	N/A
	R131H	31	281,744	1.10E-04	0
	V372I	49	251,318	1.95E-04	0
	R619C	Not present	N/A	N/A	N/A
	Q639P	Not present	N/A	N/A	N/A

N/A, or Not Applicable, indicates that the allele could not be identified in gnomAD.

mother and sister of Patient 2 (the father was unavailable, but was not known to have a neuropathy) (Figure 2.1A). G327R is absent in gnomAD (Table 2.2) and affects a glycine residue conserved between humans, mice, worms, and yeast (Figure 2.1B). Therefore, it was prioritized for functional evaluation using a yeast complementation assay. This assay demonstrated that wild-type human *GARS1* can support formation of yeast colonies in the absence of the endogenous yeast ortholog *GRS1* (Figure 2.1C). However, G327R *GARS* did not rescue the deletion of *GRS1*, indicated by an absence of yeast growth; these data strongly suggest that G327R impairs gene function (Figure 2.1C). This is consistent with the loss-of-function pattern observed in pathogenic *GARS1* variants from patients with peripheral neuropathies,[170,171,301,371] and strengthens the argument that G327R *GARS* is a pathogenic allele. In sum, G327R *GARS1* adds to the increasing catalog of loss-of-function missense *GARS1* mutations linked to dominant peripheral neuropathies.

2.3.2 Newly identified HARS1 *variants in patients with dominant peripheral neuropathy*

Mutations in *HARS1* have been found to segregate with dominant peripheral neuropathy in several multi-generational pedigrees[174]; however, only a small number of pathogenic *HARS1* mutations have been described.[174,175]. Here, we discuss three newly identified *HARS1* variants in three families with a motor-predominant axonal neuropathy. In the first family, the variant V155G was present in five individuals with motor neuropathy, and absent in one unaffected individual (Figure 2.2A). In the second family, the variant Y330C was present in two individuals with motor-predominant peripheral neuropathy, but absent in two unaffected individuals (Figure 2.2A). The third family comprised an affected daughter with a mild to moderate motor-predominant peripheral neuropathy, who was heterozygous for S356N *HARS1*. Notably, her unaffected mother was also heterozygous for S356N, indicating that this variant either does not cause a fully penetrant phenotype or is not pathogenic (Figure 2.2A). Neither V155G nor Y330C are present in gnomAD (Table 2.2). S356N is present at a low frequency, with an allele count of 6/282,884. All three affected amino acids are evolutionarily conserved from human to mouse, worm, and yeast; Y330 is also conserved in bacteria (Figure 2.2B) This suggests that these variants may be deleterious to enzymatic function.

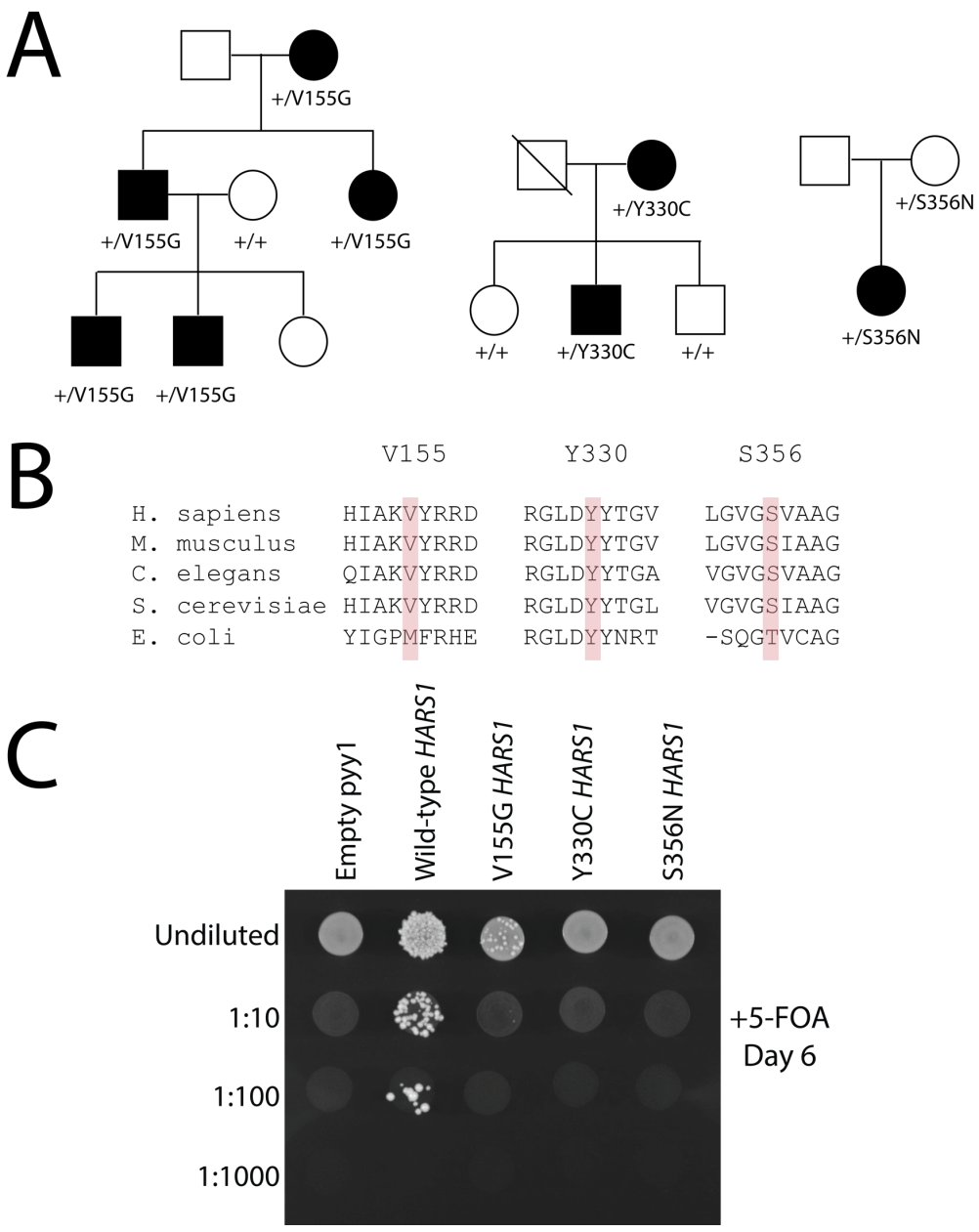


Figure 2.2. Identification and characterization of V155G, Y330C, and S356N *HARS1*.

(A) Pedigrees of three identified families with *HARS1* variants. A diagonal line indicates the individual is deceased. Genotypes are shown for individuals with available DNA samples. **(B)** Conservation analysis of V155, Y330, and S356N. Each residue is shown in pink, with flanking amino acids from evolutionarily diverse species. **(C)** Yeast complementation assay for V155G, Y330C, and S356N. Representative images from haploid yeast lacking *HTS1*, transformed with either an empty vector ("Empty pyy1"), wild-type *HARS1*, V155G *HARS1*, Y330C *HARS1*, or S356N *HARS1*. Yeast were spotted in 1:10 dilutions on media containing 5-FOA, and were grown at 30°C.

To test this possibility, *in vitro* aminoacylation assays were performed. Here, to carefully assess the effect of each individual mutant on substrate recognition, assays were performed with a fixed concentration of enzyme and varying concentrations of substrate. The initial velocity of the reaction and substrate concentration were fit to the Michaelis Menten equation, which provided estimates of K_m (the concentration of the substrate which permits the enzyme to reach half of its maximum velocity, or V_{max}) and k_{cat} (the turnover number). From this, k_{cat}/K_m was calculated, a value that represents the overall efficiency of the enzyme in both binding the substrate and performing the reaction. In the first reaction, the tRNA concentration was varied, while ATP and histidine concentrations were held at saturating concentration. None of the three HARS1 mutants showed increased K_m values for tRNA (Table 2.3). In the second reaction, histidine concentration was varied. Here, Y330C and V155G exhibited a 25-fold and 85-fold increase in K_m value, respectively, indicating these mutations required significantly increased histidine concentration to reach half of V_{max}, and suggesting a decreased affinity for histidine relative to the catalytic turnover (Table 2.3). In the third reaction, ATP concentration was varied. Here, the K_m for Y330C increased 40-fold, and maximum velocities could not be reached for S356N and V155G even at saturating concentrations (Table 2.3). These data suggest that a major contributor to the decrease in the catalytic activity for these enzymes comes from a significant increase in K_m for ATP. In all three experiments, each HARS1 mutation decreased substrate turnover (k_{cat}), as well as the overall catalytic efficiency (k_{cat}/K_m) (Table 2.3).

To determine how these *in vitro* deficiencies impact cellular function *in vivo*, the three mutations were tested for the ability to rescue yeast growth when the endogenous yeast ortholog, *HTS1*, was deleted. Wild-type *HARS1* was able to support colony formation, indicating that human *HARS1* can at least partially complement loss of *HTS1* (Figure 2.2C). In contrast, V155G *HARS1* lead to significantly reduced yeast growth, and neither Y330C nor S356N supported any growth (Figure 2.2C). These data are consistent with all three mutations reducing gene function. In this assay, Y330C and S356N *HARS* are functionally null alleles, while V155G is hypomorphic, with enough gene function to support limited colony formation.

In sum, the above functional studies indicate that V155G, Y330C, and S356N impair *HARS1* function. For V155G and Y330C, the combination of the functional data and the segregation

Table 2.3. Steady state kinetics of $tRNA^{His}$ aminoacylation by wild-type and neuropathy-associated mutations in HARS1.

	Variable Substrate								
	tRNA			**histidine**			**ATP**		
	K_m (μM)	**k_{cat} (s^{-1})**	**k_{cat}/K_m ($\mu M^{-1} s^{-1}$)**	**K_m (μM)**	**k_{cat} (s^{-1})**	**k_{cat}/K_m ($\mu M^{-1} s^{-1}$)**	**K_m (μM)**	**k_{cat} (s^{-1})**	**k_{cat}/K_m ($\mu M^{-1} s^{-1}$)**
WT HARS	0.782 ± 0.101	5.4 ± 0.2	6.9	**8.0 ± 4.0**	**4.1 ± 0.4**	**0.5 ± 0.4**	44.2 ± 5.5	5.8 ± 0.2	0.13
S356N HARS	0.199 ± 0.059	0.30 ± 0.02	1.5	10.8 ± 9.1	0.39 ± 0.04	0.04	*ND*	*ND*	$1.5x10^{-4} \pm 1.2x10^{-5}$
Y330C HARS	0.330 ±0.072	0.56 ± 0.03	1.7	202.9 ± 83.23	0.74 ± 0.10	0.004	1,763 ± 544	0.48	0.000272
V155G HARS	0.979±0.159	3.05 ± 0.14	3.1	687.2 ± 200.0	2.08 ± 0.16	0.003	ND	ND	$3.8x10^{-4} \pm 2.7x10^{-5}$

Values reported are the mean ± standard error of three independent experiments.

with disease in the affected families provides strong evidence that these variants are pathogenic. Although S356N clearly affects HARS1 function, the minimal genetic data (*i.e.*, the incomplete penetrance within a small pedigree) makes it difficult to know whether S356N is pathogenic. Importantly, in the absence of convincing genetic evidence, a loss-of-function effect is not sufficient to determine pathogenicity, because loss of function is likely necessary but not sufficient for an ARS mutation to cause dominant peripheral neuropathy.[68] To further inform the interpretation of this variant, further investigation is required in an informative model system that can reveal dominant, neurotoxic properties of ARS mutations.

2.3.3 A MARS1 *variant of uncertain significance in Charcot-Marie-Tooth disease*

Here, we evaluate a *MARS1* variant found in an 11-year-old girl who presented with a severe motor peripheral neuropathy that began in her first year of life. Exome sequencing was performed, and a heterozygous *MARS1* variant, A397T, was identified. The patient's unaffected mother did not carry the mutation, and the father was not available for testing (Figure 2.3A). *MARS1* variants have been previously described in patients with adolescent or adult-onset dominant peripheral neuropathy.[216–218] The genetic evidence is weak for all of these cases, although one variant has been functionally evaluated in a yeast complementation assay and reduces gene function when tested in the yeast ortholog *MES1*[216]. Based on the motor neuropathy presentation of the A397T/+ proband, the fact that A397 is conserved to yeast (Figure 2.3B), and the absence of A397T from gnomAD (Table 2.2), this variant was prioritized for further study. To determine if this variant reduced gene function similar to known pathogenic dominant variants in other ARS genes, it was assessed for a loss-of-function effect in a yeast complementation assay.

Here, we performed a yeast complementation assay with the human *MARS1* coding sequence. Wild-type human *MARS1* supported robust yeast growth in the absence of the yeast gene *MES1* (Figure 2.3C). However, four independent clones of A397T *MARS1* only lead to sporadic colony formation, suggesting that this mutation impairs gene function (Figure 2.3C).

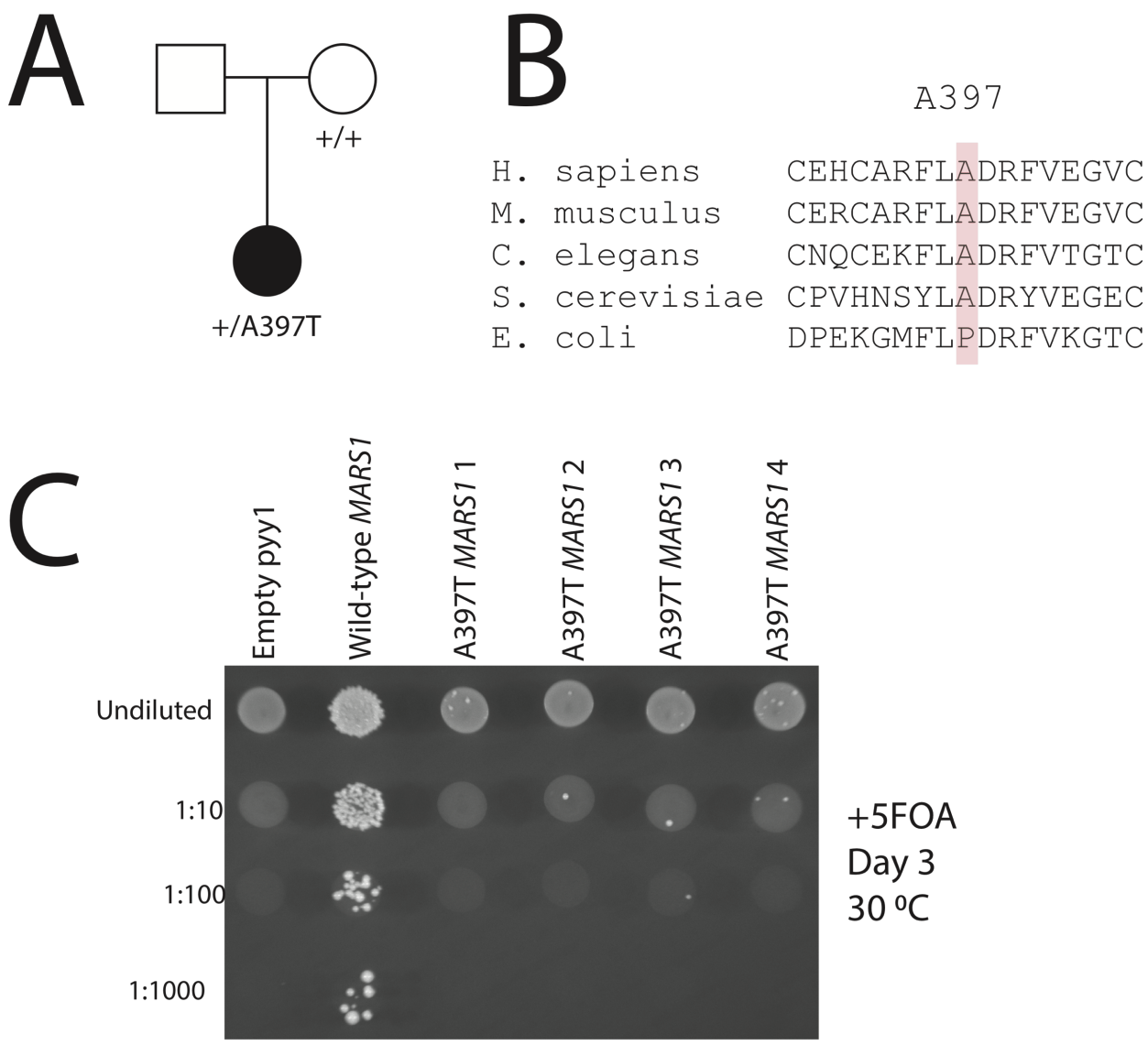


Figure 2.3. Identification and characterization of A397T *MARS1*.

(A) Pedigree information of an individual heterozygous for A397T *MARS1*. **(B)** Conservation analysis of A397, shown in pink. Surrounding amino acids from evolutionarily diverse species are shown. **(C)** Representative images from haploid yeast lacking *MES1*, transformed with either an empty vector ("Empty pyy1"), wild-type *MARS1*, or four independently generated A397T *MARS1* constructs. Yeast were spotted in 1:10 serial dilutions on media containing 5-FOA, and were grown at 30°C.

2.3.4 Identification and characterization of MARS1 *variants in patients with recessive disease*

Previous studies have shown that bi-allelic *MARS1* variants cause a severe recessive disorder predominantly characterized by interstitial lung and liver disease; to date, a total of 12 mutations have been found across 36 patients.[210–212,214,215,390] We identified a case of *MARS1*-mediated recessive disease in a male infant who presented with transfusion-dependent anemia, hypothyroidism, cholestasis, developmental delay, and interstitial lung disease. Through whole-exome sequencing, the patient was found to be compound heterozygous for two *MARS1* variants; a Y330C variant was inherited from his mother and, interestingly, the same R618C mutation that was previously identified in a small pedigree with CMT disease[216] was inherited from his father (Figure 2.4A). The R618 residue is conserved to *E. coli*; in contrast, Y330 is conserved in mouse and worm, but not in yeast or bacteria (Figure 2.4B). R618C is found in 3/282892 alleles in gnomAD, and Y330C is found in 2/251470 alleles (Table 2.2).

To determine the functional consequences of R618C and Y307C MARS1, a yeast complementation assay was performed, modeling these variants in the human open-reading frame as previously described. Consistent with previous data studying R618C in the yeast gene *MES1*, R618C *MARS1* failed to support yeast growth, supporting its loss-of-function designation. Y307C *MARS1* lead to yeast growth comparable to that of wild-type *MARS1*, indicating that in this assay it does not substantially impair function (Figure 2.4C). Considering that a complete absence of MARS1 function should be incompatible with life, it is expected that at least one of two *MARS1* alleles retain some function. Due to the phenotypic similarities of this patient and other *MARS1*-mediated recessive disease, it is likely that Y307C is a pathogenic allele, and that the yeast complementation assay does not have sufficient resolution to detect mildly hypomorphic variants.

2.3.5 NARS1 is a candidate gene for dominant peripheral neuropathy

Here, we present data implicating an additional ARS gene—*NARS1*—in dominant peripheral neuropathy. NARS1 is the sixth dimeric[49], cytoplasmic ARS to be mutated in patients with dominant peripheral neuropathy. We identify two large families with cases of dominant axonal

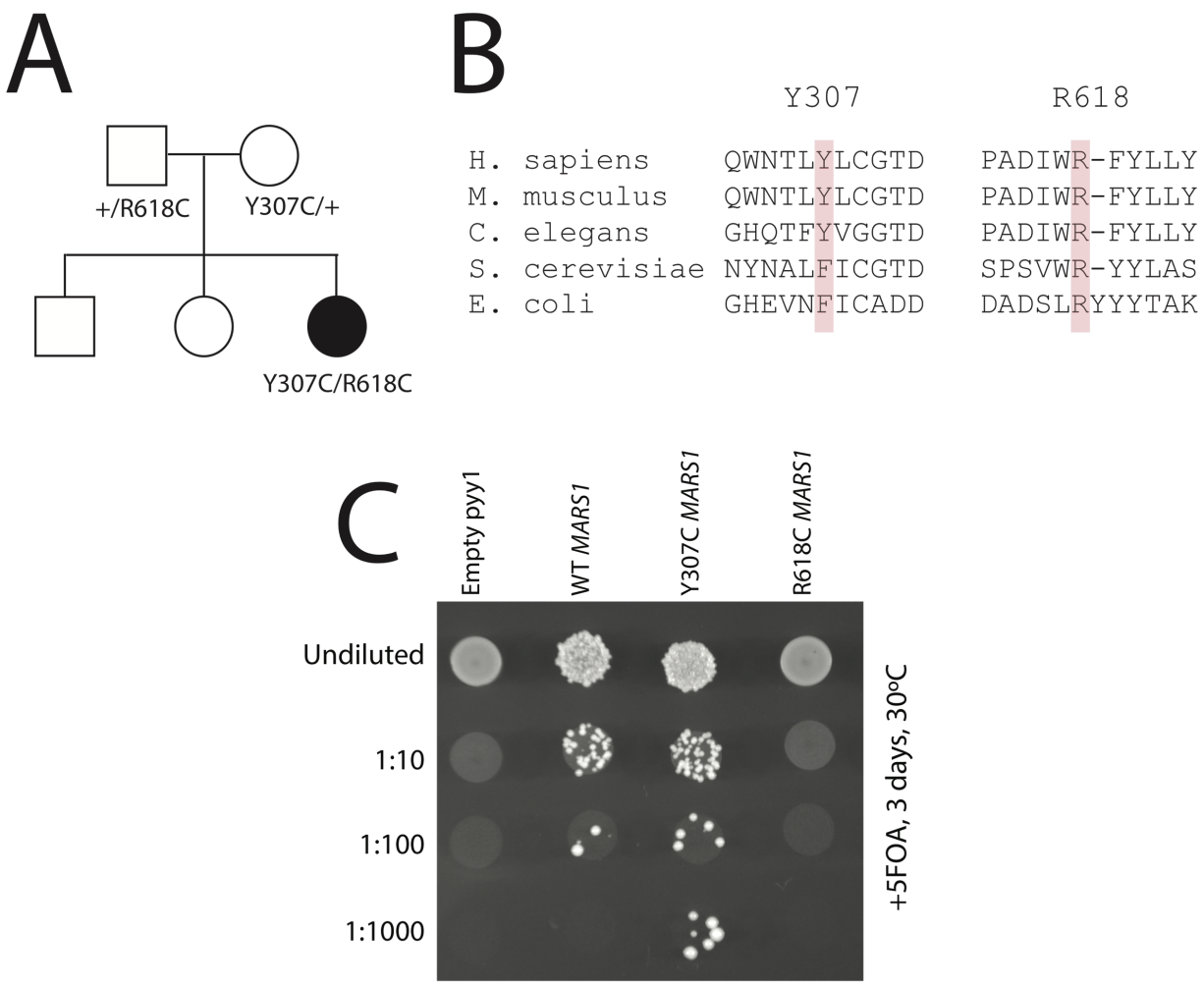


Figure 2.4. Identification and characterization of R618C and Y307C *MARS1*.

(A) Pedigree displaying the inheritance pattern of R618C and Y307C *MARS1*. **(B)** Conservation analysis of *MARS1* variants. Y307 and R618 are shown in pink, surrounded by flanking amino acid sequences from evolutionarily diverse species. **(C)** Representative images from haploid yeast lacking *MES1*, transformed with either an empty vector ("Empty pyy1"), wild-type *MARS1*, Y307C *MARS1*, or R618C *MARS1*. Yeast were spotted in 1:10 serial dilutions on media containing 5-FOA, and were grown at 30°C.

neuropathy. In the first family (Family A), the variant S461F *NARS1* co-segregates with disease in five affected individuals (three unaffected individuals were genotyped and do not carry this allele) (Figure 2.5A). In the second family (Family B), the variant C342Y *NARS* was identified in three affected individuals (six unaffected individuals were genotyped and did not carry this allele) (Figure 2.5A). Finally, one case of a *de novo* in-frame deletion (ΔM236 *NARS1*) was found in a patient who had sensorimotor neuropathy in addition to cerebellar ataxia, cognitive impairment, and abnormal eye movements (Figure 2.5A). None of these three variants are found in gnomAD (Table 2.2), indicating that they are rare. All three affect conserved residues: S461 *NARS1* is conserved between humans, mouse, worm, and yeast; C342 *NARS1* is conserved between humans, mouse, and worms; M236 *NARS1* is conserved between humans and mouse (Figure 2.5B)

A yeast complementation assay was performed to assess the consequence of these variants on gene function. Wild-type human *NARS1* supported yeast growth in the absence of the yeast ortholog *DED81*. However, none of the three patient *NARS1* variants (S461F, C324Y, or ΔM236) lead to colony formation, indicating that all three are loss-of-function alleles in this assay (Figure 2.5C). These data are consistent with the pattern of loss-of-function mutations found in other ARS-mediated dominant neuropathies. Together with the pedigree information for S461F and C342Y, these experiments suggest that *NARS1* is a candidate gene for dominant peripheral neuropathy. However, because *NARS1* has not been previously implicated in this phenotype, additional models to test for neurotoxic dominant properties are required to more confidently link it to disease. In this case, determining if one of the *NARS1* variants with stronger genetic evidence (S461F or C342Y) can cause peripheral neuropathy in a mouse model would be the strongest evidence to implicate *NARS1* in dominant disease.

2.3.6 Expanding the phenotypic spectrum of TARS1-mediated recessive disease

Bi-allelic loss-of-function mutations in threonyl-tRNA synthetase (*TARS1*) were recently identified in two patients with triochothiodystrophy (TTD)[253], a rare neuro-ectodermal condition marked in part by brittle hair.[391] One of these two patients presented with the characteristic TTD "tiger-tail" hair banding pattern, along with ichthyosis, follicular keratosis, delayed physical

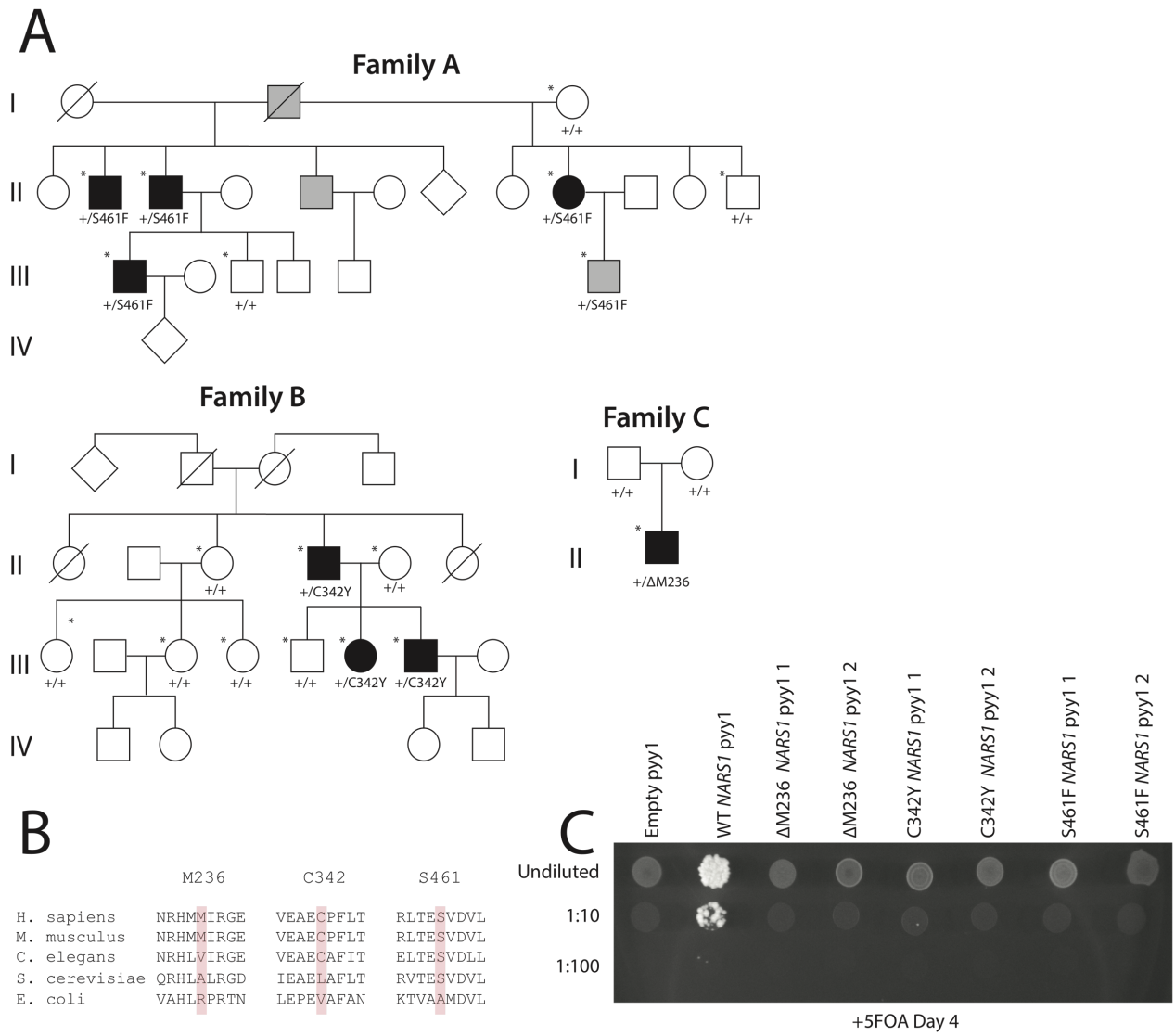


Figure 2.5. Identification and characterization of *NARS1* variants in individuals with dominant peripheral neuropathy.

(A) Pedigrees of three families with *NARS1* variants. Hearsay cases of peripheral neuropathy are marked in gray. Clinically evaluated patients are marked with an asterisk. In Family A, the Generation III individual marked with gray is oligosymptomatic. **(B)** Conservation analysis of M236, C324, and S461, marked in pink. Flanking residues from evolutionarily diverse species are shown. **(C)** Representative images from haploid yeast lacking the yeast ortholog *DED81*, transformed with an empty vector ("Empty pyy1"), wild-type *NARS1*, ΔM236 *NARS1,* C342Y *NARS1*, or S461F *NARS1*. Yeast were serially diluted on media containing 5-FOA, and were grown at 30°C.

development, acromandibular dysplasia, and recurrent respiratory tract infections.[253] The other patient also had hair with the "tiger-tail" banding pattern, was born a collodion baby encased in a shiny membrane, and had ichthyosis.[253] Functional evidence from yeast complementation assays and primary fibroblasts convincingly demonstrated that these mutations significantly impair TARS1 function and stability.[253]

However, these are not the only phenotypes associated with reduced *TARS1* function. Here, we report three additional patients with bi-allelic *TARS1* mutations. These patients all present with infantile-onset or childhood-onset severe neurological symptoms, including seizures, microcephaly, intellectual disability, spasticity, and ataxia. They also have musculoskeletal and craniofacial dysmorphism. Although neurological symptoms are the most prominent aspect of the presentations, there are some skin and hair phenotypes reminiscent of TTD: the proband from Family 3 has thin and fragile hair, as well as translucent and easily scratched skin. The proband from Family 2 is noted to have sparse eyebrows.

Six *TARS1* variants were identified in these patients. Proband 1 is a Pakistani individual born to consanguineous parents who is homozygous for a complex allele containing two variants, R131H and V372I *TARS1* (Figure 2.6A). Interestingly, these variants are both present in gnomAD at a low frequency in the South Asian population (Table 2.2). Proband 2 is an Egyptian individual who is compound heterozygous for R619C *TARS1* (inherited from their father) and Q639P *TARS1* (inherited from their mother) (Figure 2.6A). Neither of these variants are present in gnomAD (Table 2.2). Proband 3 is homozygous for R663Q *TARS1* (Figure 2.6A), which is also not found in gnomAD (Table 2.2). The older sibling of Proband 3 was stillborn at full term with multiple abnormalities. Although there is no known consanguinity in the Family 3, they are noted to come from a small Italian town. All six *TARS1* variants are conserved across diverse species; R131H and R619C are conserved between human, mouse, worm, and yeast, and V372 and R663 are conserved between human, mouse, worms, yeast, and bacteria (Figure 2.6 B). In this analysis, Q639 is only conserved between human and mouse (*E.coli* also show a Q at this residue, although the surrounding region does not appear to have high conservation with mammalian species) (Figure 2.6B)

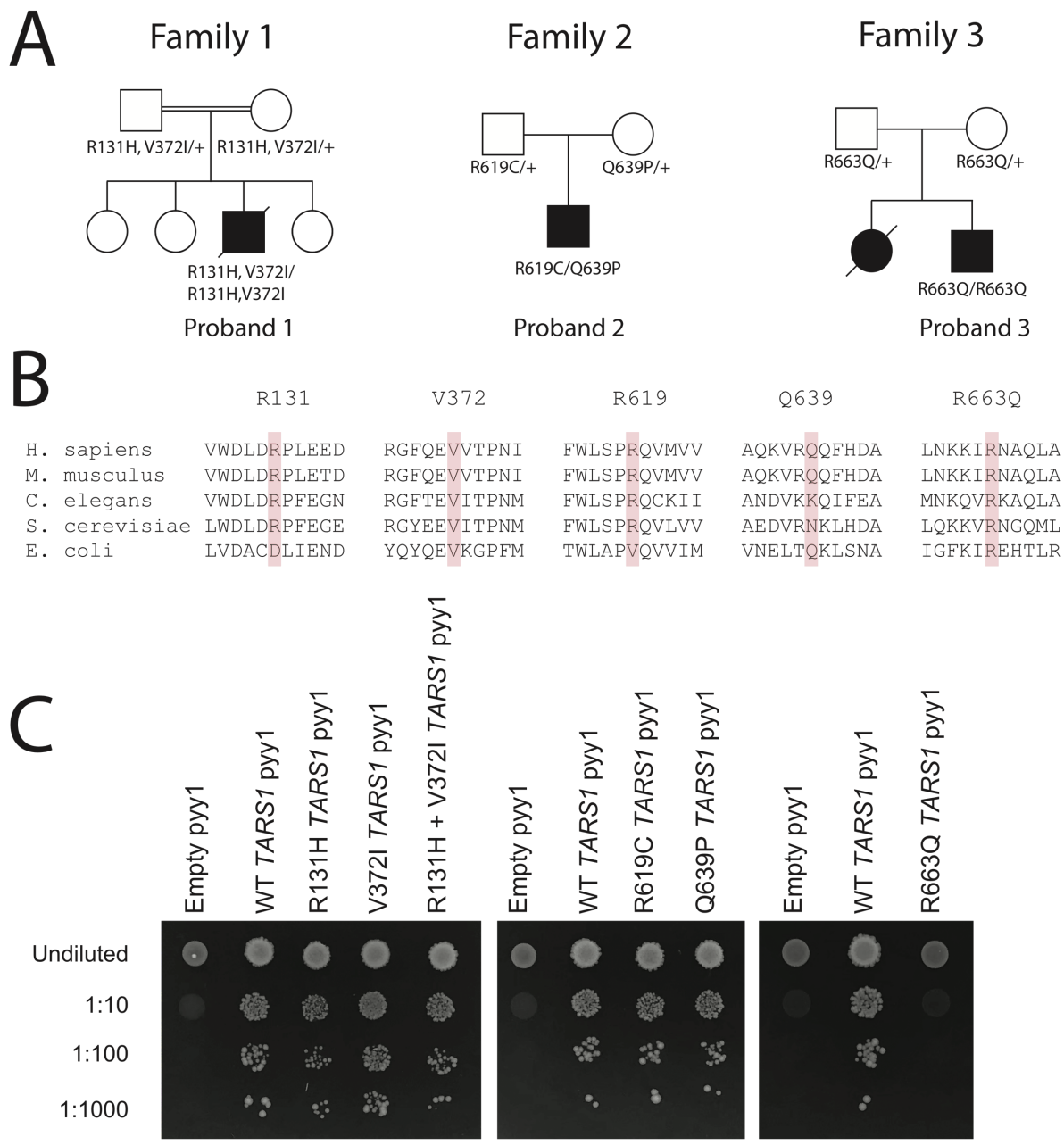


Figure 2.6. Identification and characterization of recessive *TARS1* variants.

(A) Pedigrees of three families with *TARS1* variants. **(B)** Conservation analysis of *TARS1* variants. R131, V372, R619, Q639, and R663 are all marked in pink. Each affected residue is shown with surrounding sequences of evolutionarily diverse species. **(C)** Representative images from haploid yeast lacking *THS1*, transformed with an empty vector ("Empty pyy1"), wild-type *TARS1*, or mutant *TARS1* (in the first panel, R131H and V372, tested independently and *in cis*; in the second panel, R619C and Q639P; in the third panel, R663Q). Yeast were spotted on media containing 5-FOA in serial 1:10 dilutions, and were grown at 30°C.

To assess the functional consequence of these variants, we performed a yeast complementation assay in a strain with the endogenous threonyl-tRNA synthetase gene, *THS1*, deleted. Wild-type human *TARS1* supported yeast growth in the absence of *THS1* (Figure 2.6 C). R131H and V372I *TARS1* were tested individually and *in cis*, to determine if one variant or both variants together impacted TARS1 function. Interestingly, when tested individually and *in cis*, R131H and V372I *TARS1* supported yeast growth comparable to wild-type *TARS1* (Figure 2.6 C). A similar result was obtained for both R619C and Q639P *TARS1*, which were identified *in trans*; neither variant significantly reduced yeast growth (Figure 2.6 C). One possibility is that this yeast complementation assay does not have the resolution to detect mild to moderate defects in TARS1 function, and that a more sensitive assay (for example, an *in vitro* aminoacylation assay) is required to determine if they impact enzyme activity.

In contrast, the R663Q *TARS1* mutation, which was found in the homozygous state, does not support any yeast growth, indicating that it significantly impairs TARS1 function (Figure 2.6 C). In light of what is known about the relationship between impaired ARS function and severe multisystem disorders, and previous work demonstrating the role of reduced *TARS1* function in disease[253], these data indicates that R663Q *TARS1* is likely a pathogenic allele. This case serves as a useful reminder that the function of a human ARS in a yeast system may not directly translate to the function in an affected patient cell—although this mutation does not support any yeast growth, a complete loss of function in a human patient would be incompatible with life. Therefore, it is probable that R663Q *TARS1* is a pathogenic mutation that significantly impairs, but does not ablate, TARS1 function.

2.4 Discussion

Here, we present the identification and functional characterization of 15 previously unreported ARS variants across 4 studies of dominant disease and 2 studies of recessive disease. These efforts contribute to expanding the known genetic and phenotypic heterogeneity of ARS-mediated disease and to defining the features of pathogenic ARS variants.

Two of these studies examine novel patient variants in two well-characterized ARS disease genes, *HARS1* and *GARS1*. In these cases, the patients' presentations—adolescent to adult-onset dominant peripheral neuropathy with mild to moderate severity—is consistent with previously described *HARS1*- or *GARS1*-mediated disease. Although the G327R *GARS1* variant arose *de novo*, its reoccurrence in two individuals with similar phenotypes provides strong genetic evidence of pathogenicity. For *HARS1,* two of the three variants—Y330C and V155G *HARS1*—segregate with disease in the patient's families, although one (S356N *HARS1*) shows incomplete penetrance. Additionally, all four of these novel mutations impair gene function, consistent with previous reports of pathogenic *GARS1* and *HARS1* mutations.[170,174,175,371] In total, these data support the pathogenicity of G327R *GARS1*, Y330C *HARS1*, and V155G *HARS1*, whereas S356N *HARS1* requires further investigation.

The *MARS1* cases described here provide an example of both recessive and dominant phenotypes associated with a single gene. The patient with recessive developmental delay, cholestasis, and interstitial lung disease presents with a phenotype previously linked to bi-allelic loss-of-function *MARS1* mutations. For this individual, although only the R618C *MARS1* variant significantly impaired *MARS1* function in a yeast complementation assay, it is likely that Y307C *MARS1* is also a pathogenic allele based on the similarity of the phenotype with the known hallmarks of *MARS1*-mediated recessive disease. Of note, this patient inherited the R618C *MARS1* allele from his father. Although nerve conduction velocity studies were not available, the father was reportedly healthy. Explicit questioning did not reveal any family history of a peripheral neuropathy that might match the phenotype previously attributed to R618C. It is still possible that R618C causes a mild, late-onset phenotype with reduced penetrance; however, the absence of a family history of peripheral neuropathy contributes to the uncertainty surrounding R618C *MARS1* as a pathogenic variant.

A newly identified *MARS1* variant, A397T, was identified in an individual with early-onset dominant peripheral neuropathy. While our functional data are suggestive of pathogenicity, the minimal genetic information available for this variant makes it impossible to confidently implicate A397T in the patient's phenotype. There is also a significant disparity between the severe childhood-onset phenotype associated with A397T and the later-onset CMT phenotype

seen in patients with R618C, P800T, or R737W *MARS1*.[216–218] It is possible that differential *MARS1* impairment correlates with differential severity, or that environmental or genetic modifiers might lead to variable phenotypes. Alternately, it is possible that none of these *MARS1* variants are pathogenic. To date, no *MARS1* mutations have been shown to segregate with CMT in large, multigenerational families, nor have any been tested for a dominant neuropathy phenotype in a multicellular model system. Whether *MARS1* will ultimately be confirmed as a *bona-fide* CMT disease gene will have significant bearing on a proposed mechanism of disease, as it would be the first implicated ARS gene to function as a monomer.[392] This would rule out a dominant-negative effect, which is dependent on homo-dimerization, as a predominant unifying mechanism.

By this logic, the argument for a dominant-negative mechanism is strengthened by implicating yet another homodimeric, cytoplasmic enzyme in dominant disease. Here, we present data to support the argument that variants in *NARS1*—which encodes a homodimeric[49], cytoplasmic enzyme—causes dominant peripheral neuropathy. We identify three loss-of-function *NARS1* variants in individuals with dominant sensorimotor neuropathy. Two of these variants segregate with this phenotype in multigenerational pedigrees, which, combined with the functional data, builds a strong case for pathogenicity (the third variant is *de novo*).

Two recent studies have implicated *NARS1* in recessive disorders that include neurodevelopmental defects, seizures, and microcephaly.[221,222] Interestingly, the phenotypic spectrum of recessive *NARS1* disease also includes peripheral neuropathy,[222] indicating that the peripheral nervous system is sensitive to decreased NARS1 function; this also supports a dominant-negative mechanism for dominant *NARS1*-mediated disease. One of these studies also identified eight patients with mono-allelic *de novo NARS1* mutations, who presented with global developmental delay, intellectual disability, seizures, dysmorphic features, spasticity, ataxia, and microcephaly, as well as some cases of demyelinating neuropathy.[222] Six of these patients were heterozygous for a premature stop codon, R534* *NARS1*. The effect of this truncation on mRNA and protein stability was not thoroughly addressed, but the authors determined that NARS1 enzymatic activity in patient lymphoblast cells was reduced by approximately 80% compared to unaffected control cells.[222] Although the authors propose that a dominant-negative mechanism is

responsible, it is also possible that these patients have a non-coding *NARS1* variant *in trans*, which would not be detected by whole exome sequencing. This would also explain the phenotypic similarity between these patients and the other individuals with bi-allelic *NARS1* mutations presented in the same study.

The ataxia and intellectual impairment described in these patients may overlap with the phenotypes of the ΔM236/+ *NARS1* individual described here, who presents with both peripheral neuropathy and other central nervous system disorders. It is possible that the ΔM236/+ *NARS1* individual also has a non-coding variant *in trans* that was not detected with whole exome sequencing. It is also possible that this patient may have multiple genetic disorders, with a second unrelated variant accounting for the additional central nervous system phenotypes. Notably, this patient is heterozygous for the *SQSTM1* variant A33V, which has been associated with ALS and FTD.[393,394] Interestingly, bi-allelic variants in *SQSTM1* cause a neurodegenerative disorder including ataxia, cognitive decline, and gaze palsy.[395] This description matches the patient's additional phenotypes of cerebellar ataxia, cognitive impairment, and abnormal eye movement. Therefore, this patient may have an undetected non-coding variant in *SQSTM1 in trans* with A33V, and their multifaceted phenotype may be a result of both *SQSTM1* recessive disease and *NARS1* dominant disease.

We also have expanded the limited phenotypic spectrum of described *TARS1* phenotypes from trichothiodystrophy (TTD) to include a broader range of neurological symptoms. Here, the functional consequences of some patient variants are unclear—specifically, the complex allele containing R131H and V327I *TARS1*, and the individual variants R619C and Q639P *TARS1*. These require experimental approach with higher resolution than a yeast complementation assay. However, the R663Q *TARS1* variant clearly impairs *TARS1* function, and is likely the cause of the neurological and dysmorphic symptoms in a homozygous individual. It is also striking that this individual has thin and brittle hair, similar to the phenotype originally described in TTD patients. One possibility is that the hair and skin are less susceptible to mutations that mildly impair TARS1 (like the R131H and V372I allele or the R619C/Q639P variants *in trans*) than the central nervous system, but that more significant impairment affects both systems. Of note, the records of the previously reported TTD patients[253] were from early in life, and both individuals

were lost to follow-up. It is unknown whether these patients may have also developed more pronounced neurological symptoms. Moving forward, it will be critical to identify additional patients and develop appropriate model systems to understand the full phenotypic spectrum of *TARS1*-related disease, and to understand how reduced threonine-tRNA charging causes these symptoms.

This body of work has added to the known range of dominant and recessive disorders caused by mutations in aminoacyl-tRNA synthetase genes. It has also continued to demonstrate the relevance of testing patient variants for loss-of-function effects. However, caution is warranted when interpreting negative results from a yeast complementation assay given the limited resolution for hypomorphic alleles. Caution is also required when interpreting positive results from a yeast complementation assay, particularly for dominant peripheral neuropathy variants, as all loss-of-function mutations are not necessarily dominantly toxic to neurons. Overall, this work contributes to a building a large catalog of pathogenic ARS variants that will be an invaluable resource for research and clinical communities. It will also play an essential role in deciphering the mechanism of ARS-mediated disorders and understanding how mutations in these genes affect protein translation, cellular health, and tissue function.

Chapter 3

Designing Predictive Models to Assess Threonyl-tRNA Synthetase (*TARS1*) for a Role in Recessive and Dominant Phenotypes

3.1 Introduction

To date, five of the thirty-seven ARS genes (*AARS1, GARS1, HARS1, YARS1, WARS1*) have been confidently implicated in dominant peripheral neuropathy.[68] It is currently unknown if there are any unique characteristics of these five genes that might explain their role in this disease, or, if the pathology is due to a defect in tRNA charging, whether mutations in any other ARS gene could also cause dominant peripheral neuropathy. The only known shared characteristics of AARS1, GARS1, HARS1, YARS1, and WARS1 is that they perform their canonical role of tRNA charging in the cytoplasm, and that they all function as homodimeric enzymes (Table 1.1). This raises the possibility that mutations in any additional cytoplasmic, homodimeric ARS enzymes could also cause disease.

Previous research into the mechanism of pathogenic ARS mutations has yielded two important findings: 1) the majority of recessive and dominant ARS mutations cause a reduction of gene function *in vitro* or in yeast complementation assays (as discussed in Section 1.3.4) and 2) expression of dominant ARS mutations in multicellular organisms causes neuronal dysfunction, including morphological and behavioral defects. Dominant ARS variants have been studied in: *Drosophila* models of *GARS1* and *YARS1* mutations, where they cause neuromuscular junction defects and motor impairment[306,384]; zebrafish models of *HARS1* mutations, where they lead to shortened and misguided axons[350]; and *C. elegans* models of *HARS1* and *AARS1* mutations, which affect neuron morphology.[174,175,307] For example, when *HARS1* mutations are expressed in GABA-ergic neurons in worms, 40-50% of the worms have aberrant neuron projections from the

ventral nerve cord to the dorsal nerve cord, with a branching phenotype that suggests impaired axon guidance.[174] These worms also display a thinning of the dorsal nerve cord and impaired locomotion in a swimming assay. However, these models are limited in their ability to recapitulate the human peripheral neuropathy phenotype. In humans, the neuropathy usually progresses in a length-dependent manner, with the longest axons of the peripheral nervous system affected first. The neurons of flies, fish, and worms are relatively short, making it difficult to understand this length-dependent aspect of ARS toxicity. Additionally, these models rely on transgenic approaches that over-express ARS proteins. While this may help exacerbate a phenotype in shorter neurons, it does not accurately reflect the approximately equal expression of wild-type and mutant alleles in humans.

Mouse models can circumvent some of these limitations, and offer the benefit of an established mammalian system for studying human genetic diseases. There are currently three mouse models of dominant ARS disease, all caused by mutations in glycyl-tRNA synthetase (*Gars1*). Only one of these lines was designed to study a human variant—the ΔETAQ *GARS1* mutation, which causes a severe, early-onset SMA-like phenotype in humans.[170] *Gars*$^{\Delta ETAQ/+}$ mice develop neuromuscular junction (NMJ) defects by 6 weeks of age and motor defects by 12 weeks.[170]. The other two *Gars1* mouse models derive from spontaneous or ENU-induced mutagenesis. The spontaneous P234KY *Gars1* mutation causes severe neuromuscular dysfunction by 3 weeks of age and dramatically shortened lifespan.[69] The milder C201R *Gars1* mutation causes notable behavioral defects in grip strength by 1 month of age, although it does not affect lifespan.[347] Based on the early-onset phenotypes in these two lines and their similarity to the ΔETAQ model, it is likely that the human equivalent of P234KY *Gars1* and C102R *Gars1* would be more akin to the severe SMA-like phenotype caused by ΔETAQ. Although these mice lines are relevant resources for studying the early-onset form of *GARS1*-mediated neuropathy, they are not representative of the less severe, adult-onset form of ARS-mediated peripheral neuropathy.

Other genetic causes of dominant axonal peripheral neuropathy have been modeled in mice. These include mutations in the microtubule motor protein KIF1B-β,[396] the dynein heavy chain protein DHC,[397] the neurofilament protein NEFL,[398,399] and the chaperone protein HSPB1.[373] These mice show impaired neuromuscular function in a series of behavioral assays that include

rotarod tests, measurements of grip strength, and gait analysis. Although some of these mice display mild symptoms as early as three months of age,[397] others do not show overt signs of a neuropathy until 6 months,[373] 9 months,[399], or even 12 months of age.[396] It is likely that a mouse model of an ARS mutation that causes a later-onset, milder human phenotype would more closely resemble these models.

Currently, only five ARS genes—*AARS1, GARS1, HARS1, WARS1,* and *YARS1*— are confidently implicated in dominant peripheral neuropathy. It is currently unclear if there are special characteristics of these five genes that can enable them to cause dominant peripheral neuropathy, or if variants in any ARS could cause this disorder. The only known shared function between them is their canonical role of tRNA charging in the cytoplasm. Therefore, if there is a common mechanism of peripheral neuropathy, it is likely to be related to altered axoplasmic tRNA charging, which could also be caused by mutations in other cytoplasmic ARS genes. Additionally, these five ARS genes all encode homodimeric enzymes (Table 1.1), suggesting that homodimerization is important for disease pathology. Therefore, we predict that mutations in another homodimeric, cytoplasmic ARS could also cause dominant peripheral neuropathy.

We sought to test this hypothesis in the absence of convincing patient data, relying on the model organisms that have been previously used to characterize known dominant pathogenic variants in other ARS genes. These pathogenic variants impair gene function in yeast complementation assays,[174,175,307] and cause morphological and behavioral defects when over-expressed in *C. elegans* neurons.[174,175,307] We set out to identify similar variants in an ARS gene that is not yet implicated in dominant neuropathy and that encodes a homodimeric, cytoplasmic enzyme—threonyl-tRNA synthetase (*TARS1*)[56]. At the time, *TARS1* had not been implicated in dominant *or* recessive disease. This presented an additional opportunity to discover recessive variants in *TARS1* and define the related phenotypes.

To design *TARS1* variants that were likely to be deleterious and dominantly toxic, missense mutations were introduced at highly conserved residues. These were tested in yeast complementation models to identify those that reduced gene function; loss-of-function variants were prioritized for neuronal toxicity studies in *C. elegans* (Figure 3.1). A loss-of-function

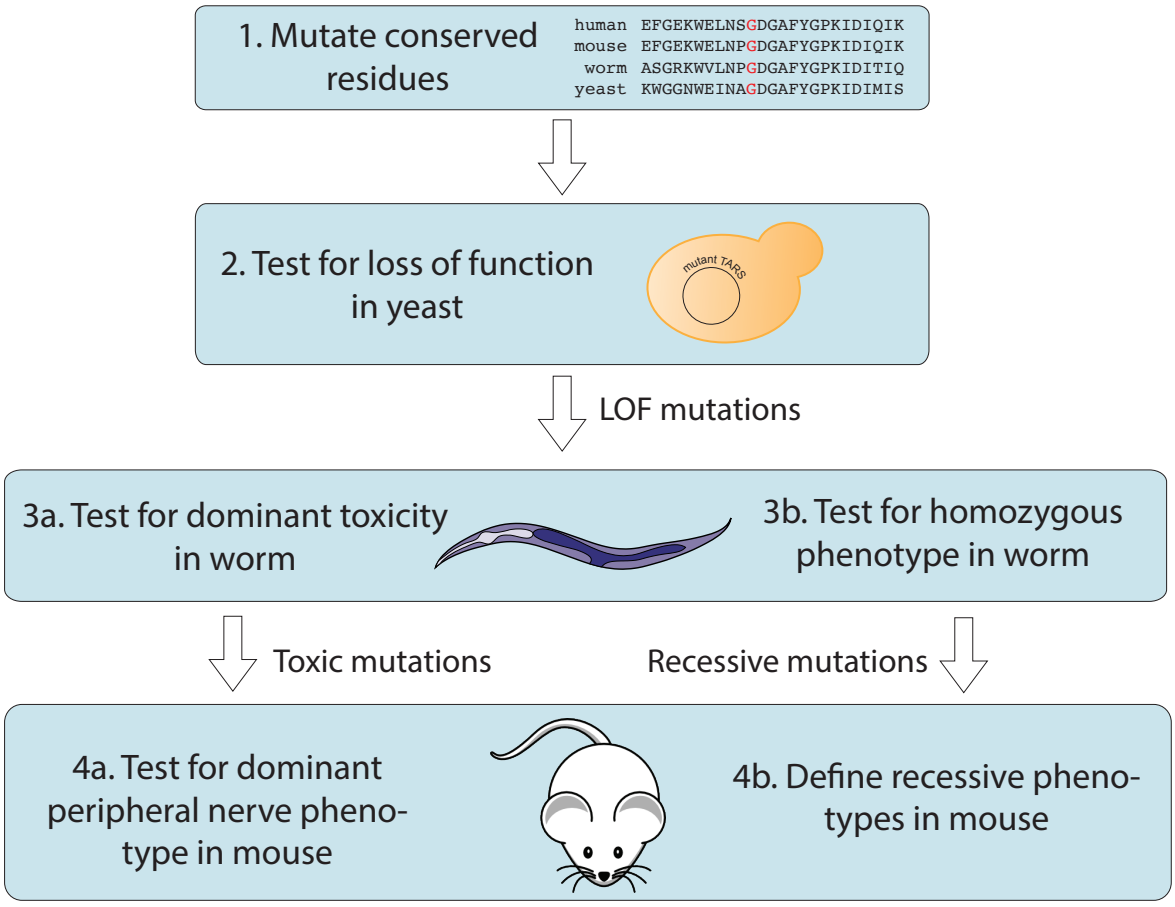


Figure 3.1. Flow-chart of the prediction pipeline designed to identify novel ARS candidates for dominant and recessive phenotypes.

mutation that was dominantly toxic to *C. elegans* neurons was then modeled in mouse to determine if it caused a peripheral neuropathy (Figure 3.1). Additionally, this pipeline was effective at identifying a novel hypomorphic *TARS1* allele, which lead us to a preliminary characterization of recessive *TARS1* phenotypes in worm and in mouse (Figure 3.1). Although *TARS1* has now been linked to a form of recessive trichothiodystrophy,[253] this mouse model provides an important resource to further explore the phenotypic spectrum of *TARS1*-related disease and to investigate whether reduced threonine-tRNA charging differentially affects mammalian tissues.

The studies presented in this Chapter represent a collaborative effort between the author and multiple mentors with expertise in *C. elegans* or mouse biology. The author of this chapter performed the yeast complementation assays (Figure 3.2), generation and characterization of *tars-1* knock in worms (Figure 3.4 through 3.7), and lead efforts to breed, genotype, and characterize mouse lines (Figure 3.8 through Figure 3.18). Stephanie Oprescu generated an initial panel of *THS1* variants that lead to the identification of the three candidate *TARS1* variants studied in this Chapter. Dr. Asim Beg generated and characterized worms over-expressing the *tars-1* alleles in GABA-ergic neurons (Figure 3.3). All mouse lines were generated by Dr. Thomas Saunders and the University of Michigan Transgenic Core, and guidance and assistance in mouse husbandry, genotyping, dissection, and behavioral analysis was provided by Dr. Miriam Meisler, Dr. Young Park, Dr. Guy Lenk, and Jennifer Pierluissi. Mouse behavioral testing was performed by the author, Dr. Young Park, and Steven Whitesall of the Molecular and Integrative Physiology Phenotyping Core. Analysis of mouse footprints was performed by the author, Dr. Young Park, and Jennifer Pierluissi. Nerve conduction studies were performed by John Hayes of the Michigan Mouse Metabolic Phenotyping Center. Preparation of histology sections was performed by Histoserv, Inc., and analysis of pup H&E and PAS histology was performed by Dr. Jerrold Ward. Dr. Marina Grachtchouk dissected adult mouse skin for histology, and Dr. Andrzej Dlugosz interpreted the pathology and provided guidance on analyzing the mouse hair phenotype. Matthew Pun and Molly Kuo assisted with the computational analysis of the mouse proteome.

3.2 Materials and methods

3.2.1 Generation of TARS1 *expression constructs*

The open reading frame (ORF) of *TARS1* was amplified from HeLa cell cDNA, using primers with the attB1 and attB2 gateway recombination sequences (primer sequences in Appendix A). These amplicons were purified with Qiagen Spin Miniprep columns and recombined into pDONR221 using Gateway cloning technology (Invitrogen). The recombination reaction was then transformed into Top10 cells (Invitrogen) to isolate clonal populations. Individual bacterial colonies were selected and grown in media containing kanamycin, which selected for the kanamycin resistance cassette on pDONR221. Plasmids were then isolated using the Qiagen Miniprep kit and genotyped by digesting with *Bsr*GI (New England Biolabs) to detect the presence of the *TARS1* insert. Clones with successful insertions were analyzed by Sanger sequencing to ensure absence of mutations introduced by amplification errors. To introduce variants into the *TARS1* ORF, site-directed mutagenesis was performed with the QuickChange II XL Site-Directed Mutagenesis Kit (Agilent) (primer sequences in Appendix A). The reaction was transformed into Top10 cells and grown in LB containing kanamycin to select for pDONR221. Plasmid DNA was isolated and sequenced as above, to ensure successful mutagenesis. Then, the Gateway LR reaction was used to recombine the wild-type or mutant *TARS1* into the vector pYY1. This vector has a 2-micron origin of replication, resulting in a high copy number per cell, as well as the *ADH1* promoter, resulting in strong constitutive *TARS1* expression. Recombinants were transformed into Top10 cells, which were plated on ampicillin to select for the ampicillin resistance cassette on pYY1. Then, plasmids were extracted, purified, and digested with *Bsr*GI to identify successfully recombined clones.

3.2.2 Yeast complementation assay

Yeast complementation assays were performed with the Δ*THS1* strain (Horizon Discovery, Clone ID 21471). Yeast viability was maintained with a pRS316 vector that expressed wild-type *THS1* from its endogenous promoter. pRS316 also carries the auxotrophic marker *URA3*, and has a yeast centromere sequence which results in a low copy number per cell. The pYY1 vector

(expressing wild-type *TARS1,* mutant *TARS1*, or an empty control) was transformed into yeast with a standard lithium acetate transformation, performed at 30°C with 200ng of plasmid. Yeast were grown on solid media without uracil and leucine, which selected for cells with both pRS316 and pYY1. Yeast were grown for 3 days at 30°C, then individual colonies were picked into 2mL liquid media lacking uracil and leucine. These cultures were grown for 2 days at 30°C, shaking at 275 rpm. Then, 1mL of saturated culture was centrifuged at 15,000 rpm for 1 minute and cell pellets were re-suspended in 50µl water. Yeast were serially diluted to 1:10, 1:100, or 1:1000 using water. 10µl of each dilution (included undiluted yeast) was spotted on complete media containing 5-FOA (Teknova), which selects for cells that have spontaneously lost the pRS315 vector expressing *URA3* and *THS1*.[400] After 3 to 5 days, yeast growth was visually inspected.

3.2.3 Cloning tars-1 expression constructs

Wild-type *tars-1* was amplified from *C. elegans* cDNA with primers containing the attB1 and attB2 Gateway recombination sequence (sequences found in Appendix A), then purified and recombined into pDONR221 using the Gateway BP reaction (Invitrogen). Clones were isolated as described above, and successful recombination into vectors was confirmed by *Bsr*GI digest followed by Sanger sequencing. Site-directed mutagenesis using the QuickChange II XL Site-Directed Mutagenesis Kit (Agilent) was performed on wild-type *tars-1* in pDONR221 (mutagenesis primer sequences in Appendix A). Plasmids were transformed into Top10 cells, purified, and verified via Sanger sequencing. Wild-type or mutant *tars*-1 was recombined into the expression vector using the Gateway LR reaction. This resulting expression vector contains a *unc-25* promoter and *let858* termination sequence, which is active specifically in worm GABA-ergic neurons.[401] The LR reaction was transformed into Top10 cells, purified, and verified by *Bsr*GI digest.

3.2.4 Overexpressing tars-1 in GABA-ergic neurons

To inject worms with *tars-1*, the following mixes were prepared for each construct: 50ng/µl expression plasmid bearing *tars-1*, 47.5ng/µl 1kb+ ladder (which facilitates uptake of the plasmid), and 2.5 ng/µl pCFJ90 plasmid, a co-injection marker that leads to transient expression

of mCherry in the pharynx. This mix was delivered to the worm gonadal tract using standard microinjection techniques,[402] using an inverted microscope (Olympus IX71) and micromanipulator (Narishige). The oxIS12 worm strain was used for these experiments. These worms harbor a transgene for GFP expression in GABA motor neurons (Punc-47::GFP), which allows visualization of commissural processes.[403]

3.2.5 CRISPR-Cas9 genome editing in worm

CRISPR-Cas9 genome editing was performed according to previously described methods.[404] Briefly, the gonadal tract of P1 adult worms was injected with an injection mix of: 300mM KCl, 20mM HEPES, 2.5 ng/µl pCFJ90, 50ng/µl single stranded oligonucleotide homologous donor repair template (Integrated DNA Technologies), 5µM single guide (sg) RNA (Synthego), and 5µM Cas9 protein (Integrated DNA Technologies). Sequences for the repair templates and guide RNAs can be found in Appendix A. Injected worms were then placed on single 35mm plates of nematode growth media (NGM) and fresh OP50 bacteria as a food source. Approximately 2 days after injection, plates were screened for the presence of F1 progeny expressing the pCFJ90 marker, which expresses mCherry in the pharyngeal muscles. This enriches for worms that were exposed to the injection mix, increasing the likelihood of identifying a worm subjected to genome editing.

The mCherry-positive F1s were singled to individual plates and allowed to produce their own offspring (F2). Then, the F1 worms were placed in lysis buffer (50mM KCl, 10mM Tris-HCl pH 8.3, 2.5mM $MgCl_2$, 0.45% NP-40, 0.45% Tween-20, 1mg/mL proteinase K) and lysed with incubation at -80°C for one hour, incubation at 65°C for one hour, and incubation at 95°C for fifteen minutes. To genotype worms, the targeted *tars-1* region was amplified by PCR (primer sequences in Appendix A) using Q5 PCR mix (New England Biolabs). Amplicons were then purified with DNA Clean and Concentrator kits (Zymo Research) and digested with the appropriate restriction enzyme (*Eag*I for G541R or *Sac*I for R433H, New England Biolabs). Digested PCR products were separated on a 1% agarose gel and analyzed to identify successful integration of the restriction site. Then, the undigested PCR product from F1s with successful gene editing events was submitted for Sanger sequencing to confirm proper insertion of the

restriction site and the desired *tars-1* mutation. The offspring of these F1 worms were then maintained for subsequent experiments.

3.2.6 Back-crossing and balancing worm strains

To reduce possible off-target mutations caused by CRISPR-Cas9 editing, G541R/+ *tars-1* worms were back-crossed to the ancestral N2 strain five times, and R433H/+ *tars-1* worms were back-crossed six times. To analyze G541R in the heterozygous state, wild-type and G541R/+ worms were balanced with the mIN1 inversion. This inversion stretches across the *tars-1* locus,[405] which prevents recombination with the G541R allele. This balancer allele also contains a transgene driving GFP expression in the pharynx, as well as a recessive *dpy* mutation that confers the Dpy phenotype to any homozygous worms.[405] Thus, possible phenotypes and genotypes from the offspring of a balanced hermaphrodite worm are: GFP and Dpy (+/+ *tars-1*), GFP and non-Dpy (G541R/+ *tars-1*), and non-GFP non-Dpy (G541R/G541R).

3.2.7 Analysis of worm axonal morphology

To visualize changes in GABA-ergic neuron morphology, worms were crossed to a oxIS12 background, which contains a transgene expressing GFP in GABA-ergic neurons.[403] These worms were then age-synchronized by placing fertile adult hermaphrodites on a plate to lay embryos for 2-5 hours, then removing the adults and letting the offspring grow to adulthood. When worms reached L4 or P1, sets of 25-30 worms were moved to fresh plates and continuously transferred to fresh plates every other day until P7, when they were analyzed. Approximately 5µl of polystyrene beads (Polysciences Incorporated) were placed on a 10% agarose pad, and 5-7 worms were deposited in the beads. A coverslip was gently placed on top of the worms to encourage immobilization in the beads. Then, worms were imaged on a Nikon Eclipse Ti microscope. Any axon breaks, thinning of the dorsal nerve cord, or axonal branching were scored as abnormalities. This experiment was performed blinded to genotype.

3.2.8 Analysis of worm development

To identify differences in the rate of development, R433H/R433H *tars-1* worms and wild-type N2 worms were first age-synchronized by placing approximately 25 adult worms on a 60 mm plate with NGM and OP50, letting them produce embryos for 4-5 hours, and then removing the adults. After 48 hours, worms were transferred to unseeded 35mm NGM plates in batches of 4-5 worms. These worms were filmed and analyzed using the WormLab System (MBF Biosciences). Plates were filmed for 30-second intervals, with the camera set at 4.81um/pixel for R433H/R433H worms (Setting 1 on the Wormlab camera apparatus [MBF Bioscience]) and 8.47um/pixel for N2 worms (Setting 3 on the Wormlab camera apparatus [MBF Bioscience]). After filming, worms were moved to new NGM plates seeded with OP50. Filming was repeated every 24 hours up to 168 hours, or 7 days, after birth (as R433H/R433H worms increased in size, filming was performed with the camera setting at 8.47um/pixel). All videos were analyzed with the WormLab software (MBF Bioscience), and the worm length parameter was extracted to compare the size of R433H/R433H *tars-1* worms and N2 worms over the course of development.

3.2.9 Worm thrash assays

Thrash assays were performed to detect changes in worm movement. The bottom of a 66mm well (Thermo Scientific) was coated with 2.5% agar. 500µl liquid M9 media (22mM H_2KO_4P, 42 mM HNa_2O_4P, 85 mM NaCl, 1mM $MgSO_4$) was added to each well, and 1-5 worms (wild-type or R433H/R433H *tars-1*) were placed in the M9. Worms were allowed to acclimate for 30-60 seconds before they were filmed with the WormLab System (MBF Biosciences) for 1 minute. Only worms with at least 1,000 frames of high-quality video were included in subsequent analysis. To identify defects in locomotion, the WormLab "Speed" parameter was analyzed; this parameter calculates changes in the center point of the worm over time.

3.2.10 Generation of G541R and R433H Tars1 mouse lines

The G541R mutation was introduced into mouse *Tars1* using CRISPR-Cas9 mediated gene editing, performed by the University of Michigan Transgenic Animal Core. A single-stranded oligonucleotide (ssODN) was designed to introduce the G541R mutation *in cis* with silent mutations that introduced a *Hha*I cut site and destroyed the PAM site. A mixture of Cas9,

sgRNA, and ssODN was injected into 308 hybrid C57BL/6J x SJL/J zygotes, and the 280 surviving zygotes were transferred to 11 pseudopregnant female mice. These mice produced 59 offspring, which were all genotyped by PCR-amplification (primer sequences in Appendix A) and *Hha*I digestion to identify successful introduction of the repair template. PCR products were submitted for Sanger sequencing to ensure the proper integration of the entire repair template. Mice carrying the G541R *Tars1* allele were then mated to C57BL/6 mice to establish germline transmission.

A similar approach was taken to generate the R433H *Tars1* mutation. Synonymous mutations in the repair template *in cis* with R433H ablate a *Bgl*I cut site that is present in the wild-type allele, and prevent binding of the guide RNA after repair. Cas9, sgRNA, and ssODN was injected into hybrid zygotes, which were then implanted into pseudopregnant females. These mice produced 32 pups, which were genotyped to by PCR-amplification (primer sequences in Appendix A) and *Bgl*I digestion to identify mice that had incorporated the repair template. Amplicons were then submitted for Sanger sequencing to identify mice with proper integration of the repair template. These mice were then mated to C57BL/6 mice to establish germline transmission.

3.2.11 Mouse behavioral assays

To analyze gait, mouse paws were painted with non-toxic paint (the back paws with red paint and the front with paws blue paint). Mice were prompted to walk down a strip of paper fitted inside a narrow walkway. After the paint had dried, three subsequent footsteps in the middle of the strip were analyzed. The length between each step and the width of the mouse stance was measured. To test mice for general motor function, an accelerated rotarod test was used. Mice were placed on a rotating horizontal bar 10 inches above the desktop, rotating at 4 rpm. The speed was gradually increased to 40 rpm over the course of 5 minutes; if the mouse fell, the time to fall was recorded. Mice were also tested for their ability to run on a treadmill. Each mouse was placed on a treadmill in front of an electric grid. The treadmill began moving at 9 meters/minute, then increased by 1 meter/minute each minute. If a mouse fell off, it received a light electrical shock. The time until a mouse received three electrical shocks (counted as at least one second with multiple paws on the grid) was recorded. The test was stopped after 22 minutes. An

alternate test of activity was performed using a running wheel. Mice were placed in cages containing food, water, and a running wheel. This wheel recorded the total time the mice spent running on the wheel, as well as the total distance they traveled. Mice were housed in these cages for 7 days, and the total distance traveled each day was recorded. To test motor strength, a "wire hang" test was performed. Mice were placed on a wire grid lying horizontally on the table top. The grid was then turned upside down and held approximately 10 inches above an empty cage for 3 minutes. The time to fall was recorded with a stopwatch. To test grip strength, a Grip Strength Meter (Columbus Instruments) was used. Mice were held and allowed to hold the test apparatus with their fore paws. Then, gentle and uniform force was applied at the base of its tail until the force applied exceeded the mouse's ability to maintain its grip. The peak force was then recorded as the mouse's grip strength. Each mouse was tested in 5 subsequent trials. All behavioral analyses were performed blinded to genotype.

3.2.12 Mouse nerve conduction measurements

To assess defects in nerve conduction amplitude and velocity, nerve conduction studies were performed according to published protocol.[406] In short, mice were anesthetized and core temperature was maintained at 34 °C with a heating lamp. Skin temperature was monitored with an infrared probe (Fluke, Everret, WA) and kept at 32 °C. Stainless steel needle electrodes (Natus, Middleton, WI) were cleaned with 70% alcohol between animals. Sural sensory nerve conduction velocity (NCV) was determined by recording at the dorsum of the foot and antidromically stimulating with supramaximal stimulation at the ankle. Nerve conduction velocity was calculated by dividing the distance by the take-off latency of the sensory nerve action potential. Sciatic-tibial motor NCV was determined by recording at the dorsum of the foot and orthodromically stimulating with supramaximal stimulation first at the ankle, then at the sciatic notch. Latencies were measured in each case from the initial onset of the compound muscle action potential. The sciatic-tibial motor NCV was calculated by subtracting the measured ankle distance from the measured notch distance. The resultant distance was then divided by the difference in the ankle and notch latencies for a final nerve conduction velocity. Amplitude was calculated at the dorsum to ankle measurement. Amplitude was calculated from the onset of the compound muscle action potential to the peak of the negative deflection.

3.2.13 Mouse dissections

The author assisted Dr. Guy Lenk (Meisler Laboratory, University of Michigan) with dissections on two *Tars1*$^{+/+}$ mice, two *Tars1*$^{G541R/+}$ mice, and two *Tars1*$^{F538Kfs*4/+}$ mice. All mice were eleven weeks old. Mice were euthanized by cervical dislocation and decapitation. Brains were dissected and lysed with 1mL RIPA buffer (Thermo Scientific) and 1x Halt Protease Inhibitor Cocktail (Thermo Scientific) in a dounce homogenizer. Then, lysate was centrifuged at 4,500xg for 5 minutes at 4°C. Supernatants were removed and aliquoted for protein analysis.

3.2.14 Western blots from mouse brain

Total protein concentration was measured using the Thermo Scientific Pierce BCA Protein Assay Kit. 6.25µg, 12.5µg, or 25µg of lysate was analyzed. Samples were prepared with 1X Novex Tris-Glycine SDS sample buffer (Invitrogen) and 2-mercaptoethanol (BME), and were boiled at 99°C for 5 minutes. Protein samples were separated on precast 4-20% Novex Wedgewell Tris-glycine gels (Invitrogen) at 150V for 1 hour and 15 minutes. PVDF membranes (Millipore Sigma) were pre-washed in 100% methanol for 1 minute, then soaked in 1X transfer buffer (Invitrogen) and 10% methanol between two pieces of filter paper (Thermo Fisher Scientific). Samples were transferred from the Tris-Glycine gel to the PVDF membrane using a Mini Trans-Blot Electrophoretic Transfer Cell (Biorad) at 0.03A for 18-20 hours. The membrane was then blocked in 2% milk in 1X TBST overnight at 4°C. After blocking, the membrane was washed with 1X TBST three times, with each wash comprising five minutes of rocking at room temperature. Primary antibody was applied in the 2% milk solution: anti-TARS1 (Thermo Fisher PA5-30690) was applied at 1:500 dilution, and anti-actin (Sigma A5060) was applied as a loading control at 1:5,000 dilution. Primary antibody was incubated overnight at 4°C. Membranes were then washed three times with 1X TBST as above, and secondary antibodies (anti-mouse HRP [1:2,000; Thermo Fisher Scientific] for TARS1 and anti-rabbit HRP [1:5,000; EMD Millipore] for actin) were applied in 2% milk solution. The blots were rocked for 1 hour at room temperature before incubating with SuperSignal West Dura substrate (Thermo Scientific) according to the manufacturer's instructions.

3.2.15 Preparation of mouse tissues for histology

To investigate P0 pups for gross histological changes, dead pups were collected and live pups were sacrificed by decapitation. Pups were individually fixed in neutral-buffered formalin, rocking overnight at room temperature. Then, pups were placed in 70% ethanol and stored at 4°C. To investigate histological changes in the hair and skin of adult mice, three *Tars1*$^{R433H/F538Kfs*4}$ mice and their age-matched, sex-matched *Tars1*$^{R433H/+}$ littermates were sacrificed. The mice were shaved, and skin was collected from the dorsal trunk, ventral trunk, ears, tail, and paws, as well as from the area of the head with visible hair loss. Skin samples were placed on 0.45μm HA filters (Millipore) wetted in PBS and strips were cut parallel to the direction of hair follicle growth. All strips were then fixed overnight in neutral-buffered formalin at room temperature, then transferred to 70% ethanol and stored at 4°C.

Samples were shipped to Histoserv, Inc. for embedding and sectioning. Briefly, samples with bone were decalcified; then, tissues were dehydrated, and water inside of the tissues was replaced with paraffin wax. Tissues were then embedded into wax blocks of paraffin. Blocks were then sectioned and affixed to slides (two sagittal sections were taken for the P0 pups). Adult skin sections were stained with H&E; P0 pup sections were stained with either H&E or PAS, which detects glycoproteins and mucins.[407]

3.2.16 Analysis of epidermal thickness in P0 pups

Dorsal skin from H&E-stained sections was used to analyze the epidermal thickness of four *Tars1*$^{R433H/F538Kfs*4}$ mice and three *Tars1*$^{R433H/+}$ littermates. Five 1mm areas were selected, evenly spaced out across the back. In each 1 mm area, the thickness of the epidermis was measured by drawing lines in Adobe Illustrator that span the width of the epidermal layer, then using the 200um scale in each image to convert line length to um. Five measurements were made that evenly spanned the 1mm area; each measurement was made at the widest local area.

3.3 Results

3.3.1 Identification of three loss-of-function TARS1 mutations

Previous work in the Antonellis group had assessed a panel of variants in the yeast threonyl-tRNA synthetase gene, *THS1*. Three of these variants reduced yeast growth and were therefore selected for study in the human *TARS1* open reading frame. The three variants—N412Y, R433H, and G541R—were designed to affect residues conserved between human, mouse, worm, and yeast (Figure 3.2A). To assess whether these variants affected the function of human *TARS1*, a complementation assay was performed using a yeast strain with the endogenous *THS1* deleted. Yeast viability was maintained with a pRS316 vector expressing the yeast *THS1*, along with *URA3*. The pYY1 vector expressing either wild-type or mutant human *TARS1* was transformed into yeast, then yeast were plated on 5-FOA, which selects for the loss of the maintenance vector expressing *URA3* and *THS1*. Wild-type *TARS1* supported yeast growth, demonstrating that human *TARS1* can function in yeast (Figure 3.2B). Transformation with N412Y *TARS1* or G541R *TARS1* did not lead to formation of colonies, indicating that these two mutations significantly impair gene function. Transformation with R433H *TARS1* did support yeast growth, but lead to significantly fewer colonies than wild-type *TARS1*, indicating it was a partial loss-of-function allele (Figure 3.2B). Based on these data, all three variants were prioritized for studies in worm to determine whether they were dominantly toxic to *C. elegans* neurons. Of note, the variant nomenclature used in this Chapter will remain consistent with the amino acid numbers of the human alleles for ease of comprehension. However, the *C. elegans* and mouse numbers differ from the human by one amino acid residue (Table 3.1).

3.3.2 Over-expression of G541R tars-1 is dominantly toxic to C. elegans neurons

Pathogenic loss-of-function variants in *HARS1* and *AARS1* are dominantly toxic to *C. elegans* neurons when over-expressed from a transgene, causing a branching morphology in GABA-ergic neurons and a failure to reach from the ventral nerve cord to the dorsal nerve cord.[174,175,307] To determine if the *TARS1* loss-of-function variants could cause a similar phenotype, they were over-expressed in *C. elegans* GABA-ergic neurons. GABA-ergic neurons are a population of 19

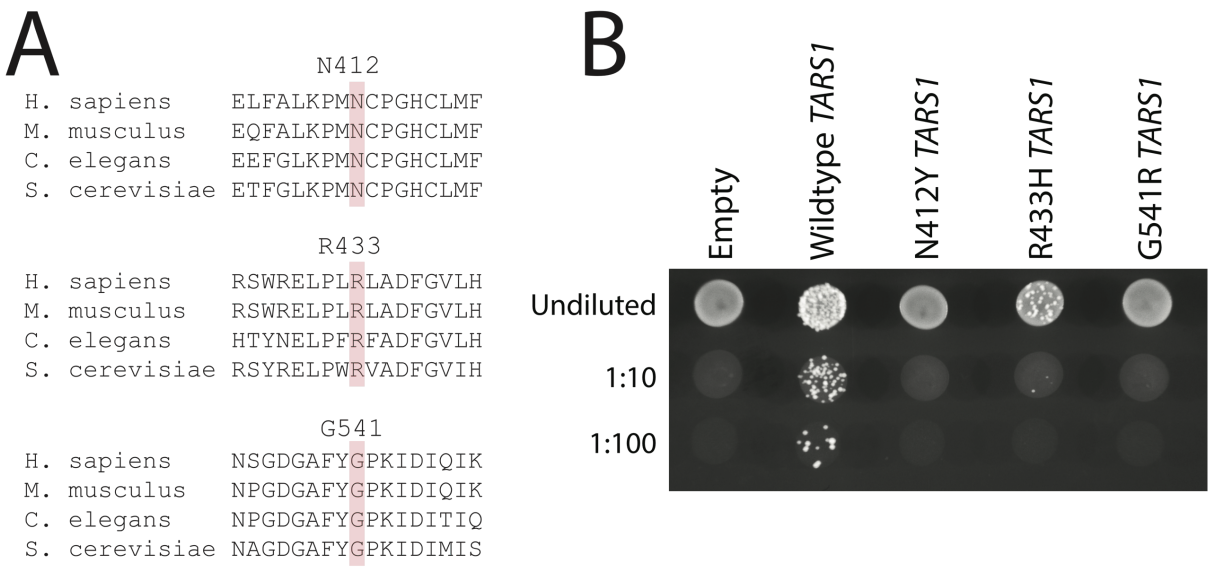


Figure 3.2. Conservation and yeast complementation analyses for three designed *TARS1* variants.

(A) Conservation analysis of N412, R433, and G541 TARS1. The targeted residues are shown in pink, surrounded by flanking sequences from evolutionarily diverse species. **(B)** Representative images of yeast haploid strains with *THS1* deleted, transformed with an empty vector ("Empty"), wild-type *TARS1*, N412Y *TARS1*, R433H *TARS1*, or G541R *TARS1*. Yeast were spotted on media containing +5FOA in serial dilutions, then grown at 30°C.

Table 3.1. Comparison of amino acid numbering between human TARS1, worm tars-1, and mouse Tars1.

***TARS1* variant (*H. sapiens*)**	***tars-1* variant (*C. elegans*)**	***Tars1* variant (*M. musculus*)**
R433H	R432H	R432H
G541R	G540R	G540R

motor neurons that innervate body wall muscles, are required for locomotion,[408] and have defined projections across the circumference of the worm that are easily visualized with GFP (Figure 3.3B). N412Y, R433H, and G541R were each introduced to the *tars-1* coding sequence, which was expressed from the GABA-ergic neuron promoter *unc-47*.[403] In this construct, *tars-1* is separated from a mCherry reporter by a SL2 splice site, which allows mCherry signal to mark proper transgene expression without introducing a tag onto the tars-1 protein[409] (Figure 3.3A). Expressing wild-type *tars-1* did not lead to any morphological defects, indicating that increased levels of wild-type *tars-1* is not toxic. Similarly, R433H *tars-1* did not cause changes in neuronal morphology. However, G541R *tars-1* caused failure of commissures to reach the dorsal nerve cord, aberrant commissure branching, and thinning of the dorsal nerve cord (Figure 3.3C) This phenotype recapitulated the morphological defects previously seen with *hars-1* and *aars-2* mutants.[174,175,307] Worms expressing N412Y *tars-1* could not be recovered. These findings were based on small numbers of animals, so statistical significance cannot be assessed. However, based on the compelling neuronal morphology defects, G541R *TARS1* was pursued to model in the endogenous locus. R433H *TARS1* was also knocked into the endogenous locus as a potential candidate for recessive *TARS*-mediated phenotypes, as it was a partial loss-of-function allele in yeast but was not dominantly toxic to worm neurons.

3.3.3 Introduction of tars-1 variants to the endogenous worm locus

A major limitation of the above *tars-1* transgenic studies is the inability to control for the expression level of the mutant protein. In this plasmid-based method, the number of plasmid copies that are incorporated into extrachromosomal arrays can be variable, as can the degree of array silencing.[410] This leads to inconsistencies from cell to cell, and from worm to worm. To address these limitations, CRISPR-Cas9 mutagenesis was used to introduce R433H or G541R into the endogenous *tars-1* locus, using previously detailed methods for site-specific gene editing in worm.[404] In addition to the desired *tars-1* missense mutation, silent mutations were included in the ssODN repair template to introduce unique restriction enzyme cut sites (*Eag*I *in cis* with G541R and *Sac*I *in cis* with R433H). These enable quick identification of worms that are heterozygous or homozygous for the mutation of interest by amplifying the locus with PCR and

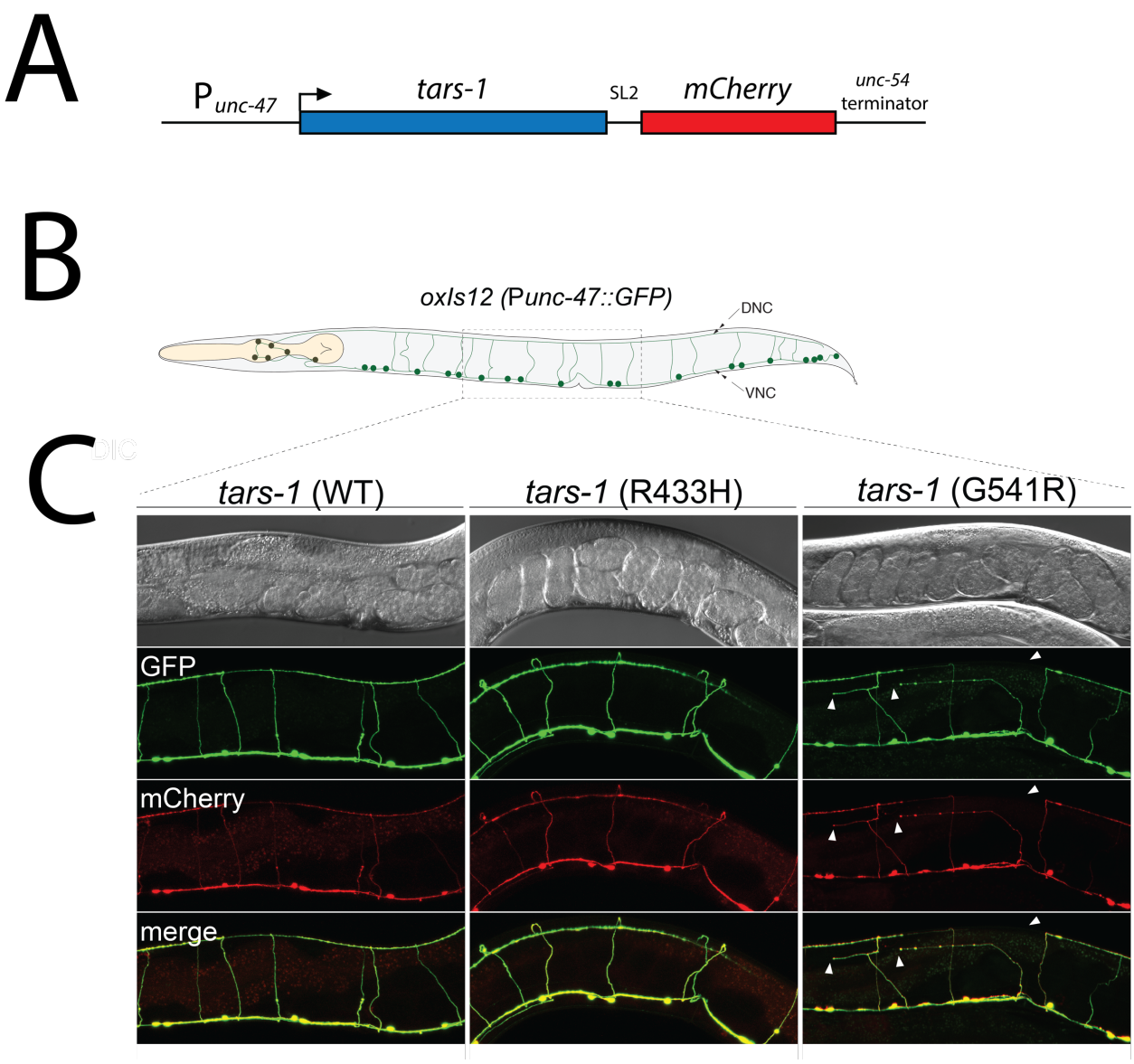


Figure 3.3. Over-expression of G541R *tars-1* is dominantly toxic to worm GABA-ergic neurons.

(A) Cartoon of the construct injected into *C. elegans* gonadal tract. A GABA-ergic neuron promoter, P_{unc-47}, drives expression of wild-type or mutant *tars-1*, followed by a mCherry reporter and the *unc-54* termination sequence. **(B)** Worms were analyzed in the oxIS12 background, which expresses GFP in GABA-ergic neurons for easy visualization of the ventral nerve cord (VNC) and its projections across the body to the dorsal nerve cord (DNC). **(C)** Images of GABA-ergic neurons from worms over-expressing wild-type *tars-1* (left), R433H *tars-1* (middle), or G541R *tars-1* (right). Transgenic expression is indicated by mCherry signal. Aberrant neuronal morphology in worms expressing G541R is highlighted with white arrows.

digesting with the relevant enzyme to detect the restriction site marker of successful homology-directed repair. This approach was used to screen the progeny of injected worms and identify worms that were heterozygous for G541R (Figure 3.4A) and worms that were either heterozygous or homozygous for R433H (Figure 3.4B).

3.3.4 G541R tars-1 is homozygous lethal in worm

An additional advantage of gene editing is the ability to stably propagate a mutation and follow its inheritance pattern across subsequent generations. Variation from the expected 1:2:1 pattern of allele segregation indicates that a mutation is deleterious. To assess the viability of G541R and R433H *tars-1* worms, individual unbalanced heterozygote hermaphrodites were moved to separate plates, where they self-fertilized to produce dozens of offspring. Upon maturation to the L4 or P1 stage of development, each progeny was individually genotyped. The observed ratio of homozygote wild-type, heterozygote mutant, and homozygote mutant was compared to the 1:2:1 ratio expected for allelic segregation. Out of 91 offspring, no G541R/G541R offspring were recovered, indicating that G541R is a loss-of-function allele (Figure 3.5A). This confirms the results from the yeast complementation assay. In contrast, R433H/R433H homozygotes were recovered, but at a lower frequency than would be expected for a benign variant (Figure 3.5B). This may indicate that R433H homozygotes were born less often R433H heterozygotes and wild-type worms, or that there were fewer homozygotes on the plate that had reached the genotyping L4/P1 (adolescent/adult) stage, leading to an under-sampling of this genotype. In either scenario, the data indicate that homozygosity—but not heterozygosity—for R433H is deleterious but not lethal. This also agrees with the yeast complementation data indicating R433H is a hypomorphic allele.

3.3.5 R433H tars-1 *causes recessive developmental delay and locomotion defects in worm*

The depletion of R433H/R433H *tars-1* worms in the above Mendelian analyses could be due to under-sampling of this genotype, which would happen if developmental delay prevented them from reaching adulthood at the same rate as wild-type or R433H/+ *tars-1* worms. To investigate this possibility, a population of R433H/R433H *tars-1* worms and wild-type N2 worms were age-

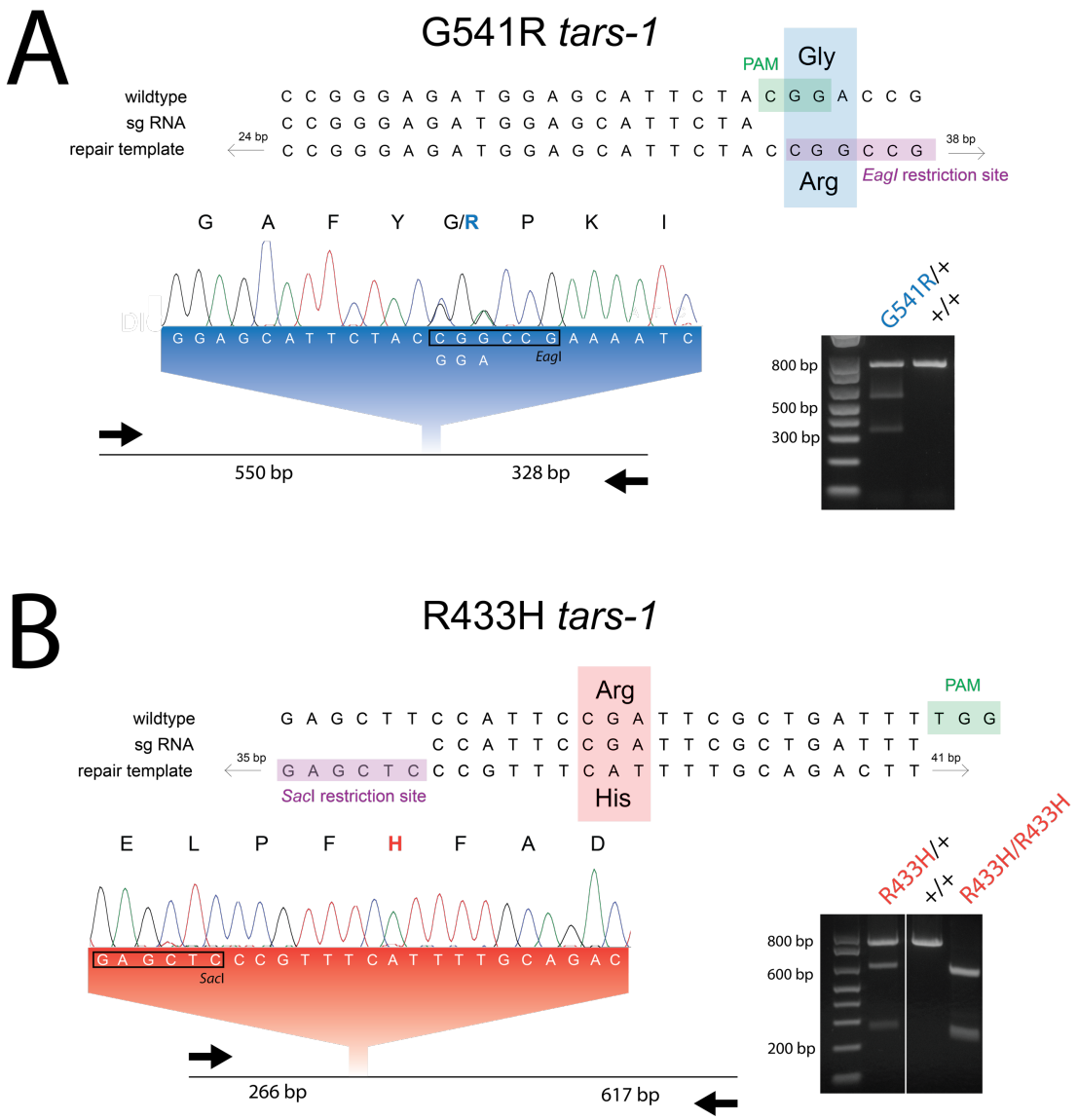


Figure 3.4. CRISPR-Cas9 strategies to introduce G541R or R433H into the endogenous *tars-1* locus.

(A) Top: Cartoon of the G541 locus in *tars-1*, aligned with the repair template sequence and guide RNA. Introduction of point mutations that change the glycine to arginine (blue) simultaneously ablate the PAM site (green) and introduce an *Eag*I restriction site (purple) for genotyping. Bottom left: cartoon depicting PCR amplification of the targeted region, with a representative chromatogram above it. Bottom right: *Eag*I digestion of amplicons from a wild-type or heterozygous worm. **(B)** Top: cartoon of the R433 locus in *tars-1*, aligned with the repair template and guide RNA sequence. The repair template encodes the arginine to histidine change (red), as well as synonymous mutations that introduce a *Sac*I cut site (purple) and prevent re-editing after repair. Bottom left: cartoon depicting PCR amplification of the targeted region, with a representative chromatogram of a homozygous worm above it. Bottom right: *Sac*I digestion of amplicons from a wild-type, heterozygous, or homozygous worm.

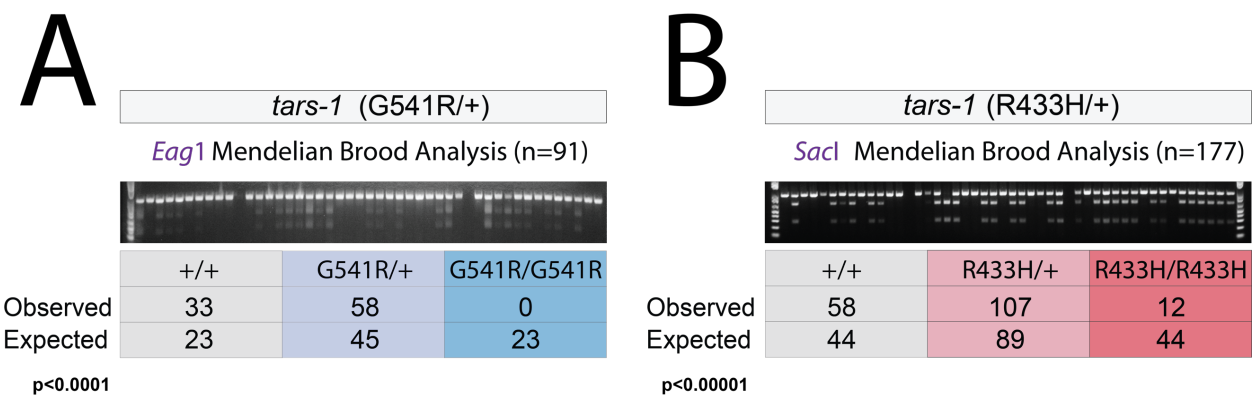


Figure 3.5. Mendelian analysis of G541R/+ *tars-1* offspring and R433H/+ *tars-1* offspring.

(A) Genotype analysis of 91 offspring from G541R/+ *tars-1* hermaphrodites. A representative genotyping gel image is shown. The observed and expected number of each genotype is listed. **(B)** Genotype analysis of 177 offspring from R433H/+ *tars-1*; a representative image from a genotyping gel is shown. The observed and expected number of each genotype is shown. For both analyses, a Chi-square was performed to determine if the difference between observed genotype frequency and expected genotype frequency was statistically significant.

synchronized and the size of worm cohorts was tracked for over 7 days. At 48 hours after worms were hatched, worm length was measured each day using the WormLab video and software system. R433H/R433H *tars-1* worms were consistently smaller than wild-type controls until Day 7 (Figure 3.6A). Whereas wild-type worms reach a mature size of approximately 1mm 3-4 days after birth, R433H/R433H *tars-1* worms do not reach this size until 6-7 days after birth.

To assess whether R433H/R433H *tars-1* causes behavioral defects, a thrash assay was performed with adult worms 9 days after they reached adulthood (P9). Here, R433H/R433H *tars*-1 worms were age-matched to wild-type N2s by synchronizing embryo production, as described in Section 3.2.7. The WormLab video capture and analysis system was used to record one-minute videos of worms swimming in M9 buffer, track their motion, and calculate locomotion parameters. R433H/R433H worms moved at a significantly reduced speed compared to wild-type worms (Figure 3.6B), indicating that reduced *tars-1* function affects the neuromuscular circuitry governing worm locomotion.

3.3.6 Morphological defects are absent in G541R/+ GABA-ergic neurons

To determine if G541R/+ worms could re-capitulate the GABA-ergic neuron defects seen when over-expressing G541R *tars-1*, they were crossed into the oxIS12 strain (which expresses GFP in the GABA-ergic neurons) and examined for neuronal abnormalities at P7. R433H/R433H *tars-1* worms at P9 were also included in these studies; if the dominant toxicity of G541R was due to a dominant-negative effect that decreased overall tars-1 function, then homozygosity for a hypomorphic allele might also reduce tars-1 function enough to cause neuronal defects. Worms were scored as having aberrant axons based on the presence of neuronal commissure branching, breakage, or incorrect guidance, as well as thinning of the dorsal nerve cord. Although some striking examples of aberrant morphology were identified in G541R/+ *tars-1* and R433H/R433H *tars-1* worms (Figure 3.7A), neither genotype caused a significant increase in the percentage of worms with morphological defects (Figure 3.7B and C). Based on the strength of the neuronal phenotype when G541R was over-expressed, G541R was selected for validation in a mammalian model, with the expectation that longer mouse peripheral nerves would be more sensitive to endogenous expression levels of a dominant ARS allele than the shorter worm neurons.

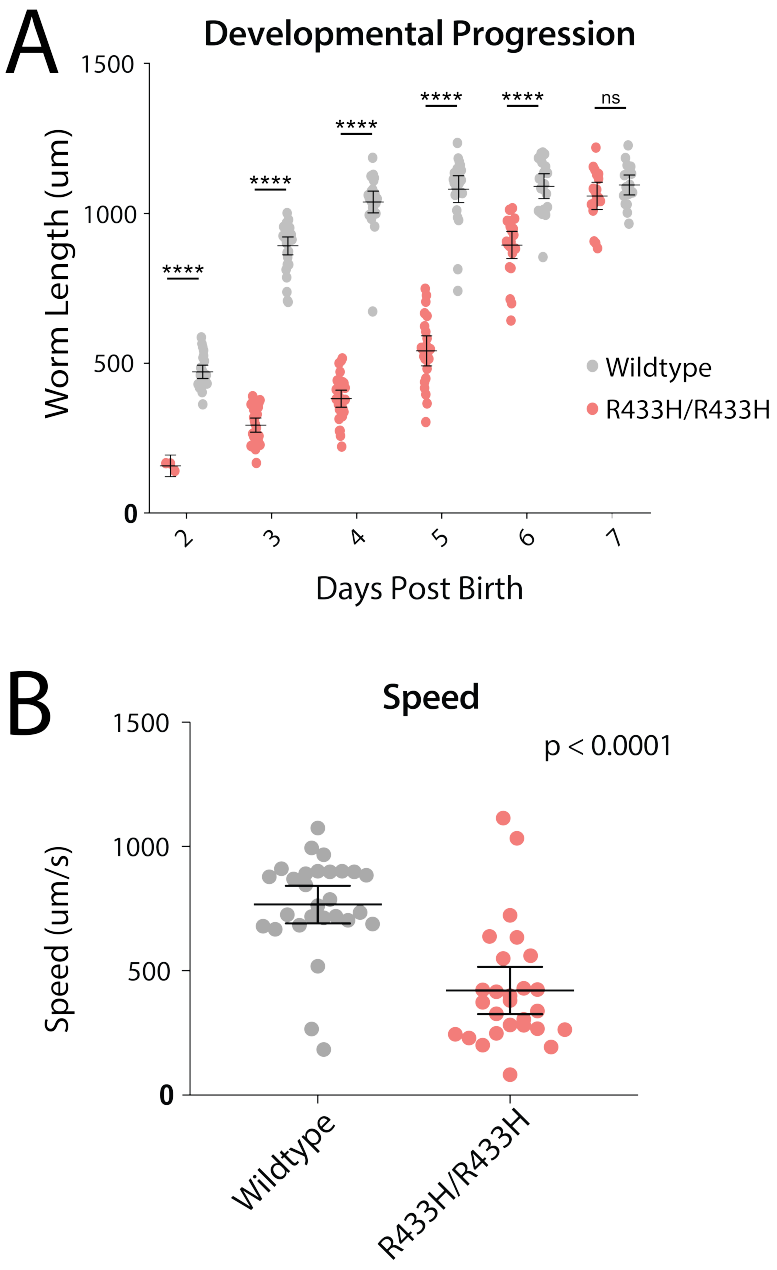
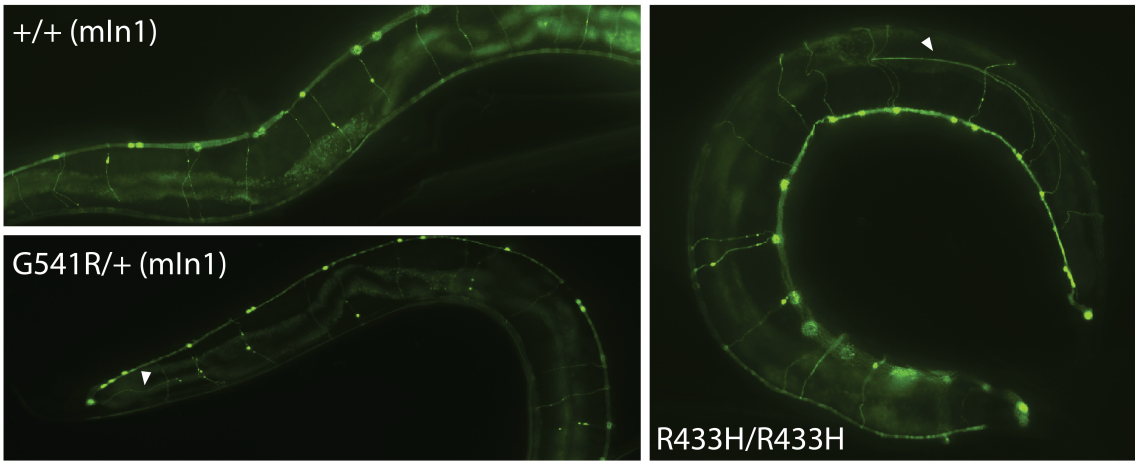


Figure 3.6. Developmental and locomotion defects in R433H/R433H *tars-1* worms.

(A) Body length measurements of R433H/R433H *tars-1* worms and wild-type *tars-1* worms (N2) across six days after hatching. 3 R433H/R433H worms were measured on Day 2, then 18-30 worms each day after. 18-30 wild-type worms were measured each day. Bars indicate the mean worm length for each day, along with 95% confidence intervals. Statistical significance was calculated with an unpaired t-test; ****, p<0.000001; ns = not significant. **(B)** Movement speed of 28 P9 wild-type worms (N2) and 27 P9 R433H/R433H *tars-1* worms is shown. Speed was calculated by analyzing the track length of the worm over time. Bars indicate mean speed and 95% confidence intervals. Statistical significance was calculated with an unpaired t-test.

A



B

Genotype	Worms with abberant axons	Total worm count	% Worms with abberant axons	
+/+ (mIN1)	15	102	14.7%	ns
G541R/+ (mIN1)	9	101	8.9%	

C

Genotype	Worms with abberant axons	Total worm count	% Worms with abberant axons	
+/+	11	101	10.9%	ns
R433H/ R433H	13	118	11%	

Figure 3.7. GABA-ergic neuron morphology in G541R/+ and R433H/R433H *tars-1* worms.

(A) Images of GABA-ergic neurons from wild-type worms with the mIn1 balancer, G541R/+ worms with the mIn1 balancer, and R433H/R433H worms. Morphological defects in G541R/+ worms and R433H/R433H worms are highlighted with white arrows. **(B)** Percentage of wild-type worms and G541R/+ worms with aberrant GABA-ergic neuron morphology. **(C)** Percentage of wild-type worms (N2) and R433H/R433H worms with abberant GABA-ergic neuron morphology. Statistical significance was calculated using Fisher's Exact Test; ns = not significant.

3.3.7 Establishment of G541R and R433H Tars1 mouse lines

The mouse *TARS1* ortholog, *Tars1*, was edited using CRISPR-Cas9 to introduce the G541R variant *in cis* with silent mutations that form a *Hha*I cut site and ablate the PAM site. Two founders carrying the G541R variant were identified and mated to establish germline transmission, then back-crossed to C57BL/6 mice twice for further analysis. Genotyping was performed by PCR-amplification of *Tars1* exon 14, followed by *Hha*I digestion to identify successful introduction of the repair template (Figure 3.8A). All PCR amplicons from the founder cohort were also Sanger-sequenced to identify insertions or deletions that might produce a frameshift and a premature stop codon. As a result, we identified a mouse carrying an 11 base pair deletion that ablates a *Hae*III cut site and that leads to a premature stop codon shortly downstream of the PAM site (F538Kfs*4). When this region is amplified and digested with *Hae*III, the presence of an undigested upper band indicates the frameshift allele (Figure 3.8B). Founder *Tars1*$^{F538Kfs*4/+}$ mice were also mated to C57BL/6 mice to establish germline transmission.

A similar approach was taken to identify mice that had undergone successful CRISPR-Cas9 editing of the R433H *Tars1* mutation. Here, silent mutations *in cis* with R433H ablate a *Bgl*I cut site and the PAM site. In this genotyping strategy, *Tars1* exon 12 is amplified with PCR and digested with *Bgl*I; amplicons that remain undigested indicate the R433H mutation (Figure 3.8C). Two *Tars1*$^{R433H/+}$ founders were identified and mated to establish germline transmission. One *Tars1*$^{R433H/+}$ line was backcrossed to C57BL/6 mice twice for further analysis.

*3.3.8 Homozygosity for G541R or F538Kfs*4 is not compatible with life in mouse models*

If G541R is a loss-of-function allele in mouse, as it is in yeast and worm, homozygosity for G541R *Tars* should not be compatible with life. To test this, three litters from *Tars*$^{G541R/+}$ x *Tars*$^{G541R/+}$ mating pairs were genotyped to assess whether the G541R allele segregated according to 1:2:1 Mendelian ratios. Out of 19 mice, no *Tars*$^{G541R/G541R}$ genotypes were detected at 3 weeks of age (Figure 3.9A). As a comparison, the same experiment was performed with the F538Kfs*4 *Tars1* mouse line; this mutation encodes a premature stop codon, producing a

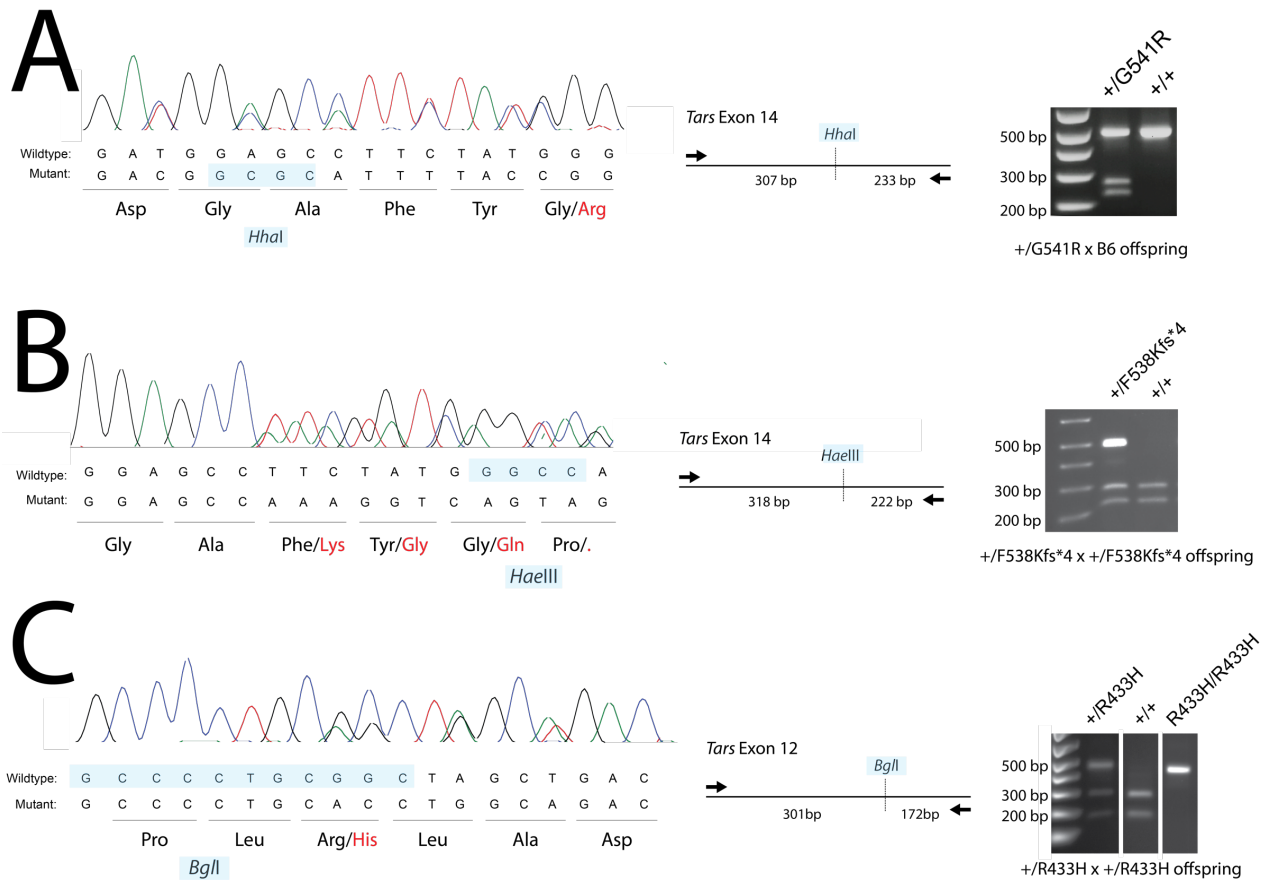


Figure 3.8. CRISPR-Cas9 genome editing to introduce G541R and R433H into the *Tars1* locus.

(A) Left: chromatogram of a G541R/+ *Tars1* mouse, with the sequence of the wild-type and mutant alleles below. The *Hha*I site introduced by the repair template is highlighted in blue, and the introduced arginine in red. Middle, right: genotyping strategy for amplifying exon 14 and identifying the G541R mutant allele by *Hha*I digest. **(B)** Left: chromatogram of a mouse with an 11 base pair deletion in Exon 14. This deletion causes a frameshift and a premature stop codon at amino acid 541 (F538Kfs*4). It also ablates a *Hae*III cut site (blue). Middle, right: genotyping strategy for amplifying exon 14 and identifying the frameshift allele by *Hae*III digest. **(C)** Left: chromatogram of a R433H/+ mouse, with the sequence of the wild-type and mutant alleles below. The nucleotides changed to convert arginine into histidine also ablate a *Bgl*I cut site in the wild-type allele (blue). Middle, right: genotyping strategy for amplifying exon 12 and identifying the R433H allele by *Bgl*I digest.

A

Genotype	Observed	Expected (Mendelian ratios)
+/+	5	4.75
G541R/+	14	9.5
G541R/G541R	0	4.75

Chi-square test: p=0.0318

B

Genotype	Observed	Expected (Mendelian ratios)
+/+	7	7
F538Kfs*4/+	21	14
F538Kfs*4/F538Kfs*4	0	7

Chi-square test: p=0.0052

C

Genotype	Observed	Expected (Mendelian ratios)
+/+	13	10.75
R433H/+	24	21.5
R433H/R433H	6	10.75

Chi-square test: p=0.2393

Figure 3.9. Mendelian ratios of mutant *Tars1* mouse lines.

(A) Genotype analysis of 19 offspring from *Tars1*$^{G541R/+}$ x *Tars1*$^{G541R/+}$ mating pairs. **(B)** Genotype analysis of 28 offspring from *Tars1*$^{F538Kfs*4/+}$ x *Tars1*$^{F538Kfs*4/+}$ mating pairs. **(C)** Genotype analysis of 43 offspring from *Tars1*$^{R433H/+}$ x *Tars1*$^{R433H/+}$ mating pairs. All mice were genotyped at approximately 3 weeks of age. Chi-square tests were performed to determine if the difference between observed genotype counts and expected genotype counts was statistically significant.

transcript that is predicted to undergo nonsense mediated decay. From a total of 28 offspring, no *Tars*$^{F538Kfs*4/F538Kfs*4}$ mice were identified (Figure 3.9B). These data indicate that both G541R *Tars1* and F538Kfs*4 *Tars1* are loss-of-function mutations in mouse. In contrast, *Tars*$^{R433H/R433H}$ mice were born at a frequency that did not deviate from expected Mendelian ratios (Figure 3.9C). This is consistent with R433H TARS1 retaining some function, as predicted by the yeast and worm studies.

3.3.9 Tars1$^{G541R/+}$ mice do not develop a detectable peripheral neuropathy by one year of age

To determine if G541R *Tars1* could cause a dominant peripheral neuropathy in mouse, cohorts of *Tars1*$^{G541R/+}$ mice and their littermate *Tars1*$^{+/+}$ controls were evaluated for behavioral indications of neuromuscular defects. First, two *Tars1*$^{G541R/+}$ founders were mated with C57BL/6 mice to establish two independent lines carrying the mutation (Line A and Line B). The Line A founder produced a litter of four males (two *Tars1*$^{G541R/+}$ and two *Tars1*$^{+/+}$). At 2 months of age, this cohort was evaluated for behavioral defects that might indicate neuromuscular impairment. Gait analysis revealed that the two *Tars1*$^{G541R/+}$ showed a 20% reduction in stride length, and took a 14% wider stance with their front paws (Figure 3.10A). However, there were no significant differences in rotarod performance (Figure 3.10B), ability to hang on a wire mesh (Figure 3.10C), or grip strength (Figure 3.10D) across five sequential days of testing.

To increase the sample size of this cohort, three *Tars1*$^{G541R/+}$ and three *Tars1*$^{+/+}$ males from the Line B founder were added. All ten mice were then subjected to the same battery of behavioral assays at 5-6 months of age. The stride length of all ten mice were assessed in a gait analysis; as before, *Tars1*$^{G541R/+}$ mice took shorter steps *Tars1*$^{+/+}$ mice (Figure 3.11B), despite no difference in body size between the two genotypes (Figure 3.11A). Stride width was only measured for the four Line A mice; again, the *Tars1*$^{G541R/+}$ mice took slightly wider stances with their front paws (Figure 3.11B). The combined Line A and Line B cohort was then tested for their grip strength, ability to hang from a wire mesh, ability to run on a treadmill, and performance on a rotarod. There was no difference in grip strength (Figure 3.12A) or latency to fall off the wire mesh (Figure 3.12B) between *Tars1*$^{G541R/+}$ and *Tars1*$^{+/+}$ mice. *Tars1*$^{G541R/+}$ mice did perform poorly on the first day of treadmill running and rotarod testing; however, this difference was not replicable

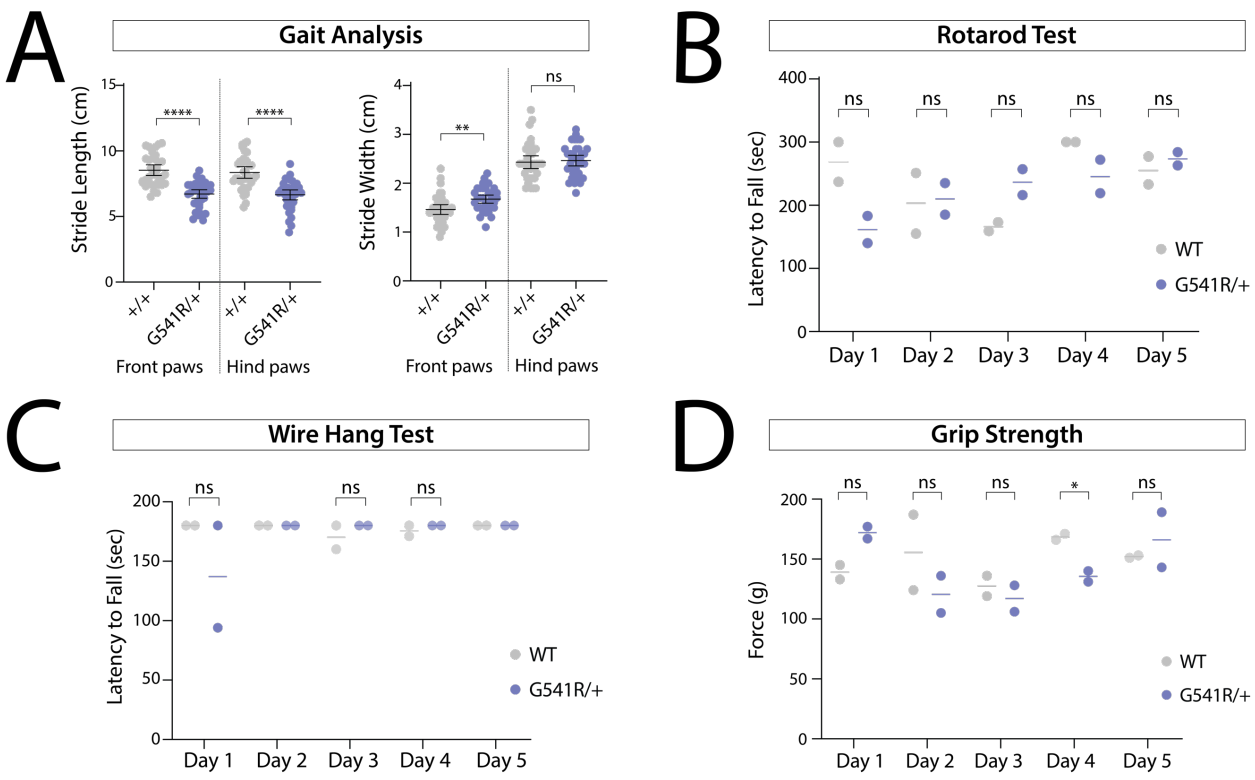


Figure 3.10. *Tars1*$^{G541R/+}$ mice have a shorter gait and wider stance at 2 months of age.

(A) Gait analysis for two male *Tars1*$^{G541R/+}$ mice and two male *Tars1*$^{+/+}$ littermates. 36-37 steps front paw steps and 36-37 hind paw steps were analyzed for each genotype. An unpaired t-test with Welch's correction was performed to determine statistical significance. **(B)** Rotarod performance data across five consecutive days **(C)** Latency to fall from a wire grid, measured for five consecutive days. **(D)** Grip strength, measured for five consecutive days. For **(B)-(D)**, unpaired t-tests were performed to determine statistical significance. **** $p<0.0001$, ** $p<0.01$, * $p<0.05$, ns = not significant.

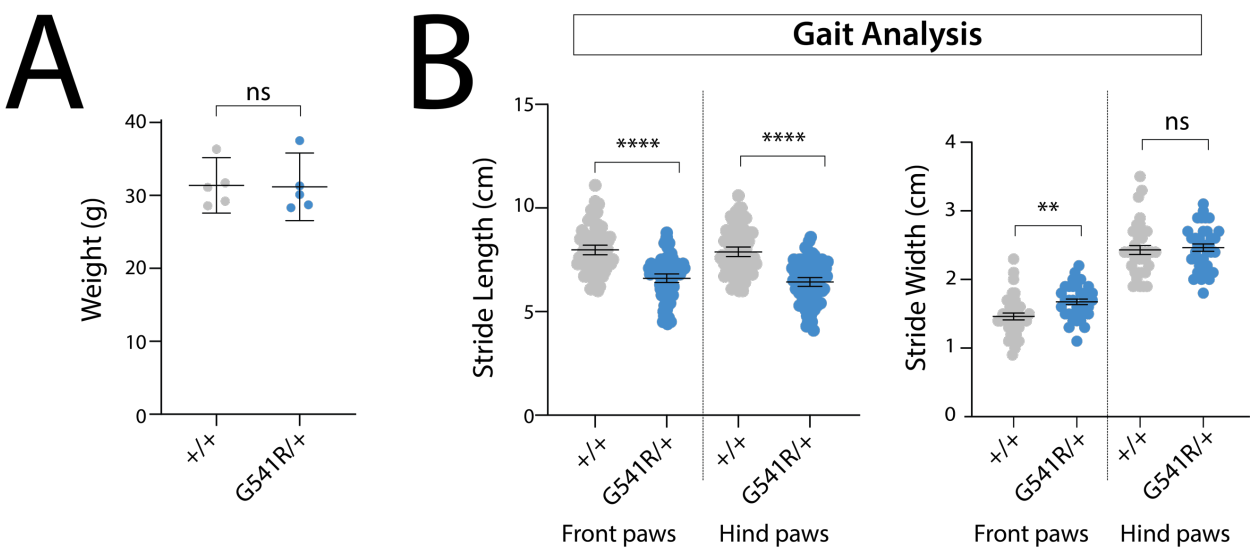


Figure 3.11. *Tars1*$^{G541R/+}$ mice have a shorter gait and wider stance at 5-6 months of age.

(A) Weight of 5 male *Tars1*$^{G541R/+}$ mice and 5 male *Tars1*$^{+/+}$ littermates at 5-6 months of age. **(B)** Gait analysis for 5 male *Tars1*$^{G541R/+}$ mice and 5 male *Tars1*$^{+/+}$ littermates. For stride length measurements, 85-90 front paw steps and 85-90 hind paw steps were analyzed for each genotype. For stride width measurements, only 2 male *Tars1*$^{G541R/+}$ mice and 2 male *Tars1*$^{+/+}$ mice were assessed; 36 steps front paw steps and 36 hind paw steps were analyzed for each genotype. For **(A)** and **(B)** an unpaired t-test with Welch's correction was performed to determine statistical significance. **** $p<0.0001$, ** $p<0.01$, ns = not significant.

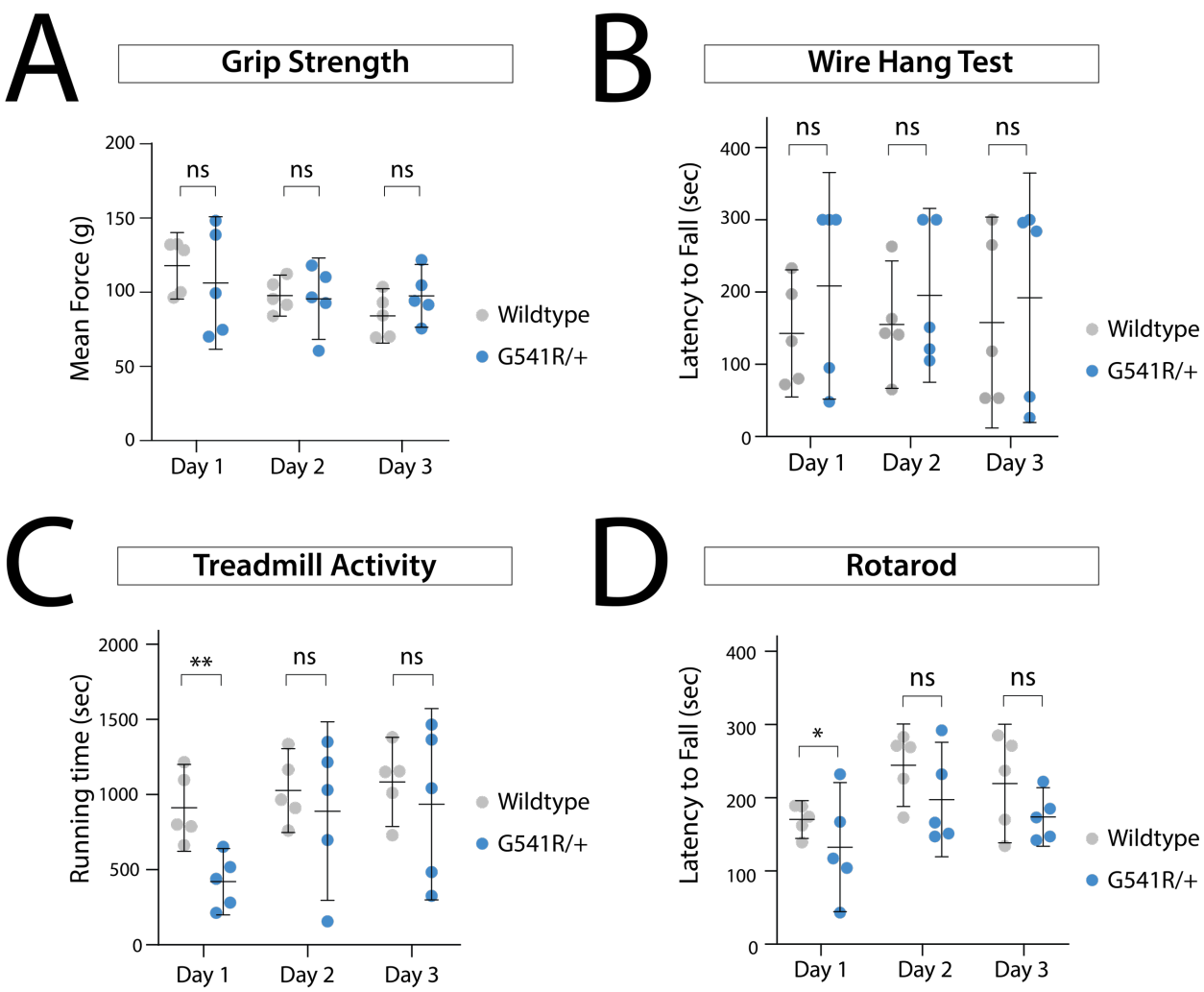


Figure 3.12 No difference between *Tars1*$^{G541R/+}$ mice and *Tars1*$^{G541R/+}$ mice in motor function or grip strength at 5-6 months of age.

(A) Grip strength of 5 male *Tars1*$^{G541R/+}$ mice and 5 male *Tars1*$^{+/+}$ littermate, measured for three consecutive days. **(B)** Latency to fall from a wire grid, measured for three consecutive days. **(C)** Length of time spent running on a treadmill, measured for three consecutive days. **(D)** Rotarod performance data, measured for three consecutive days. All bars represent mean value with 95% confidence intervals. Unpaired t-tests were performed for each assay to determine statistical significance. ** $p<0.01$, * $p<0.05$, ns = not significant.

in the subsequent two days of testing (Figure 3.12C and Figure 3.12D). Similar gait changes have been seen in other mouse models of dominant peripheral neuropathy[373,399] and are indicative of neuromuscular dysfunction. However, additional behavioral evidence was required to build a compelling argument for neuromuscular dysfunction in *Tars1*$^{G541R/+}$ mice. We pursued three strategies to more rigorously test the phenotype. One *Tars1*$^{G541R/+}$ line (Line B) was back-crossed an additional time to C57BL/6 mice to increase the amount of C57BL/6 genetic background, which served the additional purpose of continuing to cross away any possible off-target mutations induced during CRISPR-Cas9 mutagenesis. Then, a larger cohort of *Tars1*$^{G541R/+}$ mice was generated—8 *Tars1*$^{G541R/+}$ and 8 *Tars1*$^{+/+}$ from a total of three litters born within three weeks of each other— to provide the statistical power to detect subtle differences between the groups. Finally, this larger cohort was assessed at 1 year of age, which is more consistent with phenotypes observed in other mouse models of axonal CMT disease.[396] Indeed, because ARS-mediated peripheral neuropathy is a degenerative disorder, older mice were expected to show a more severe phenotype.

Surprisingly, when gait was assessed for 1-year-old mice, *Tars1* $^{G541R/+}$ mice took longer strides and smaller stances than their wild-type littermates (Figure 3.13A), the opposite of what was noted for the previous cohorts. Although these differences are statistically significant, it is difficult to interpret their biological relevance. However, we can conclude that this cohort of *Tars1* $^{G541R/+}$ mice does not display the hallmarks of neurodegenerative gait impairment. They are also not deficient at running on a wheel (Figure 3.13B), although the lack of a statistically significant signal may be partially due to the wide range of running activity of the wild-type mice. Finally, motor nerve conduction amplitude, sciatic motor nerve conduction velocity, and sural nerve conduction velocities were measured. Axonal peripheral neuropathy is marked by a primary defect in the axon, which causes decreased nerve conduction amplitudes.[411] This degeneration can also affect the integrity of the myelin sheath and cause secondary decreases in nerve conduction velocities as well.[411] However, consistent with the absence of a behavioral phenotype, *Tars1* $^{G541R/+}$ mice did not show any electrophysiological defects in nerve conduction studies.

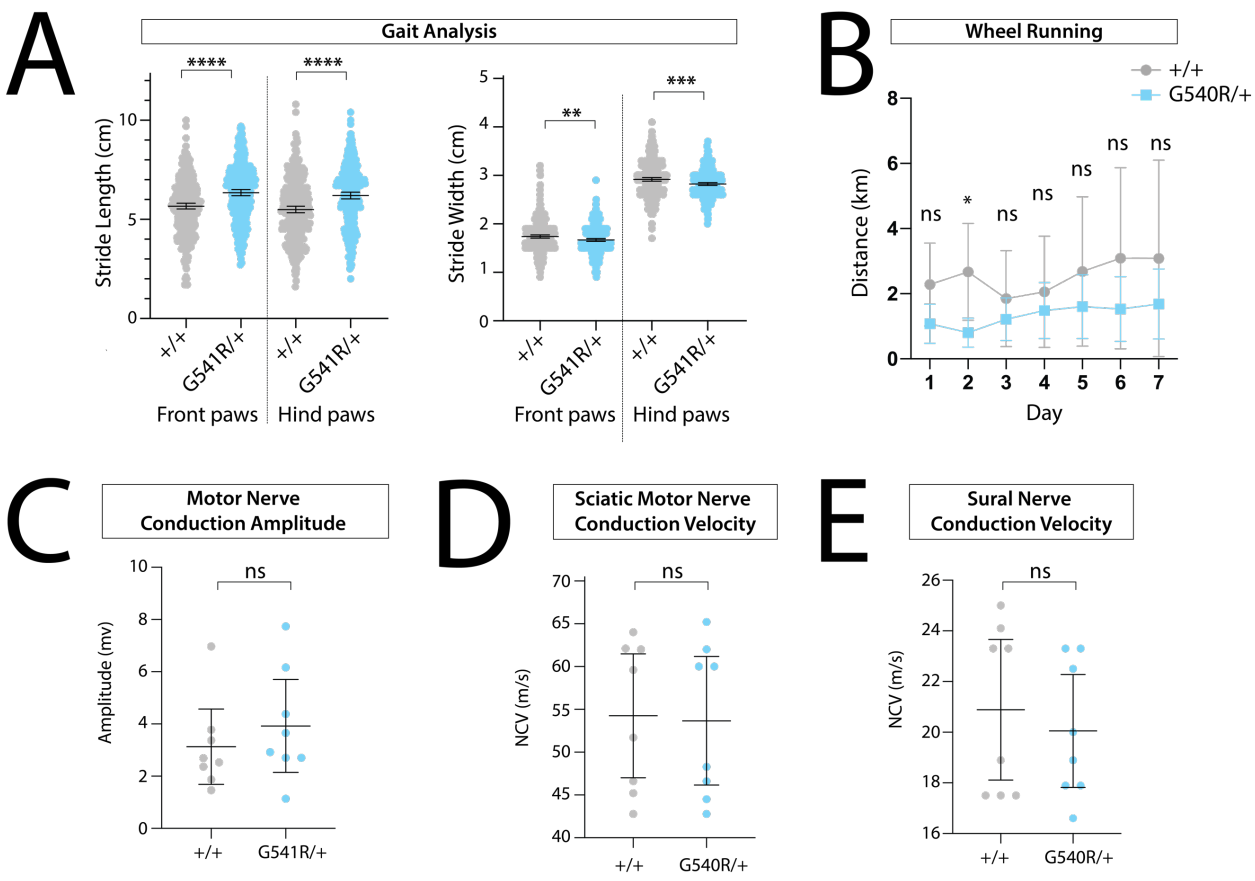


Figure 3.13. Behavior and nerve conduction analysis for *Tars1*$^{G541R/+}$ mice at 12 months of age.

(A) Gait analysis for 10 male *Tars1*$^{G541R/+}$ mice and 10 male *Tars1*$^{+/+}$ littermates at 1 year of age. For stride length and width measurements, 330-353 front paw steps and 330-353 hind paw steps were analyzed for each genotype. **(B)** Distance traveled on a wheel, measured for 7 consecutive days. **(C)** Nerve conduction amplitudes in the motor nerves of *Tars1*$^{G541R/+}$ and *Tars1*$^{+/+}$ mice **(D)** Nerve conduction velocities for the sciatic motor nerves of *Tars1*$^{G541R/+}$ and *Tars1*$^{+/+}$ mice **(E)** Nerve conduction velocities for the sural nerves of *Tars1*$^{G541R/+}$ and *Tars1*$^{+/+}$ mice. All bars represent the mean value, with 95% confidence intervals. For all comparisons, an unpaired t-test with Welch's correct was performed to determine if the differences between *Tars1*$^{G541R/+}$ and *Tars1*$^{+/+}$ mice were statistically significant. **** $p<0.0001$, *** $p<0.001$, ** $p<0.01$, * $p<0.05$, ns = not significant.

3.3.10 Tars1$^{G541R/+}$ *mice have reduced Tars1 levels*

To determine if G541R *Tars1* reduced Tars1 protein levels, protein lysates were obtained from the brains of 11-week-old mice with +/+, G541R/+, and F538Kfs*4/+ genotypes. Two mice from each genotype were assessed. Total TARS1 abundance was analyzed with immunostaining. Here, we used an antibody that recognized both human TARS1 and mouse Tars1, and included lysate from yeast expressing human TARS1 as a size control (both human TARS1 and mouse Tars1 are predicted to be 83.4 kilodaltons). To more accurately compare protein abundance across genotypes, *Tars1*$^{+/+}$ lysates were included in a gradient of 6.25µg, 12.5µg, and 25µg. 25µg of Tars1$^{F538Kfs*4/+}$ lysate was detected at a similar intensity as 12.5µg of Tars1$^{+/+}$ lysate, consistent with decreased expression from the F538Kfs*4 allele (Figure 3.14) (One important caveat is that the TARS1 antibody recognizes an epitope between 346 and 657 amino acids, and may not be able to recognize a truncated protein at 541 amino acids.) Notably, 25µg of *Tars1*$^{G541R/+}$ lysate lead to an intermediate Tars1 band intensity, between that of *Tars1*$^{F538Kfs*4/+}$ and *Tars1*$^{+/+}$. This suggests that although G541R does not fully destabilize TARS1 protein, it is not expressed at wild-type levels. This may explain why over-expression of G541R caused neuronal phenotypes in worm, but endogenous levels of expression in worm or mouse did not.

3.3.11 P0 deaths are enriched for Tars1$^{R433H/F538Kfs*4}$ *mice*

Many individuals with recessive ARS-mediated disease are compound heterozygous for a hypomorphic missense allele and a null allele.[68] The above studies in yeast and worm indicate that R433H *TARS1* is a hypomorphic allele, and the mouse studies of F538Kfs*4 *Tars1* homozygous lethality indicate that it is a null allele. To investigate a genotype relevant to human patients, *Tars1*$^{R433H/R433H}$ mice were crossed to *Tars1*$^{F538Kfs*4/+}$ mice to obtain compound heterozygous offspring. One expectation was that these offspring would develop phenotypes similar to those seen in patients with other ARS-mediated recessive disease, such as microcephaly, developmental delay, or small stature.[126,221,233] Another expectation was that reduced *Tars1* function would preferentially affect tissues with a high requirement for threonine in proteins. To identify these proteins, we sorted the mouse proteome by threonine content and removed uncharacterized proteins, along with proteins with a Uniprot annotation score under three, which eliminated poorly annotated proteins from the list. It should be noted that this

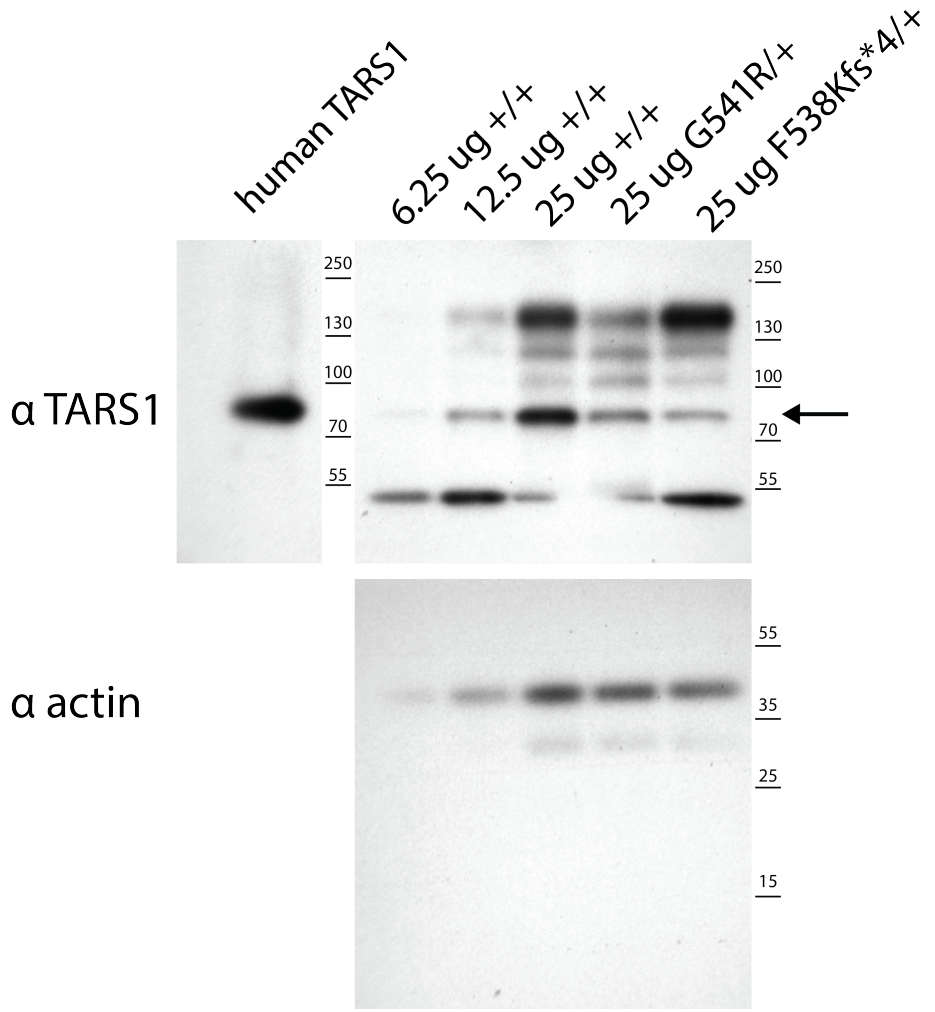


Figure 3.14. *Tars1*$^{G541R/+}$ mice show reduced Tars1 expression in brain tissue.

Representative western blot image for Tars1 protein in brain tissue of a *Tars1*$^{G541R/+}$ mouse, *Tars1*$^{+/+}$ mouse, and *Tars1*$^{F538Kfs*4/+}$ mouse. Human TARS1, expressed in yeast, is shown on the left as a size control. For *Tars1*$^{+/+}$ samples, 6.25µg, 12.5µg, and 25µg lysate was loaded. For *Tars1*$^{G541R/+}$ and *Tars1*$^{F538Kfs*4/+}$ samples, 25µg was loaded. The top blot shows signal from a TARS1 antibody; the bottom blot shows signal from an actin antibody, as a loading control. An arrow indicates the expected size of TARS1 and Tars1. The full blot is shown.

preliminary search may miss proteins with a lower overall threonine percentage but a localized high threonine content. The identified proteins composed of at least 15% threonine are primarily glycoproteins and mucins (Table 3.2).

Mucins are the primary structural component of the mucus membranes that coat all wet epithelial surfaces in the body, including the lungs and the stomach.[412] Mucin backbones are composed of tandem repeats of proline, threonine, and/or serine. These repeats are densely modified by O-linked glycosylation, which maintains the integrity of the mucous membrane.[412] The importance of threonine to these proteins is also supported by previous reports that intestinal mucin synthesis can be improved with threonine supplementation[413] or impaired with threonine reduction.[414] Threonine deprivation and knockdown of *TARS1* can decrease translation of a mucin protein (MUC1) in pancreatic cancer cells, without affecting global translation[415]. Therefore, it is possible that mice with reduced Tars1 function will be unable to incorporate sufficient threonine into mucin proteins to support proper mucous formation in tissues like the lungs or the gut.

Offspring of the *Tars1*$^{R433H/R433H}$ and *Tars1*$^{F538Kfs*4/+}$ mice were genotyped at 3 weeks of age. Based on Mendelian segregation, 50% of offspring were expected to be R433H/+ and 50% to be R433H/F538Kfs*4. However, there was a significant depletion of R433H/F538Kfs*4 at this stage of life (Figure 3.15A), indicating decreased viability prior to genotyping. Since no mortality was observed in developing pups, this death likely happens *in utero* or shortly after birth. In general, there is a base rate of neonatal mortality in laboratory mice, especially for the C57BL/6 strain.[416] However, an analysis of neonate deaths across four litters showed that pups that died at P0 were strikingly enriched for the R433H/F538Kfs*4 genotype. Out of 15 genotyped neonatal deaths, 13 were R433H/F538Kfs*4 mice (Figure 3.15B).

3.3.12 P0 Tars1$^{R433H/F538Kfs*4}$ *mice lack air in their lungs and PAS+ material in their bronchiolar club cell epithelia*

To gain insights into the reduced viability of *Tars1*$^{R433H/F538Kfs*4}$ mice, a cohort of four P0 *Tars1*$^{R433H/F538Kfs*4}$ pups and three age-matched *Tars1*$^{R433H/+}$ littermates was collected for histology studies. The four *Tars1*$^{R433H/F538Kfs*4}$ pups were all found dead within a few hours after

Table 3.2. Threonine-rich proteins in the mouse proteome.

Gene	Protein	% Threonine	Tissue with strongest expression[417]
Muc19	Mucin-19	22%	Submandibular gland
Muc4	Mucin-4	19%	Lung
Muc20	Mucin-20	18%	Lung, kidney
Muc6	Mucin-6	18%	Brain, liver, stomach, intestine
Timd4	T-cell immunoglobulin and mucin domain-containing protein 4	17%	Brain, liver, stomach, intestine
Cd164	Sialomucin core protein 24	17%	Kidney, liver, lung, pancreas, thymus
Emcn	Endomucin	17%	Lung, heart, kidney
Gp1ba	Platelet glycoprotein Ib alpha chain	17%	Platelets[418]
Plin4	Perilipin-4	16%	White adipose tissue[419]
Defb26	Beta-defensin	16%	Testis[420]
Prg4	Proteoglycan 4	16%	Liver
Nup62	Nuclear pore glycoprotein p62	15%	Widely expressed
Muc16	Mucin-16	15%	Eye, uterus, heart, lung, gastric tract, gall bladder[421]
Plet1	Placenta-expressed transcript 1 protein	15%	Colon, lung, salivary gland, keratinocytes, thymus[422]
Ovgp1	Oviduct-specific glycoprotein	15%	Oviduct[423]
Defb22	Beta-defensin 22	15%	Testis

A

Mice at 3 weeks of age

Genotype	Observed	Expected (Mendelian ratios)
R433H/+	36	25.5
R433H/F538Kfs*4	15	25.5

Chi-square test: p=0.0033

B

Neonatal mortality (0-1 days after birth)

Genotype	Observed	Expected
R433H/+	2	7.5
R433H/F538Kfs*4	13	7.5

Chi-square test: p=0.0045

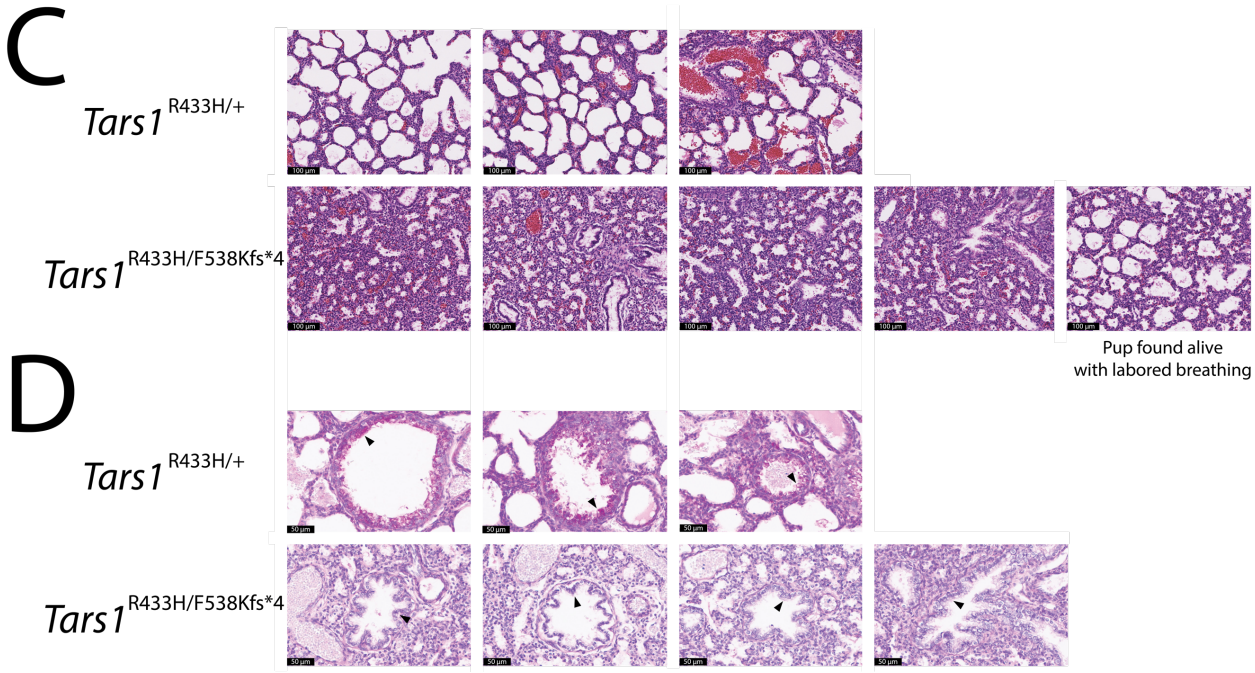


Figure 3.15. Neonatal lethality of *Tars1*$^{R433H/F538Kfs*4}$ mice.

(A) Genotype analysis of *Tars1*$^{R433H/R433H}$ and *Tars1*$^{F538Kfs*4/+}$ offspring, genotyped upon weaning at 3 weeks of age. The observed and expected number of each genotype is shown. **(B)** Genotype analysis of 15 deceased pups, identified within one day after birth. The observed and expected number of each genotype is shown. For **(A)** and **(B)**, a Chi-square test was used to determine if the difference between the number of observed and expected genotypes was statistically significant. **(C)** H&E staining of lung sections from three *Tars1*$^{R433H/+}$ P0 pups (top row) and five *Tars1*$^{R433H/F538Kfs*4}$ P0 pups (bottom row). All *Tars1*$^{R433H/+}$ pups were alive when identified at P0. The first four *Tars1*$^{R433H/F538Kfs*4}$ pups were identified dead at P0; the fifth was found alive with a gasping, labored breathing pattern. **(D)** PAS staining of lung sections from three *Tars1*$^{R433H/+}$ P0 pups (top row) and four *Tars1*$^{R433H/F538Kfs*4}$ P0 pups (bottom row). Black arrows highlight the magenta PAS signal in the bronchioles of *Tars1*$^{R433H/+}$ mice (top row), and the absence of PAS signal in the collapsed bronchioles of *Tars1*$^{R433H/F538Kfs*4}$ mice (bottom row).

birth. Interestingly, in a separate litter, a *Tars1*$^{R433H/F538Kfs*4}$ pup was found immediately after birth (traces of birth fluids were still visible) with visibly labored breathing and a failure to right itself. This additional pup was included in the cohort to investigate its breathing phenotype. All pups were fixed in formalin overnight, washed with 70% ethanol, and then sent for processing at Histoserv, where sagittal sections were taken for H&E staining and Periodic Acid Schiff (PAS) staining (which detects glycoproteins and mucins[407]).

The four P0 *Tars1*$^{R433H/F538Kfs*4}$ mice that were found dead had no air in their lungs. Whereas the alveoli of *Tars1*$^{R433H/+}$ mice were expanded with air (Figure 3.15 C, top), the alveoli of *Tars1*$^{R433H/F538Kfs*}$ mice were collapsed (Figure 3.15 C, bottom). Based on the otherwise mature body development of these pups, this indicates that they died either shortly before or upon birth. Interestingly, the additional *Tars1*$^{R433H/F538Kfs*4}$ pup found alive immediately after birth had only partially expanded alveoli, which correlates with the observed labored breathing. Moreover, although the bronchioles of *Tars1*$^{R433H/+}$ mice are replete with the magenta PAS+ signal of secretory cells, this signal is absent from the collapsed bronchioles of *Tars1*$^{R433H/F538Kfs*4}$ mice. (Figure 3.15D) Further work is needed to determine whether this is a result of an absent population of cells or a result of decreased glycoprotein production, or both (discussed further in Chapter 5).

3.3.13 Surviving Tars1$^{R433H/F538Kfs*4}$ *mice have reduced body weight*

Tars1$^{R433H/F538Kfs*4}$ mice that passed the P0 mark survived to adulthood with no apparent breathing abnormalities. However, these mice were on average smaller than their *Tars1*$^{R433H/+}$ littermates (Figure 3.16A). Male and female mice from nine litters were weighed once every week until 23 weeks of age. Only litters containing both genotypes were included in the analysis. Reduced body weight was more consistent in females (Figure 3.16B) than males, who reach a normal body size by 7 weeks of age (Figure 3.16C). This reduced size is consistent with growth restriction phenotypes in human patients with ARS-mediated recessive disease, as well as the reduced size of mice homozygous for a hypomorphic *Cars1* mutation, which is being investigated by another graduate student in the Antonellis group, Molly Kuo.

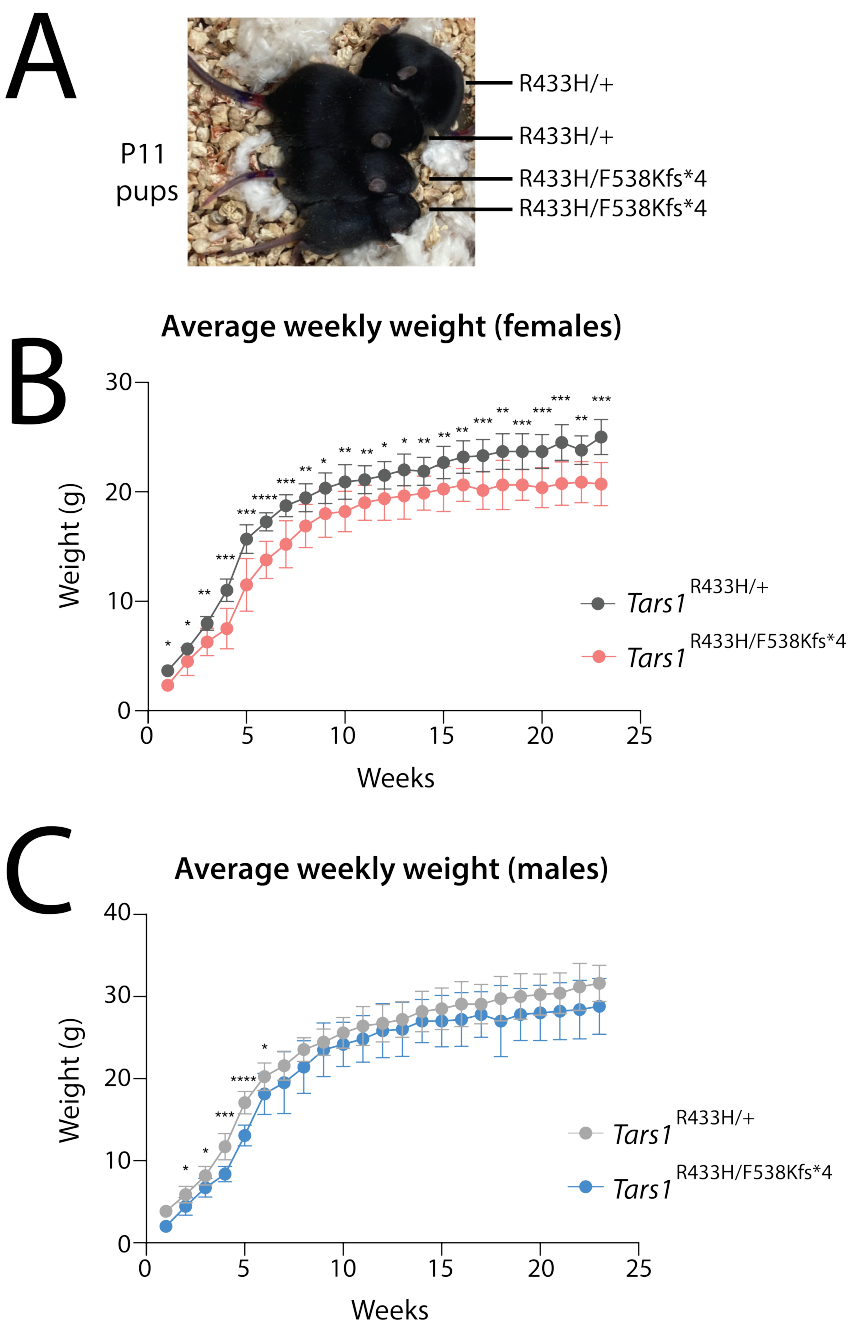


Figure 3.16. Reduced body weight of *Tars1*$^{R433H/F538Kfs*4}$ mice.

(A) Image of four littermates at P11, grouped together for comparison of body size. Each mouse is labeled with its genotype. **(B)** The average weekly weights of female *Tars1*$^{R433H/F538Kfs*4}$ mice and female *Tars1*$^{R433H/+}$ littermates are shown, until 23 weeks of age. **(C)** The average weekly weights of male *Tars1*$^{R433H/F538Kfs*4}$ mice and female *Tars1*$^{R433H/+}$ littermates are shown, until 23 weeks of age. For **(B)** and **(C)**, bars represent the mean value and one standard deviation. An unpaired t-test was performed for each week to determine if the difference between the two genotypes was statistically significant. **** $p<0.0001$, *** $p<0.001$, ** $p<0.01$, * $p<0.05$. All values in **(C)** that are not marked with an asterisk are not significantly different.

3.3.14 Tars1$^{R433H/F538Kfs*4}$ *mice display hair and skin defects*

As these studies were underway, a report of two human individuals with bi-allelic *TARS1* variants and triochothioydstrophy (TTD) phenotypes was published.[253] Both individuals presented with the classic "tiger-tail" banding pattern of the hair shaft and ichthyosis. One individual had follicular keratosis, the other was born as a "collodion baby", encased in a tight shiny membrane.[253] Interestingly, the described "collodion baby" phenotype was reminiscent of two *Tars1*$^{R433H/F538Kfs*4}$ mice that were found dead shortly after birth. The bodies of these mice were encased in a tight membrane (Figure 3.17 A and B) with visible blisters (right image). One mouse (right image) was submitted for histopathology as described above; however, there was no immediately apparent signatures of a mouse collodion baby phenotype, such as a thickened stratum corneum. Further investigation is required to fully define this specific phenotype. Interestingly, the *Tars*$^{R433H/F538Kfs*4}$ P0 pups did show evidence of other skin and hair abnormalities. In the cohort of four *Tars*$^{R433H/F538Kfs*4}$ and three *Tars*$^{R433H/+}$ littermates, *Tars*$^{R433H/F538Kfs*4}$ mice had a thinner epidermal layer than control mice (Figure 3.17 B and C); on average, the epidermis was 35% thinner. Additionally, these four mice had follicular hypoplasia (*i.e.* less mature and/or fewer hair follicles).

Adult *Tars1*$^{R433H/F538Kfs*4}$ mice also displayed a striking hair phenotype. Of the mouse cohort that was followed and weighed over the course of 23 weeks, 10 out of 14 *Tars1*$^{R433H/F538Kfs*4}$ mice (71.4%) lost hair on their heads and/or upper back by 23 weeks of age, compared to 1 out of 23 *Tars1*$^{R433H/+}$ littermates (4.35%). Hair loss onset occurred between 13 and 23 weeks of age (Figure 3.18A). It followed a stereotypic pattern of bald spots on the head and/or along the scapula of the upper back (Figure 3.18B). In more advanced stages, it would reach across the entire upper back (Figure 3.18C). To more thoroughly define this phenotype, histopathology was performed on hair samples from the affected regions for three *Tars1*$^{R433H/F538Kfs*4}$ mice and three *Tars1*$^{R433H/+}$ littermates; one pair was 2 months old, another was 12 months old, and the third was 14 months old. Analysis of H&E staining did not reveal any gross abnormalities in hair follicles, although this analysis was complicated by the asynchronous hair cycling of adult mice. Further work is required to define this phenotype (discussed in Chapter 5).

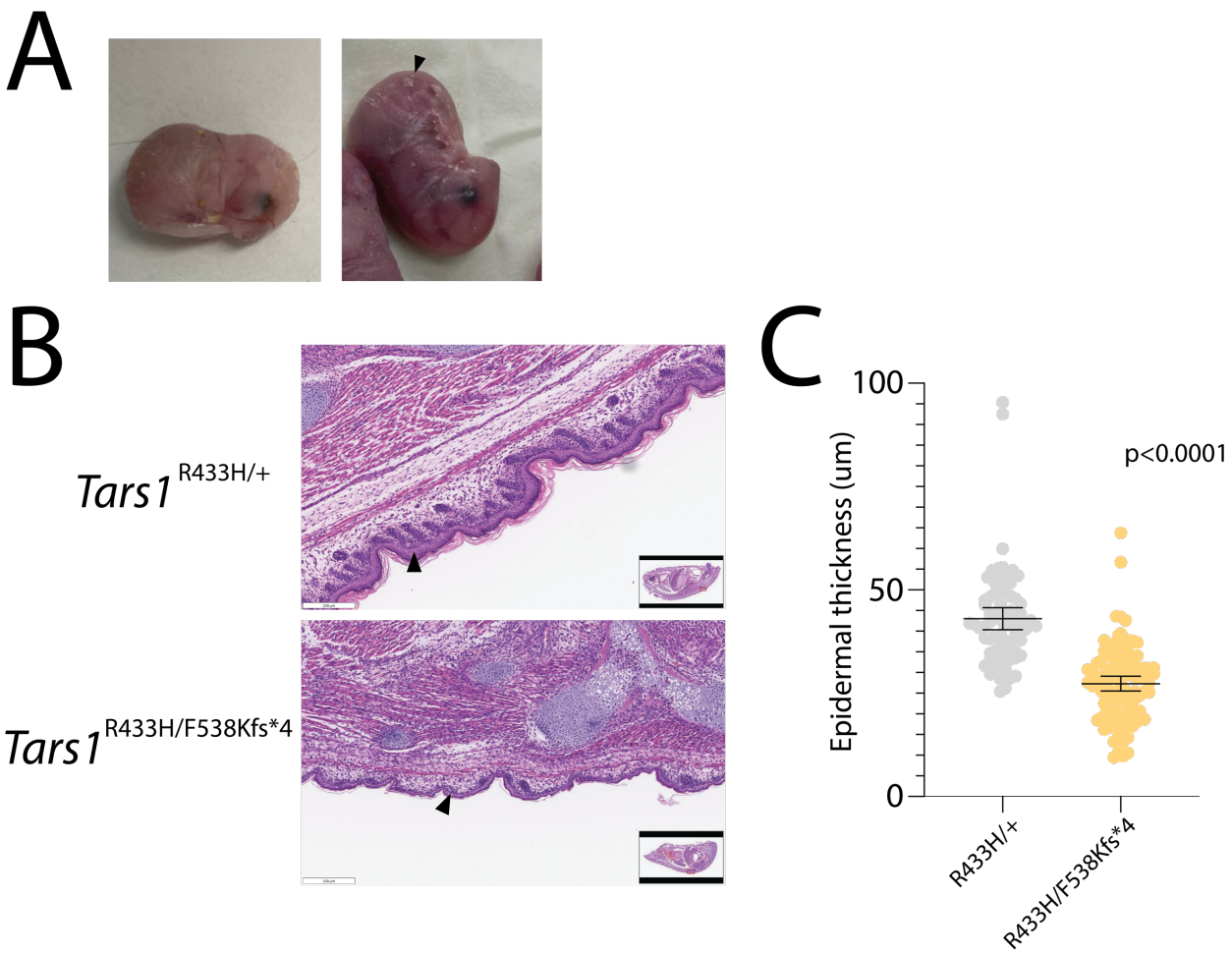


Figure 3.17. Skin abnormalities in *Tars1*$^{R433H/F538Kfs*4}$ P0 pups.

(A) Images of two *Tars1*$^{R433H/F538Kfs*4}$ pups born encased in a tight membrane. A black arrow points to a blister on the right-hand mouse. **(B)** H&E staining of dorsal skin sections in P0 pups. Black arrows point to the dark pink epidermal layer. The upper image shows skin from a *Tars1*$^{R433H/+}$ mouse, and the bottom image shows skin from a *Tars1*$^{R433H/F538Kfs*4}$ mouse. **(C)** Measurements of epidermal thickness of four *Tars1*$^{R433H/F538Kfs*4}$ P0 Pups and three *Tars1*$^{R433H/+}$ P0 littermates. The mean epidermal thickness is decreased by 35% in *Tars1*$^{R433H/F538Kfs*4}$ mice. Bars indicate the mean value and 95% confidence interval. Statistical significance was determined with an unpaired t-test with Welch's correction. N=75 for R433H/+ values and n=100 for R433H/F538Kfs*4 values.

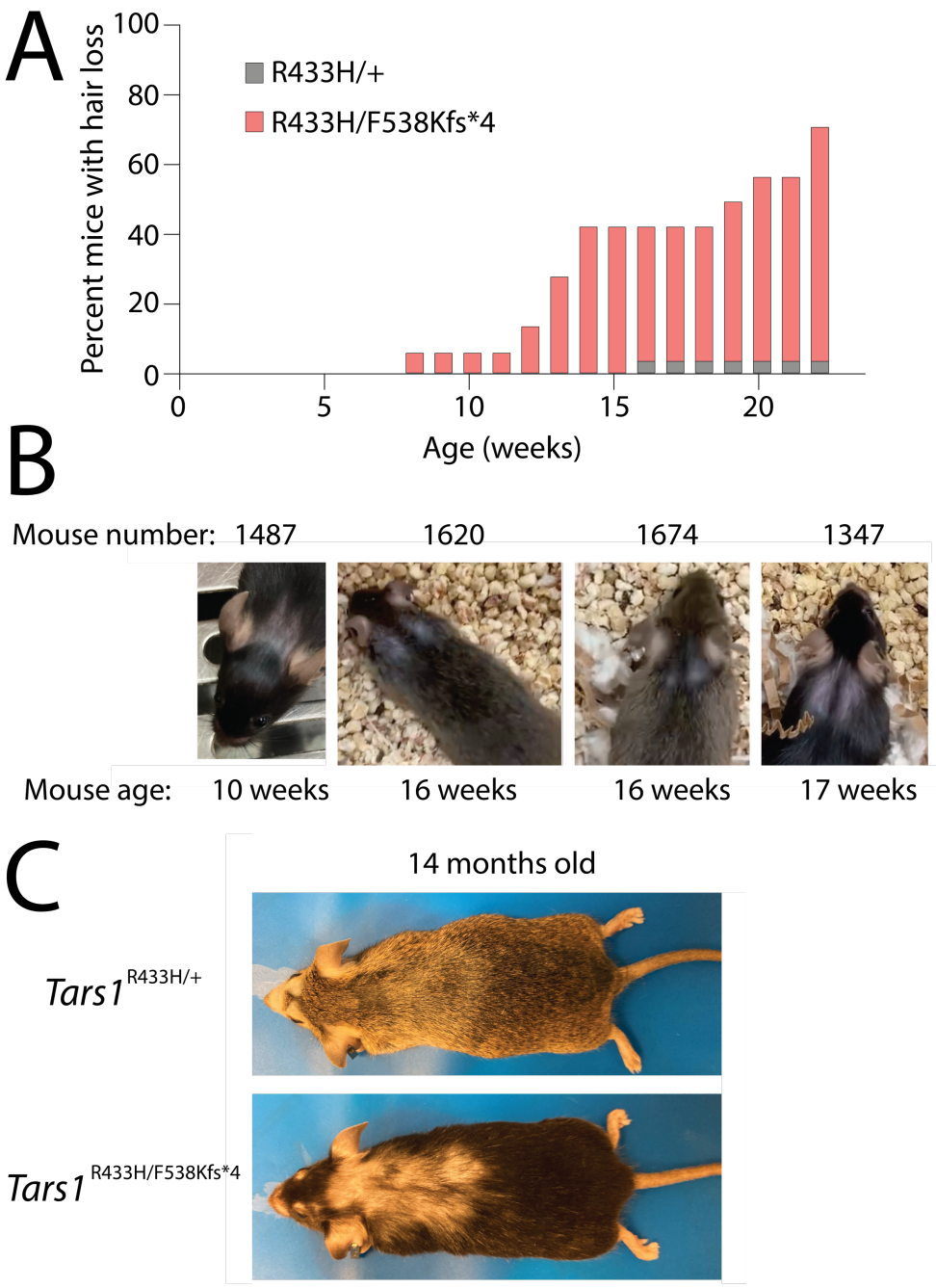


Figure 3.18. Adult onset hair loss in *Tars1*$^{R433H/F538Kfs*4}$ mice.

(A) The cumulative percentage of *Tars1*$^{R433H/F538Kfs*4}$ mice (pink) and *Tars1*$^{R433H/+}$ (gray) mice with hair loss on the back of their head or upper back is shown, until 23 weeks of age. **(B)** Representative images of four individual *Tars1*$^{R433H/F538Kfs*4}$ mice with hair loss; the consistent pattern of upper back bald patches is shown for #1620, 1674, and 1347. **(C)** (Bottom image) A representative image of extended hair loss stretching from the head to the middle of the back in a *Tars1*$^{R433H/F538Kfs*4}$ mouse, 14 months of age. His *Tars1*$^{R433H/+}$ littermate is shown above, with signs of barbering by the nose and mild age-related hair thinning on the back.

3.4 Discussion

The work presented in this Chapter aimed to use the established characteristics of pathogenic ARS variants in model organisms (*i.e.,* loss-of-function in yeast, dominantly toxic to *C.* elegans neurons) to build a pipeline for predicting the pathogenicity of variants in *TARS1* for both recessive and dominant phenotypes. Through this process, we successfully identified a hypomorphic *TARS1* allele, R433H. R433H reduced yeast growth in a complementation assay, caused developmental delay and thrashing defects in homozygous *C. elegans*, and caused premature death linked to lung failure, decreased body size, and hair defects in mouse when modeled *in trans* with a null allele. This pipeline also identified a candidate dominant variant, G541R *TARS1*, that was loss-of-function in a yeast complementation assay, was homozygous lethal to both worm and mouse, and was dominantly toxic to *C. elegans* neurons in an over-expression system. However, G541R *TARS1* did not cause reproducible neuromuscular defects in heterozygous worms or mice, likely due to insufficient abundance of the mutant protein in relevant tissues. Therefore, it is not possible to draw conclusions about *TARS1* as a candidate for dominant peripheral neuropathy based on this specific variant.

Reduced levels of the G541R Tars1 protein is the most likely reason that heterozygous worms and mice did not develop neurological phenotypes. In general, pathogenic dominant ARS variants do not significantly reduce protein abundance, to the extent that this can be detected in patient fibroblasts or lymphoblasts *in vitro*.[170,301,348] However, it is also useful to consider the limitations of mice when modeling dominant axonal neuropathies. Mouse peripheral neurons are much shorter than humans, and may not be as sensitive to a variant that produces a late-onset, mild neuropathy in humans. For example, mouse models of patient variants in *HSPB1* and *MFN2* (other dominant axonal Charcot-Marie-Tooth disease genes) show a significant discrepancy between transgenic models with high expression levels and knock-in variants with endogenous expression levels. While transgenic models develop multiple symptoms of peripheral neuropathy by 6 months of age,[373,424,425] mice expressing endogenous levels of the pathogenic allele show no significant signs of neuromuscular defects as late as 18 or 20 months of age.[426,427]

If G541R *TARS1* does not produce fully stable TARS1 protein, this could also explain the discrepancy seen between the two worm models of G541R *tars-1*. Over-expressing G541R in GABA-ergic neurons may have compensated for the reduced stability of the protein and allowed it to reach toxic levels that caused morphological defects. Endogenous levels of expression may not have been sufficient to produce these defects. Another possibility is that the axons in worm are not long enough to be sensitive to endogenous expression levels of a toxic ARS allele; further work with established pathogenic ARS alleles is required to determine if these morphological phenotypes can be reproduced with knock-in models. It will also be informative to determine if homozygosity for partial loss-of-function alleles can also produce these aberrant morphologies. If so, this would demonstrate that a loss of ARS function can cause morphological defects in GABA-ergic neurons, which would support a dominant-negative mechanism for the dominant alleles that produce these same defects. However, homozygosity R433H *tars-1* did not produce this phenotype (Figure 3.7C), indicating that testing this hypothesis may require variants with a more significant impact on ARS function.

Although modeling G541R in mouse *Tars1* did not provide evidence that *TARS1* alleles can cause dominant peripheral neuropathy, the homozygous lethal nature of G541R (and of F538Kfs*4) answers a different critical question. Threonyl-tRNA synthetase is unique among the ARS gene family; in addition to the gene encoding the cytoplasmic enzyme (*TARS1*) and the gene encoding the mitochondria enzyme (*TARS2*), there is a third paralog, *TARS3 (Tarsl2* in mouse). *TARS3* likely arose from a duplication event in eukaryotic history, and is found across mammalian species, as well as in birds and houseflies.[428] Mouse *Tars1* and *Tarsl2* share 76.4% identity across all domains except the N-terminal extension, which is only 10.3% identical.[429] Although *Tarsl2* has not been fully characterized, it is reported to be ubiquitously expressed in mouse, localized to both the cytoplasm and the nucleus, and capable of charging tRNA with threonine *in vitro*.[429] Although there is no known genetic redundancy for any cytoplasmic or mitochondrial ARS, redundancy between *TARS1* and *TARS3* had not been tested prior to this dissertation. If there was redundancy, we would not expect *TARS1* variants to cause recessive disease. Our studies address this question by demonstrating that homozygosity for either of two loss-of-function *Tars1* alleles (G541R or F538Kfs*4) is incompatible with life, indicating that *Tarsl2* cannot fully compensate for loss of *Tars1*. It remains to be seen whether *Tarsl2* might partially compensate, perhaps extending the

viability of homozygous null *Tars1* embryos *in utero*. Regardless, if this threonyl-tRNA synthetase duplication has remained intact over mammalian evolution but is not redundant with *Tars1*, it may have acquired other non-canonical activities in the cell. The function of *TARS3* should be the subject of future investigation.

Finally, we identified a hypomorphic *Tars1* allele (R433H) and, after determining that it reduced gene function in yeast and worm, used it to generate a mouse model of *Tars1*-mediated recessive phenotypes. Preliminary investigation of mice that are compound heterozygous for R433H and the null allele F538Kfs*4 revealed that these mice frequently die at birth without air in their lungs. *Tars1*$^{R433H/F538Kfs*4}$ mice that survive are smaller than their littermates and develop hair loss in consistent patterns on their heads and upper back. Although additional work is required to further define these phenotypes, this model will be a valuable resource to investigate the effects of reduced *Tars1* function across different cell populations, and provide mammalian tissues to investigate how reduced *Tars1* function impacts protein translation. These future studies can inform the assessment of patients with recessive *TARS1*-mediated disease and aid efforts to develop therapies for ARS-mediated recessive diseases more broadly.

One possibility is that proteins with a particularly high threonine content, such as mucins, are especially sensitive to decreased *Tars1* activity. This could lead to defects in the tissues that rely heavily on these proteins, such as the lung. Interestingly, the gut is also dependent on mucin synthesis.[412] Although preliminary investigation of gut histology in P0 mice did not identify any changes in PAS signal, a careful analysis of gut mucin production should be included in future characterizations of this mouse. Another possibility is that decreased *Tars1* activity reduces global protein translation by triggering phosphorylation of eIF2α. This might affect cells with a high demand for protein translation, such as differentiating stem cells. For example, if aging hair follicle stem cells cannot properly translate the proteins required for differentiation, this may explain a failure to regrow hair in the adult *Tars1*$^{R433H/F538Kfs*4}$. These possibilities are further discussed in Chapter 5.

Overall, research described in this Chapter has demonstrated the efficacy of using a tiered model organism approach to predict the pathogenicity of *TARS1*. Although it is still unclear whether

variants in *TARS1* can cause dominant peripheral neuropathy, the lessons learned from studying G541R *TARS1* will pave the way for future iterations of this process. Additionally, this work provides a template for predicting the pathogenicity of any ARS gene that has not yet been implicated in dominant or recessive disease, and for building the appropriate model systems to further investigate the mechanisms of these dominant or recessive phenotypes.

Chapter 4
Testing Neuropathy-Associated *AARS1* Alleles for a Dominant-Negative Effect

4.1 Introduction

Hereditary peripheral neuropathies are a group of phenotypically and genetically heterogeneous clinical phenotypes. These diseases are characterized by decreased sensory and/or motor neuron function in the distal extremities. This leads to sensory loss and muscle atrophy, which often begins in the feet and lower leg muscles and may progresses to include the hands and forearms of the upper extremities.[285] If a genetic peripheral neuropathy is restricted to impaired motor neuron function, it is classified as distal hereditary motor neuropathy (dHMN).[295] If the symptoms include both sensory and motor neuron dysfunction, it is classified as Charcot-Marie-Tooth (CMT) disease.[285] CMT disease can arise from a primary defect in the myelinating Schwann cells of the peripheral nervous system, which is classified as CMT Type I.[287] These defects are usually caused by mutations in genes that are important for myelin production or function, such as *PMP22*, which encodes a critical component of the myelin sheath; mutations in *PMP22* account for over half of CMT disease cases.[430] Alternately, CMT disease can be caused by a primary defect in the axon of the peripheral neuron, classified as CMT Type 2. The most common form of CMT Type 2 is caused by mutations in *MFN2,* which accounts for 20-30% of CMT Type 2 cases.[431] Like ARS genes, *MFN2* is a ubiquitously expressed, essential gene required for a basic cellular function (mitochondrial fusion).[290] Peripheral neurons are thought to be particularly sensitive to defects in *MFN2*, because mitochondrial function is required across the long axons of the peripheral neurons, including in distal regions of the axon like dendrites and synapses.[290]

Mutations in five aminoacyl-tRNA synthetases have been linked to dominant peripheral neuropathies: alanyl-(*AARS1*),[106] histidyl-(*HARS1*),[174] glycyl-(*GARS1)*,[160] tryptophanyl-(*WARS1*),[264] and tyrosyl-(*YARS1*) tRNA synthetases.[432] In addition, methionyl-tRNA synthetase (*MARS1*) variants have been identified in patients with CMT disease[216]; however, there is not convincing genetic evidence for pathogenicity of these *MARS1* alleles. It remains to be seen how many additional ARS genes will be implicated in dominant peripheral neuropathy. Defining the locus and allelic heterogeneity of this disorder will be critical both for patient diagnosis and defining disease mechanisms.

As discussed in Chapter 1, bi-allelic ARS variants that reduce enzyme function cause recessive disorders that are early-onset and that affect multiple tissues.[68,377] In some cases, the constellation of recessive phenotypes includes peripheral neuropathy,[126,433,222] demonstrating that peripheral neurons are sensitive to reduced ARS function. However, this reduction of ARS function is likely greater than 50%, because heterozygosity for a null ARS allele is not sufficient to cause a highly penetrant peripheral neuropathy—null alleles are not found in CMT patient populations, but are found in individuals who, to the best of our knowledge, are unaffected. Additionally, mice that are heterozygous for a *Gars* null allele do not develop a peripheral neuropathy.[69] Based on all of these observations, haploinsufficiency is unlikely to be the disease mechanism for ARS-associated dominant neuropathy.

The pathogenic ARS variants linked to dominant peripheral neuropathy are exclusively missense mutations or small in-frame deletions; the absence of frameshift alleles and premature stop codons indicate that an expressed mutant protein is required for pathogenicity. It is possible that these missense alleles act as neomorphs, exposing novel binding interfaces that facilitate aberrant protein interactions and lead to dysregulated neuronal pathways (Section 1.3.5). However, if there is a common mechanism to explain the role of all five ARS enzymes implicated in dominant neuropathy, it is unlikely that all mutations will cause the same neomorphic interaction. Rather, any common mechanism would likely be related to the shared canonical role in tRNA charging.

One notable commonality is that all five enzymes function as homodimers in the cytoplasm. This raises the possibility of a dominant-negative mechanism, in which decreased function of the mutant subunit reduces the function of the wild-type subunit in the dimeric holoenzyme; this would lower the overall ARS activity in the cell below 50%. A dominant-negative mechanism requires: 1) the mutant allele to reduce protein function, 2) the mutant allele to be stably expressed, and 3) the mutant allele to interact with the wild-type allele, repressing its function. This mechanism is supported by an abundance of data showing that the majority of neuropathy-associated ARS variants reduce enzyme activity.[110,176,307,170,171,264,434] Pathogenic variants also do not appear to affect the abundance of ARS protein detected in patient cells,[170,301,333,347] nor do they abolish dimerization[176,264,300,350432]; these observations are also consistent with a dominant-negative mechanism. Furthermore, models have shown that pathogenic ARS variants impair protein translation,[264,306,350] which is consistent with the expected decrease in tRNA charging resulting from a dominant-negative effect. However, no study has adequately, explicitly addressed a dominant-negative mechanism for ARS-related dominant neuropathy. Addressing this question will test for a unifying mechanism of disease for all five implicated dimeric ARS enzymes. It will also provide a relevant framework to assess the pathogenicity of newly identified variants in patients with ARS-related CMT disease.

Here, we generated a yeast model to test human *AARS1* variants for a dominant-negative effect. We focused on well-characterized alleles in two critical domains of the enzyme: the anti-codon binding domain (R329H)[106] and the amino acid activation domain (G102R).[109] We found that R329H and G102R *AARS1*, as well as three additional *AARS1* variants, are dominantly toxic to yeast growth when co-expressed with wild-type *AARS1*. We then engineered a dimer-disrupting variant in the C-terminal domain and expressed it *in cis* with the pathogenic *AARS1* variants. These double-mutants rescued the impaired yeast growth, demonstrating that the dominant toxicity of mutant AARS1 is dependent on dimerization with wild-type AARS1, and that these *AARS1* variants can be classified as dominant-negative (or antimorphic) alleles.

The author performed all experiments presented in this Chapter, with the important exception of assessing R326W, R329C, and R329S *AARS1* for a dominant-negative effect (Figures 4.9 and

4.10), which was performed by Sheila Marte. Additionally, Dr. Mike Shy and Shawna Feely contributed clinical and genetic information on four additional individuals with *AARS1* variants.

4.2 Materials and methods

4.2.1 Yeast complementation and dominant toxicity assays

All yeast assays were performed using the ptetO7-*ALA1* strain from the Yeast Tet-Promoters Hughes Collection (YSC1180-202219317, Horizon Discovery). *AARS1* variants were generated using site-directed mutagenesis (Agilent QuikChange II XL Site-Directed Mutagenesis Kit) against the *AARS1* open reading frame in pDONR221 (primer sequences found in the Appendix A). All clones were verified via Sanger sequencing to ensure that the desired mutation was generated and that no amplification errors were present. The Gateway cloning (Invitrogen) LR reaction was used to recombine the wild-type or mutant *AARS1* locus into pAG425GAL-ccdB (Addgene #14153), which is a Gateway-compatible vector with a 2-micron origin of replication that produces a high vector copy number, a *GAL1* promoter to drive high expression of the target gene in a galactose-inducible fashion, and a *LEU2* auxotrophic marker.

To assess the function of *AARS1* variants independent of wild-type *AARS1*, a p413 vector (ATCC #87370) with no *AARS1* insert ('Empty') was introduced into ptetO7-*ALA1* using lithium acetate yeast transformation. The p413 vector contains a *ADH1* promoter to drive constitutive expression of the target gene, a centromeric origin of replication to produce a low plasmid copy number per cell, and a *HIS3* auxotrophic marker for selection. This transformation was followed by pAG425 expressing wild-type or mutant *AARS1*. Colonies were grown on media lacking histidine and leucine (DO Supplement -His/-Leu, Takara Bio) to select for the presence of both vectors. After transformation, colonies were grown in 2mL liquid media in a 14mL round-bottom conical tube (Fisher Scientific) for two days at 30°C, shaking at 275 rpm until saturated. Yeast were then diluted to 1:10, 1:100, and 1:1000 in water. 10µl of serial dilutions were spotted on plates containing glucose, galactose/raffinose (Takara Bio Minimal SD Bases), or galactose/raffinose with 10µg/ml doxycycline (Fisher Scientific BP26531). Plates were imaged after four days of growth.

To assess the dominant toxicity of *AARS1* variants in the presence of wild-type *AARS1*, Gateway Cassette C (Invitrogen) was cloned into the p413 vector just downstream of (3' to) the *ADH1* promoter using the *Sma*I restriction site, then sequence verified to confirm correct orientation. The LR Gateway reaction was used to recombine wild-type *AARS1* from pDONR221 into p413. This construct was transformed into the ptetO7-*ALA1* strain, followed by *AARS1* (wild-type or mutant) in the pAG425 vector. Yeast were grown and spotted as detailed above. Complementation and dominant toxicity assays were performed side-by-side with the same pAG425 plasmid aliquots to enable direct comparisons.

4.2.2 Yeast protein isolation

The ptetO7-*ALA1* strain was transformed with mutant *AARS1* in pAG425 and grown on media lacking leucine (DO Supplement -Leu, Takara Bio). One colony was picked and placed into 3mL media and grown for 2-3 days shaking at 275rpm at 30°C until saturated, reaching an optical density (OD_{600}) of approximately 2. Yeast were then centrifuged at 1000xg for 10 minutes, washed once with water, transferred to a 1.5mL Eppendorf tube, then centrifuged at 15,000 rpm for 1 minute. The supernatant was removed and the pellet was stored at -80°C. The pellet was thawed in 150µl yeast lysis buffer (50mM Na-HEPES pH 7.5, 100mM NaOAc, 1mM EDTA, 1mM EGTA, 5 mM MgOAc, 5% glycerol, 0.25% NP-40, 3 mM DTT) with 1X Halt Protease Inhibitor Cocktail (Thermo Fisher Scientific). Approximately 100µl of 0.5mm cold glass beads (Biospec Products) were added to each sample. Samples were vortexed at 4°C for three minutes, followed by two minutes resting on ice, followed by three additional minutes of vortexing at 4°C. To remove the lysate from the beads, a 26-gauge needle (BD PrecisionGlide) was used to make a hole in the bottom of the 1.5mL tube, which was then immediately inserted into a 14mL round bottom conical tube. Lysates were centrifuged at 200xg at 4°C for 5 minutes. The lysates were collected from the bottom of the conical tube and transferred to a 1.5mL Eppendorf tube, then were centrifuged at 13,200 rpm for 5 minutes at 4°C. Supernatants were collected for measurement using the Thermo Scientific Pierce BCA Protein Assay kit, and 50µg of protein per sample was analyzed by Western blot (see below).

4.2.3 Co-immunoprecipitation of wild-type AARS1 and mutant AARS1

The LR Gateway reaction was used to recombine the wild-type or mutant *AARS1* open reading frame from pDONR221 into pDEST40 (Thermo Fisher Scientific) or pTM3xFLAG (gift from Moran Laboratory, University of Michigan). These vectors allowed differential tagging of the mutant and wild-type *AARS1* alleles; wild-type AARS1-3xFLAG was expressed from pTM3xFLAG using a CMV promoter, and either wild-type or mutant AARS1-6xHis was expressed from pDEST40 using a CMV promoter. 100mm plates (Falcon) were seeded with 1.5-2 million HEK293T cells; the following day, these were transfected with 0.5pmol plasmid using Lipofectamine 3000 (Invitrogen). 48 hours after transfection, cells were harvested using Trypsin-EDTA (Gibco, Fisher Scientific) and centrifuged at 2000rpm for 2 minutes at 4°C. Cells were then washed once with 1X PBS (Thermo Fisher Scientific), centrifuged again (as above), and then resuspended in 1mL lysis buffer (20mM Tris-HCl pH 8, 137mM NaCl, 2mM EDTA, 1%NP-40, 0.25% sodium deoxycholate) with 1X Halt Protease Inhibitor Cocktail (Thermo Fisher Scientific) for 2 hours rocking at 4°C. Samples were then centrifuged for 15 minutes at 13,200 rpm at 4°C. The supernatant was collected and protein concentration was measured using the Thermo Scientific Pierce BCA Protein Assay kit.

To conjugate beads with individual antibodies, 25µl of Dynabeads Protein G (Fisher Scientific) were aliquoted per sample. All immunoprecipitations were performed using a MagnaRack (Invitrogen). Each aliquot was washed twice with 500µl Conjugation Buffer (0.5% BSA, 0.1% Triton X-100 in PBS), then suspended in 500µl Conjugation Buffer with 2µg 6xHis antibody (abcam 18184) or 2µg FLAG antibody (BioLegend 637302). Beads and antibody were incubated overnight rocking at 4°C.

Prior to immunoprecipitation, lysates were pre-cleared to remove any proteins with non-specific affinity for the magnetic beads. An additional 25µl of Dynabeads per sample was aliquoted and washed once with lysis buffer. Then, 1mg of cell lysate in 500µl lysis buffer was added to the beads and rocked at 4°C for two hours. Supernatant from the antibody-conjugated beads was then removed, and the pre-cleared lysates were added. Samples were incubated for 3 hours rocking at 4°C. For anti-6xHis IPs, samples were washed four times with 1mL High Salt Buffer

(10mM Tris-HCl pH 7.5, 400mM NaCl, 1 mM EDTA, 1mM EGTA, 0.5% NP-40). For anti-FLAG IPs, samples were washed three times with 1mL Low Salt Buffer (10mM Tris-HCl pH 7.5, 137mM NaCl, 1 mM EDTA, 1mM EGTA, 0.5% NP-40). On the last wash, samples were moved to fresh 1.5 mL tubes to prevent co-elution of proteins bound to the tube walls. Samples were re-suspended in 50µl wash buffer with 50µl 2x Tris Glycine Buffer (Invitrogen). 4µl BME was added before samples were boiled at 99°C for 5 minutes and the supernatant was collected for Western blot (see below). Samples were divided in half and loaded in duplicate for immunoblotting with anti-AARS1, anti-6xHis, or anti-FLAG.

4.2.4 Disuccinimidyl suberate crosslinking experiments

To determine the degree of AARS1 dimerization in patient cells, AARS1 protein was crosslinked with disuccinimidyl suberate (DSS) and analyzed by Western blot. Patient and control fibroblasts were grown at 37°C in 5% CO_2 and standard growth media (DMEM supplemented with 10% FBS, 2mM L-glutamine, 100U/mL penicillin, and 50ug/mL streptomycin [Invitrogen]). Approximately 1 million cells were harvested from each sample with Trypsin-EDTA (Gibco, Fisher Scientific) and centrifuged at 2,000 rpm for 2 minutes at 4°C. They were then washed once with 1X PBS (Thermo Fisher Scientific), transferred to a 1.5mL tube, and centrifuged again (as above). Cells were then re-suspended in 50mM HEPES 0.5% NP-40. The sample was divided in two, and 50mM DSS (Thermo Fisher Scientific) was added to one aliquot to a final concentration of 5mM. Both aliquots were incubated at room temperature for 30 minutes. The crosslinking reaction was then quenched with a final concentration of 30mM TrisCl pH 7.5 at room temperature for 15 minutes. Samples were centrifuged at 13,200 rpm for 10 minutes at 4°C, and the supernatant was collected for Western blot analysis (see below). 20µg protein was analyzed for each sample. AARS1 antibody (Bethyl Laboratories A303-473A) was used at a dilution of 1:500.

4.2.5 Co-immunoprecipitation of wild-type ALA1 and wild-type AARS1

To investigate an interaction between yeast ALA1 and human AARS1, co-immunoprecipitation experiments were performed. First, the endogenous yeast *ALA1* coding sequence was amplified

from a previously published[107] pDONR221 clone with or without a C-terminal 6xHis tag encoded in the reverse primer (see primer sequences in Appendix A). Then, Gateway cloning was used to recombine these constructs into p413 (see above). The ptetO7-*ALA1* strain was transformed with p413 to express either 6xHis-tagged or untagged *ALA1,* then subsequently transformed with pAG425 to express either R329H or G757* human *AARS1*. Colonies were grown for 2-3 days until saturated in -leu -his liquid glucose growth medium, then washed in water and re-suspended in 125-250mL -leu -his galactose liquid culture (Takara Bio Minimal SD Bases, Takara Bio DO Supplement -His/-Leu). Cultures were grown to saturation, then centrifuged at 1000xg at 4°C for 20 minutes. Yeast were washed with water and aliquoted evenly into 4-5 1.5mL tubes, where they were centrifuged at 15,000 rpm for 1 minute. The supernatant was removed, and pellets were stored at -80°C. The pellets were thawed in Yeast Lysis Buffer (50mM Na-HEPES pH 7.5, 100mM NaOAc, 1mM EDTA, 1mM EGTA, 5 mM MgOAc, 5% glycerol, 0.25% NP-40, 3 mM DTT) with 1X Halt Protease Inhibitor Cocktail (Thermo Fisher Scientific). Approximately 100µl of buffer was used for each 100mg of pellet. Cells were lysed using the methods detailed above (Section 4.2.2).

25µl of Dynabeads Protein G (Fisher Scientific) were prepared for each of the samples. Beads were washed twice with 500µl Conjugation Buffer (0.5% BSA, 0.1% Triton X-100 in PBS), then re-suspended in 500µl buffer and 2µg anti-AARS (ab226259). Beads and antibody were incubated overnight rocking at 4°C. Yeast cell lysates were pre-cleared before immunoprecipitation: 25µl of magnetic beads were aliquoted and washed once with lysis buffer, before 2mg of yeast lysate in a total of 500µl lysis buffer was added. Samples were rocked at 4°C for 1 hour. The supernatant was then removed from antibody-conjugated beads and replaced with the pre-cleared lysates. These were rocked for 2.5 hours at 4°C. After incubation, they were washed once with 500µl lysis buffer, once with 200µl lysis buffer, and then re-suspended in 100µl lysis buffer before being transferred to a fresh 1.5mL Eppendorf tube. The supernatant was then removed, and beads were suspended in 25µl lysis buffer and 25µl 2x Tris Glycine Buffer (Invitrogen). Samples were boiled for 5 minutes with 2µl BME, and supernatants were removed to analyze in western blot assays.

4.2.6 Western blot analyses

To assess the levels of specific proteins in each experiment we performed western blot analyses. Protein concentrations for each sample were measured using the Thermo Scientific Pierce BCA Protein Assay kit. Samples were prepared with 1X Novex Tris-Glycine SDS sample buffer (Invitrogen) and 2-mercaptoethanol (BME), and boiled at 99°C for 5 minutes. Protein samples were separated on precast 4-20% Novex Wedgewell Tris-glycine gels (Invitrogen) at 150V for 1 hour and 15 minutes. PVDF membranes (Millipore Sigma) were pre-washed in 100% methanol for 1 minute, then soaked in 1X transfer buffer (Invitrogen) and 10% methanol between two pieces of filter paper (Thermo Fisher Scientific). The separated protein samples were transferred to the PVDF membranes using a Mini Trans-Blot Electrophoretic Transfer Cell (Biorad) at 100V for 1 hour. Membranes were then blocked for 1 hour with a 5% milk powder solution in 1X TBST. Primary antibodies were applied in 5% milk powder and membranes were incubated by rocking overnight at 4°C. The following day, membranes were washed three times with 1X TBST. Secondary antibodies against mouse (for the 6xHis primary antibody and PGK1 primary antibody), rabbit (for the AARS1 primary antibody and actin primary antibody), or rat (for the FLAG primary antibody) (Licor) were diluted in 5% milk powder solution at a concentration of 1:20,000, along with 0.1% Tween-20 and 0.02% SDS. This solution was applied to membranes for one hour, rocking at room temperature. Membranes were then washed three times with 1X TBST before exposure using a Licor Odyssey CLx Imaging System.

For yeast protein and HEK293T co-immunoprecipitation experiments, the AARS1 antibody (Bethyl Laboratories A303-473A) was used at 1:1,000 dilution. For fibroblast DSS assays, the same AARS1 antibody was used at 1:500 dilution. For HEK293T and yeast co-immunoprecipitation experiments, the 6xHis antibody (abcam 18184) was used at a dilution of 1:3,000. The FLAG antibody (BioLegend 637302) was used at a 1:2,500 dilution. The loading control was actin (Sigma A5060, 1:5,000) for mammalian protein blots and PGK1 (ab113687, 1:3,000) for yeast protein blots. For co-immunoprecipitation studies of AARS1 and ALA1, the AARS1 antibody used was ab226259 at 1:500. Full length images of western blots are available in Appendix B.

4.3 Results

4.3.1 Pathogenic AARS1 *alleles suppress yeast cell growth in the presence of wild-type* AARS1

Yeast complementation assays have been successfully used to assess the effect of ARS variants on gene function.[307] However, yeast have not been used to determine the effect of ARS variants in the presence of the wild-type allele, which is important for defining the mechanism of dominant ARS-related neuropathy. To investigate the dominant toxicity of *AARS1* variants, we developed an assay to assess the combined effects of human wild-type *AARS1* and human mutant *AARS1* on yeast viability, using the ptetO7-*ALA1* strain. In this strain, the yeast *AARS1* ortholog, *ALA1*, is placed under control of a doxycycline-repressible promoter.[435] This strain was transformed with: (**1**) a low-copy number, centromere-bearing vector (p413) containing a wild-type *AARS1* allele; and (**2**) a high-copy number vector (*i.e.*, bearing a 2 micron origin of replication) with a galactose-inducible promoter (pAG425) directing high levels of expression of a mutant *AARS1* allele (Figure 4.1B). Although this does not reflect the approximately equal allelic expression in human cells, we expect that high levels of mutant AARS1 are required to detect a dominant toxic effect in yeast cells, which are less likely to be sensitive to defects in AARS1 than a human peripheral nerve axon. To test mutant *AARS1* alleles for a dominant toxic effect, yeast cells were grown in the presence of galactose (to express mutant *AARS1*) and doxycycline (to repress endogenous *ALA1*). Subsequent yeast growth was then solely dependent on the two forms of human *AARS1*: one wild-type and one mutant.

To evaluate neuropathy-associated *AARS1* variants for a dominant-negative effect, we focused on two well-characterized pathogenic *AARS1* variants, R329H and G102R. R329H is a recurrent mutation in the tRNA binding domain that has been identified in 9 families with CMT disease (Table 2.1). It affects a highly conserved residue, and significantly impairs AARS1 enzymatic function when assessed in an *in vitro* aminoacylation assay that evaluates enzyme activity under Michaelis-Menten conditions.[107] The G102R variant affects a highly conserved residue in the activation domain of AARS1 and was found in a family with dominant myeloneuropathy. Both G102R[109] and R329H[107] have been modeled at conserved codons in the yeast ortholog *ALA1*, and were unable to support yeast growth, indicating a loss-of-function effect *in vivo.* To distinguish

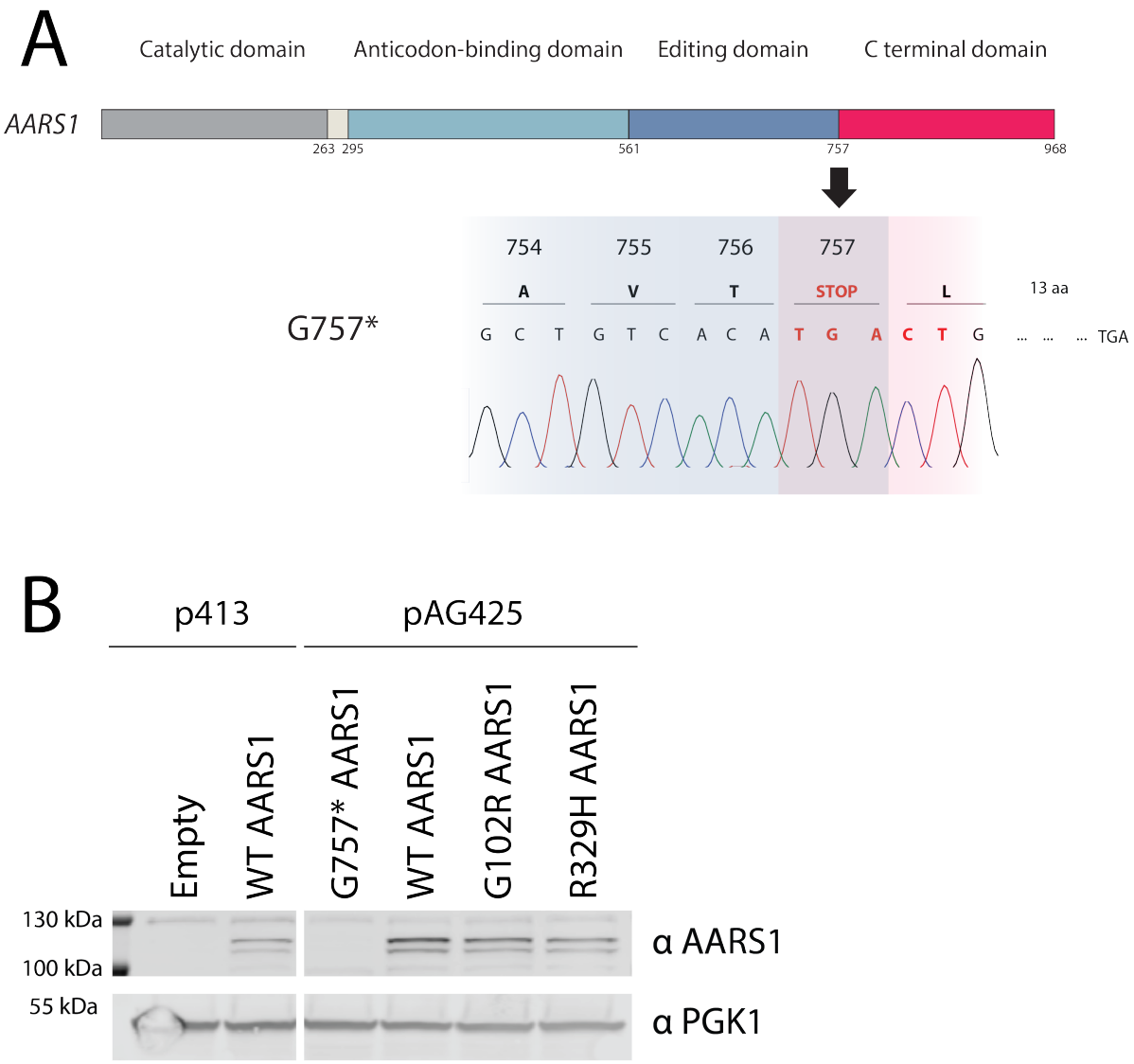


Figure 4.1. Yeast expression of human wild-type or mutant AARS1.

(A) Cartoon of the AARS1 protein structure, with the junction between the editing domain (blue) and the C-terminal domain (pink) magnified. A chromatogram illustrates the five base pair insertion introduced at this junction to create a premature stop codon, G757*. This insertion also shifts the open reading frame to generate another stop codon 13 amino acids downstream. **(B)** Western blot of yeast protein lysates, from yeast expressing: an empty p413 vector or wild-type AARS1 from p413 (left); wild-type or mutant AARS1 from pAG425 (right). The top blot was probed with an antibody against AARS1, and the bottom with an antibody against the yeast housekeeping gene PGK1. Full-length AARS1 is predicted to migrate at 107kDa. The lower band likely corresponds to a downstream open-reading frame beginning at M44, which is predicted to produce a 102kDa AARS1 protein. A representative image from four biological replicates is shown

any dominant-negative properties of these variants from a purely loss-of-function effect, we also generated a premature stop codon, G757* (Figure 4.1A), which does not generate detectable levels of AARS1 protein (Figure 4.1B). Therefore, it is expected to be a loss-of-function allele that does not exert dominant-negative effects. This is a more precise negative control than an empty vector because it includes the AARS1 coding sequence, meaning yeast transformed with the G757* allele must replicate and express an almost identical vector as yeast expressing the G102R or R329H alleles.

To confirm that R329H and G102R are loss-of-function alleles when tested in the human *AARS1* coding sequence, ptetO7-*ALA1* yeast were first transformed with an empty p413 vector, then with either wild-type or mutant human *AARS1* expressed from the *GAL1* inducible promoter on pAG425. When plated on galactose and doxycycline, yeast expressing G757* did not form colonies (Figure 4.2A), indicating that yeast cannot grow without *ALA1* expression or functional *AARS1*. Transformation with wild-type *AARS1* lead to robust yeast growth, confirming that wild-type AARS1 can complement loss of *ALA1*. Neither G102R *AARS1* nor R329H *AARS1* supported yeast growth (Figure 4.2A), confirming previous reports that these are loss-of-function alleles.[107,109]

To test for dominant toxicity, ptetO7-*ALA1* yeast cells were transformed with wild-type *AARS1* on the low-copy p413 vector, then with wild-type or mutant *AARS1* on the high-copy, glactose-inducible pAG425 vector. Transformed strains were spotted on galactose (to induce expression from the pAG425 vector) and doxycycline (to repress *ALA1* expression). The combination of wild-type *AARS1* and G757* *AARS1* supported growth, as did the combination of the two wild-type *AARS1* plasmids (Figure 4.2B). Importantly, over-expressing the G757* variant—a loss-of-function *AARS1* allele that does not lead to detectable protein (Figure 4.1B)—does not interfere with yeast growth. Similarly, over-expressing wild-type human *AARS1* does not interfere with yeast growth. In contrast, the combination of wild-type *AARS1* with G102R *AARS1*, and the combination of wild-type *AARS1* with R329H *AARS1*, both caused significantly reduced yeast growth (Figure 4.2B). This demonstrates that these alleles are not only loss-of-function but also dominantly toxic to yeast cell growth, even in the presence of wild-type *AARS1*. This toxicity is at least partially rescued by restoring *ALA1* expression (Figure 4.3A, 4.3B), which suggests that

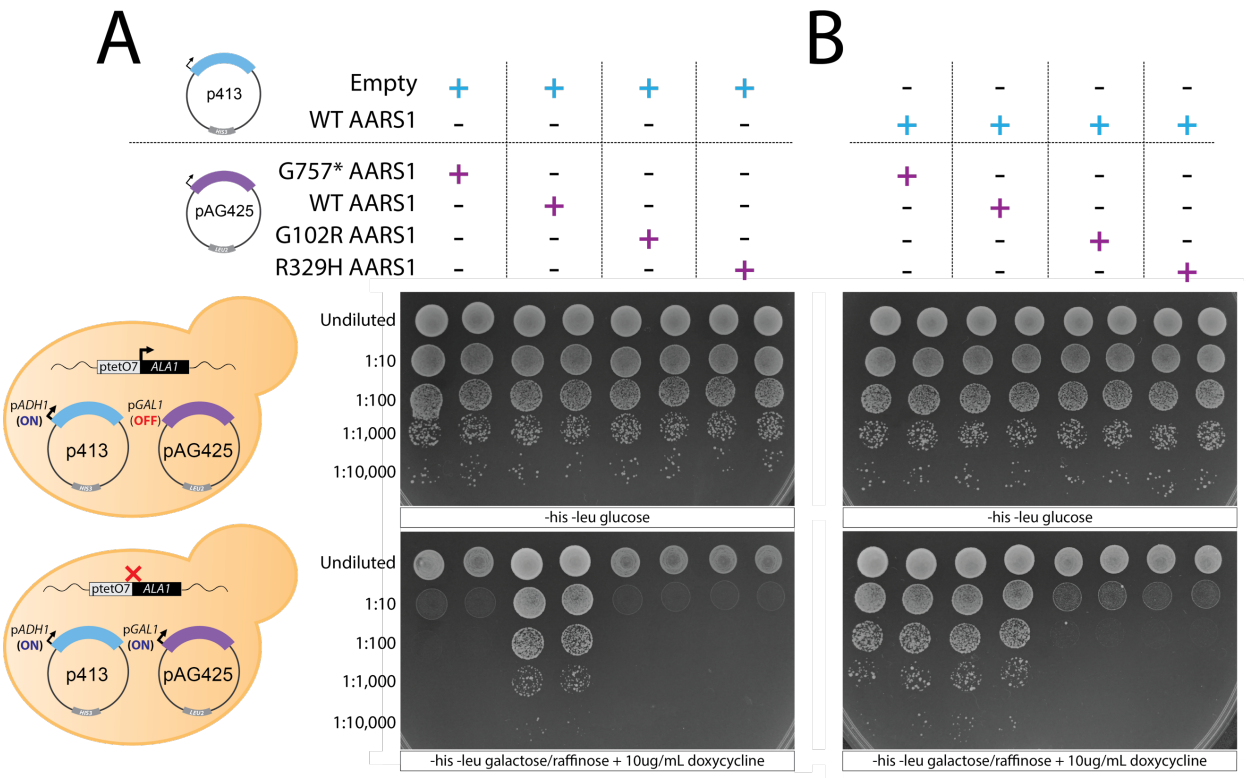


Figure 4.2. G102R and R329H *AARS1* are loss-of-function, dominantly toxic alleles in yeast.

(A) Images of yeast strains expressing an empty p413 vector and (plated left to right) G757*, wild-type, G102R, or R329H *AARS1* from a galactose-inducible, high copy number vector (pAG425). The top panel shows yeast spotted in 1:10 serial dilutions on glucose media lacking histidine and leucine. The bottom panel shows yeast spotted in 1:10 serial dilutions on galactose and raffinose media lacking histidine and leucine, with 10µg/mL doxycycline. **(B)** Images of yeast strains expressing wild-type AARS1 from p413, a low copy number vector with an *ADH1* promoter. These strains also express (plated left to right) G757*, wild-type, G102R, or R329H *AARS1* from pAG425. The top panel shows 1:10 serial dilutions on glucose media lacking histidine and leucine. The bottom panel shows 1:10 serial dilutions on galactose and raffinose media lacking histidine and leucine, with 10µg/mL doxycycline. For both **(A)** and **(B)**, 13 (for G102R) or 16 (for R329H) biological replicates were performed.

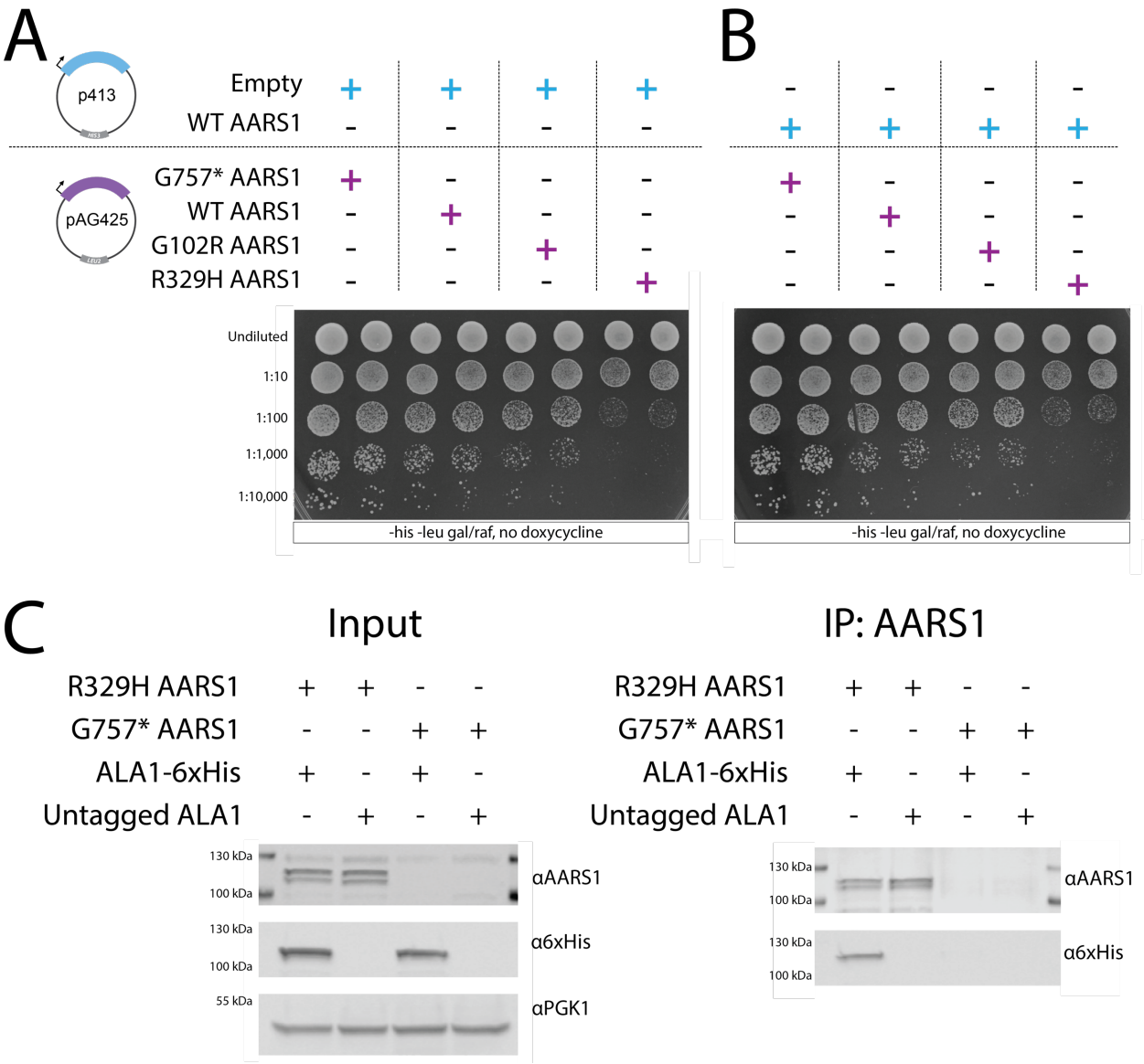


Figure 4.3. ALA1 interacts with R329H AARS1.

(A) Images of yeast strains expressing an empty p413 vector and G757*, wild-type, G102R, or R329H *AARS1* from pAG425. **(B)** Representative images of yeast strains expressing a wild-type AARS1 from p413 and G757*, wild-type, G102R, or R329H *AARS1* from pAG425. For **(A)** and **(B)**, strains are spotted in 1:10 serial dilutions on galactose and raffinose media lacking histidine and leucine. These plates do not contain doxycycline to repress *ALA1* expression. 13 (for G102R) or 16 (for R329H) biological replicates were performed. **(C)** Western blot of yeast lysate from strains expressing R329H or G757* AARS1, with 6xHis-tagged ALA1 or untagged ALA1 (left). To the right, a western blot of immunoprecipitated AARS1 and co-immunoprecipitated ALA1-6xHis. Primary antibodies include anti-AARS1, anti-6xHis, and anti-PGK1 as the Input loading control. A representative image from two technical replicates is shown.

the dominant toxicity associated with R329H and G102R *AARS1* is due to reduced alanine-tRNA charging. Interestingly, *ALA1* does not fully rescue the toxicity of R329H. This may be related to the observation that human R329H AARS1 interacts with wild-type yeast; immunoprecipitation of R329H AARS1 co-immunoprecipitates 6xHis-tagged ALA1 (Figure 4.3C). This interaction is consistent with R329H AARS1 acting as a dominant-negative allele against ALA1.

4.3.2 Pathogenic AARS1 *variants do not significantly reduce dimerization*

The data presented in Figure 4.2 demonstrate that the pathogenic, loss-of-function G102R and R329H *AARS1* variants are dominantly toxic to yeast cell growth; however, this experiment does not distinguish between a dominant-negative effect and some gain-of-function toxicity unrelated to AARS1 function. To directly test for a dominant-negative effect, we first investigated if mutant AARS1 dimerizes with wild-type AARS1. Ultracentrifugation analyses have demonstrated that isolated mutant ARS proteins retain homo-dimerization[176,300,350]; however, no studies have addressed hetero-dimerization between the AARS1 mutant and wild-type subunits. To address this, HEK293T cells were transfected with a vector expressing wild-type human AARS1 with an in-frame 3xFLAG tag, and a vector expressing wild-type or mutant AARS1 (G102R or R329H) with an in-frame 6xHis tag (Figure 4.4A). After growth for 48 hours, cells were lysed and AARS1-6xHis was immunoprecipitated. Co-immunoprecipitated proteins were subjected to a western blot with an anti-FLAG antibody to detect AARS1-3xFLAG. The reciprocal immunoprecipitation was also performed, by immunoprecipitating AARS1-3xFLAG and immunoblotting for AARS1-6xHis. Both approaches detected comparable interactions between wild-type and wild-type, wild-type and G102R, and wild-type and R329H (Figure 4.4B and 4.4C). To assess dimerization of endogenous AARS1 in patient cells, fibroblasts from a patient heterozygous for R329H *AARS1* were crosslinked with disuccinimidyl suberate (DSS), along with two independent control fibroblast cell lines. In untreated lysates (-DSS), AARS1 is detected between 100 and 130 kDa. In DSS-treated lysates, there is an additional band that migrates between the 130kDa and 250kDa markers, consistent with a dimeric AARS1 protein (Figure 4.5A). The percentage of AARS1 in dimeric form was not significantly different between control and patient cell lines (Figure 4.5B). Combined, these data indicate that mutant

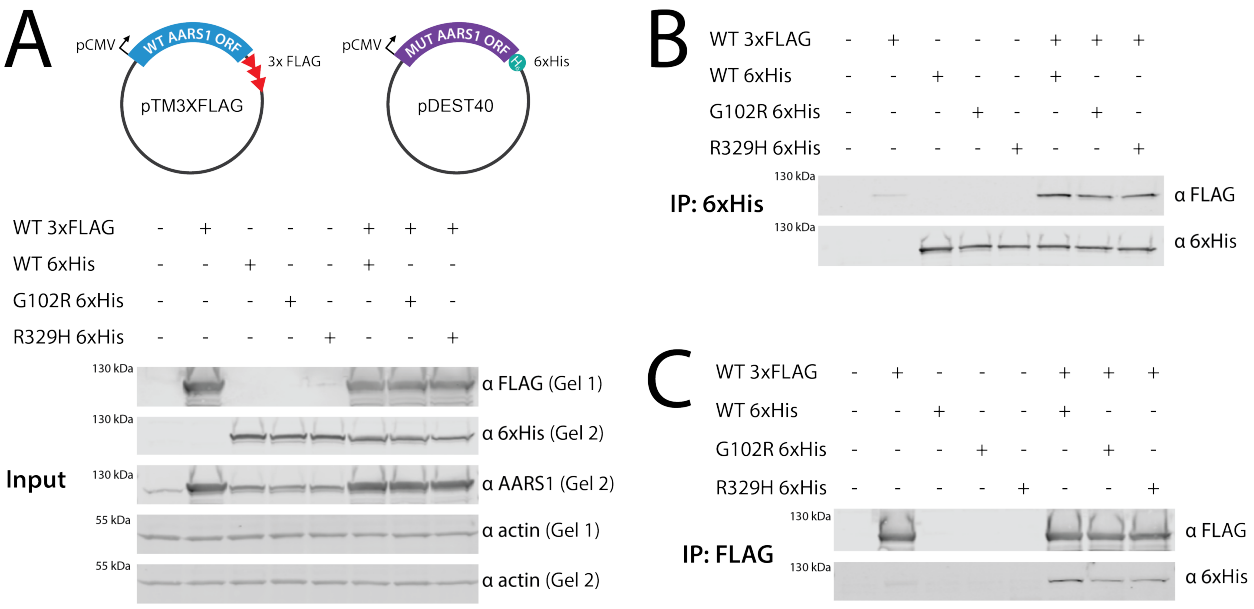


Figure 4.4. R329H and G102R do not impair dimerization with wild-type AARS1.

(A) (Top) Cartoon of plasmid expressing wild-type AARS1 tagged with 3xFLAG and plasmid expressing wild-type or mutant AARS1 tagged with 6xHis. (Bottom) Western blot of HEK293T cells expressing wild-type AARS1-3xFLAG and/or wild-type, G102R, or R329H AARS1-6xHis. **(B)** Western blot demonstrating immunoprecipitation of 6xHis-tagged AARS1 protein and co-immunoprecipitation of 3xFLAG-tagged wild-type AARS1. Five biological replicates were performed for R329H; three biological replicates were performed for G102R. **(C)** Western blot demonstrating immunoprecipitation of 3xFLAG-tagged wild-type AARS1 and co-immunoprecipitation of 6xHis-tagged wild-type, G102R, or R329H AARS1. Two biological replicates were performed for both G102R and R329H. Primary antibodies include anti-FLAG, anti-6xHis, anti-AARS1 and anti-actin as the Input loading control.

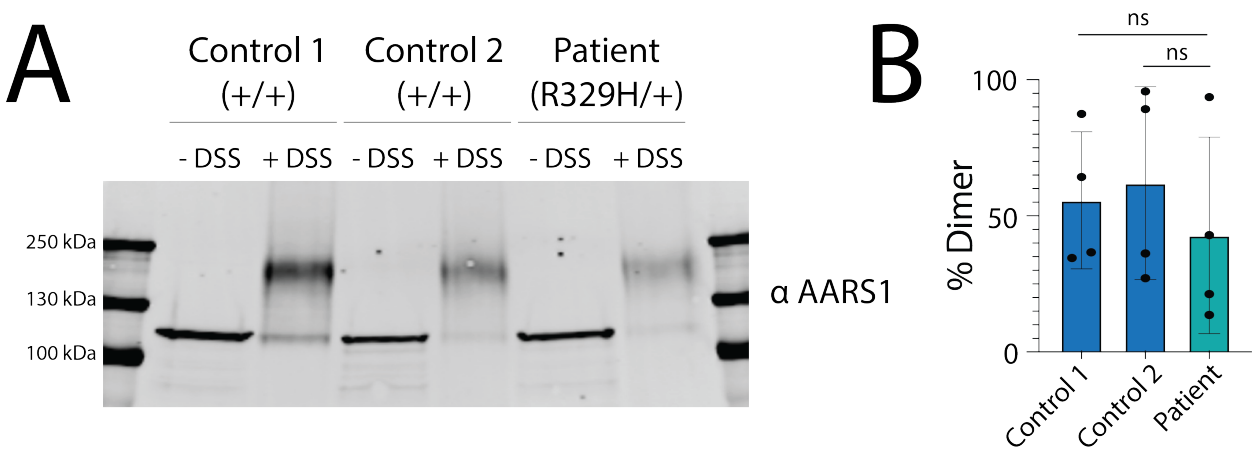


Figure 4.5. Chemical crosslinking of AARS1 in patient fibroblast cells.

(A) Western blot of fibroblast lysates from a R329H/+ *AARS1* individual and two +/+ controls. Lysates were crosslinked with DSS treatment (+DSS) or left untreated (-DSS), then separated on a gel and probed with an anti-AARS1 antibody. An image of the full blot is shown. **(B)** The percentage of AARS1 signal corresponding to the dimeric form of AARS1, quantified with ImageJ. The mean and standard deviation of four technical replicates is shown. Unpaired t-tests with Welch's correction were performed to determine if there was a statistically significant difference between R329H/+ cells and either of the two wild-type controls. ns = not significant.

AARS1 proteins retain the ability to dimerize with wild-type AARS1, which is required for a dominant-negative effect.

4.3.3 Designing dimer-disrupting AARS1 *variants*

A dominant-negative mechanism for ARS variants requires an interaction between the wild-type and mutant subunits of the homodimer. Therefore, if neuropathy-associated *AARS1* alleles function via a dominant-negative effect, placing a dimer-disrupting variant *in cis* with a dominant pathogenic variant should impair its ability to bind to the wild-type subunit. This should alleviate the dominant-negative effect and, in yeast, rescue the impaired growth. To identify dimer-disrupting variants, a series of deletions were designed in the C-terminal dimerization domain based on the published crystal structure.[30] These engineered deletions targeted amino acids that have multiple contacts with the opposite subunit (Figure 4.6A, left). This series comprised a seven amino acid deletion to encompass several contact points (ΔKNVGCLQ) as well as smaller deletions within or near this region for a more targeted approach (ΔNVG and ΔQE). This series also included a deletion of C947, a cysteine residue that forms a putative disulfide bond with C773 on the opposite subunit[30]. Finally, a stop codon at Q855 was designed to ablate the entire terminal globular domain (Figure 4.6A, right). The panel of putative dimer-disrupting *AARS1* alleles were cloned into the pAG425 vector and transformed into the ptetO7-*ALA1* yeast strain, then tested for their ability to support yeast growth in a complementation assay, with the expectation that dimerization is required for AARS1 function.

None of the deletions fully complemented loss of *ALA1*, indicating that they all reduce *AARS1* function, consistent with impaired dimerization and/or decreased protein expression (Figure 4.6B). To distinguish between these two possibilities, each allele was also evaluated for an effect on AARS1 expression. The ΔNVG allele lead to no detectable AARS1 protein, providing an explanation for its failure to complement in yeast. The C947 deletion significantly reduced AARS1 expression (Figure 4.6C) but still showed partial complementation in yeast, indicating that the C947 residue may be more important for stability than for *AARS1* function. The deletions ΔKNVGCLQ and ΔQE also reduced AARS1 expression (Figure 4.6C) and lead to less yeast growth than ΔC947 (Figure 4.6B), suggesting that these deletions reduce dimerization as well as

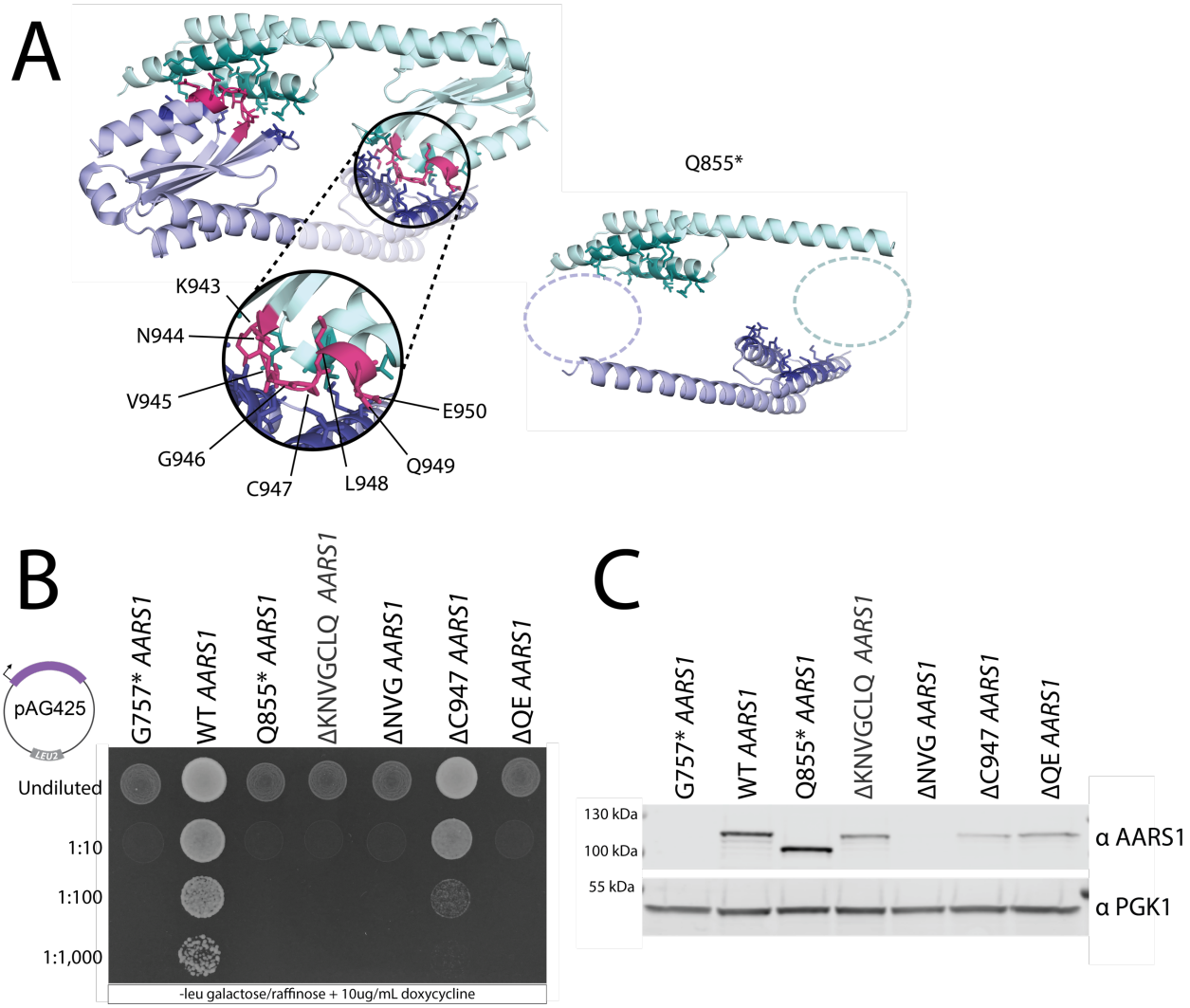


Figure 4.6. Designing and testing a deletion series in AARS1 dimerization domain.

(A) (Left) Crystal structure of the AARS1 C-terminal dimerization domain. One subunit is shown in green, the other in purple. Amino acid residues that contact the opposite subunit are shown in dark green or dark purple. The residues targeted in this assay are shown in pink and labeled. (Right) Illustration of the dimerization domain with the Q855* mutation. The dashed circles indicate the globular domain that is ablated by the premature stop codon. **(B)** Yeast strains expressing the negative control G757* *AARS1*, the positive control WT *AARS1*, or each of the five deletions designed in the dimerization domain. Yeast are spotted in 1:10 serial dilutions on galactose and raffinose media lacking leucine, with 10μg/mL doxycycline. A representative image of four biological replicates is shown. **(C)** Western blot of yeast lysates from each of the dimerization deletion strains. Yeast were grown in galactose and raffinose media lacking leucine, with no doxycycline. Blots were treated with anti-AARS1 antibody or anti-PGK1 antibody. A representative image of three biological replicates is shown.

protein abundance. Only the globular domain deletion Q855* produced a loss-of-function protein with no detectable decrease in protein levels. This made it an ideal candidate to test for reduced dimerization. Overall, these data demonstrate that the dimerization domain of AARS1 is important for both the stability and the function of the protein, indicating that AARS1 must form a homodimer to successfully charge tRNA.

To test if Q855* reduces binding to wild-type AARS1, HEK293T cells were transfected with wild-type AARS1-3xFLAG and either wild-type or Q855* AARS1-6xHis (Figure 4.7A). Immunoprecipitation of wild-type AARS1-6xHis co-precipitated wild-type AARS1-3xFLAG, indicating an interaction between the two tagged wild-type subunits (Figure 4.7B). However, immunoprecipitation for Q855* AARS1-6xhis did not co-precipitate wild-type AARS1-3xFLAG (Figure 4.7B), indicating that the Q855* truncation reduces binding to the wild-type AARS1 protein. These findings were supported by performing the reciprocal experiment. Here, immunoprecipitation of wild-type AARS1-3xFLAG co-immunoprecipitated wild-type AARS1-6xhis (Figure 4.7C). However, wild-type AARS1-3xFLAG did not co-immunoprecipitate Q855* AARS1-6xhis (Figure 4.7C). These data demonstrate that the engineered Q855* AARS1 variant reduces interactions with wild-type AARS1 and confirms that the C-terminal globular domain is required for dimerization.

4.3.4 Reducing the dimerization capacity of pathogenic AARS1 *alleles rescues yeast growth*

If G102R or R329H *AARS1* reduce yeast growth through a dominant-negative mechanism, then impairing their ability to dimerize with wild-type AARS1 should rescue yeast growth. To test this, the Q855* mutation was introduced *in cis* with either G102R or R329H using site-directed mutagenesis. These double mutants were then cloned into pAG425 and transformed into the ptetO7-*ALA1* strain. Both G102R+Q855* *AARS1* and R329H+Q855* *AARS1* produced a stable, truncated AARS1 protein (Figure 4.8A). Complementation assays studying G102R+Q855* *AARS1* and R329H+Q855* *AARS1* in the presence of an empty p413 vector showed no yeast growth, consistent with the double-mutants acting as loss-of-function alleles (Figure 4.8B). These alleles were then tested in the presence of wild-type AARS1 expressed from the p413 vector. As before, neither the control allele G757* *AARS1* nor wild-type *AARS1* repressed

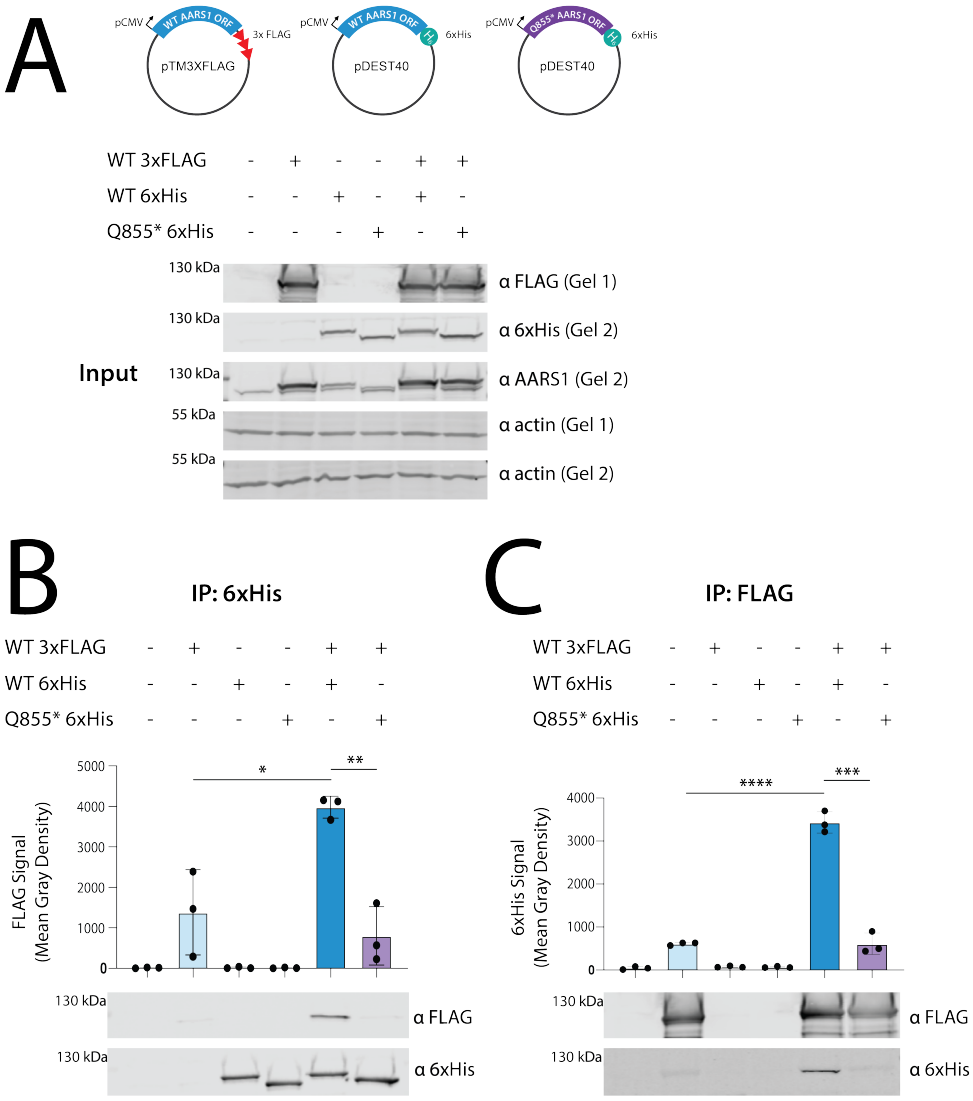


Figure 4.7. Q855* AARS1 impairs dimerization with wild-type AARS1.

(A) (Top) Cartoon of plasmid expressing wild-type AARS1 tagged with 3xFLAG and plasmid expressing wild-type or Q855* AARS1 tagged with 6xHis. (Bottom) Western blot of HEK293T cells expressing wild-type AARS1-3xFLAG and/or wild-type or Q855* AARS1-6xHis. **(B)** Western blot showing immunoprecipitation of 6xHis-tagged AARS1 protein and co-immunoprecipitation of 3xFLAG-tagged wild-type AARS1. ImageJ quantification of anti-FLAG band intensity is shown above. Bars indicate the mean value and one standard deviation for 3 biological replicates. **(C)** Western blot showing immunoprecipitation of 3xFLAG-tagged wild-type AARS1 protein and co-immunoprecipitation of 6xHis-tagged AARS1. ImageJ quantification of anti-6xHis band intensity is shown above. Bars indicate the mean value and one standard deviation for 3 biological replicates. For **(A)-(C)**, primary antibodies include anti-FLAG, anti-6xHis, anti-AARS1 and anti-actin as the Input loading control. Unpaired t-tests were performed to determine if the difference in band intensity between samples was statistically significant. **** $p<0.0001$, *** $p<0.001$, ** $p<0.01$, * $p<0.05$.

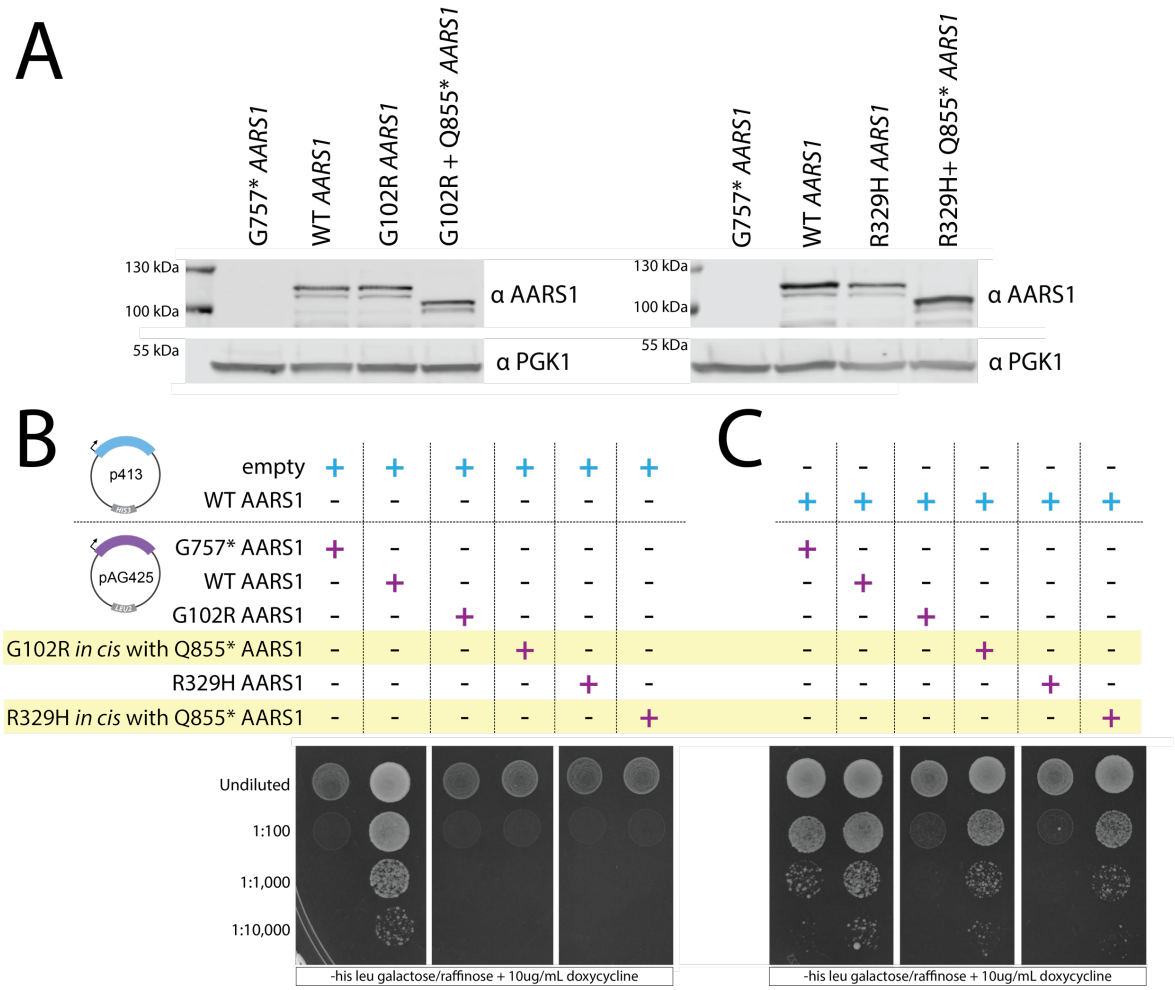


Figure 4.8. Reducing dimerization of G102R and R329H with wild-type AARS1 rescues yeast growth.

(A) Western blots of yeast lysates from yeast expressing the negative control G757* *AARS1*, the positive control WT *AARS1*, G102R *AARS1* alone or *in cis* with Q855*, or R329H *AARS1* alone or *in cis* with Q855*. Yeast were grown in galactose and raffinose media lacking leucine, with no doxycycline. Representative images were selected from three biological replicates. Blots were treated with anti-AARS1 antibody and anti-PGK1 antibody as a loading control. **(B)** Yeast strains expressing an empty 413 vector and wild-type or mutant AARS1 from pAG425. **(C)** Yeast strains expressing wild-type *AARS1* from p413, and wild-type or mutant *AARS1* from pAG425. For **(B)** and **(C)**, strains were spotted in a 1:10 dilution on galactose and raffinose media lacking histidine and leucine, with 10µg/mL doxycycline. Representative images from 8 biological replicates are shown.

yeast growth, and both G102R *AARS1* and R329H *AARS1* repressed yeast growth. However, placing Q855* *in cis* with either G102R or R329H *AARS1* ameliorated the phenotype and rescued growth. This rescued growth was comparable to the growth of yeast expressing the non-toxic control alleles G757* or wild-type *AARS1* (Figure 4.8C). These *in vivo* data demonstrate that disrupting the dimerization of G102R or R329H AARS1 with wild-type AARS1 is sufficient to rescue the dominant toxic phenotype, and shows that this phenotype is a result of mutant AARS1 dimerizing with wild-type AARS1. In sum, our yeast and biochemical data provide evidence that neuropathy-associated *AARS1* alleles are loss-of-function variants that dominantly repress yeast growth through dimerization with the wild-type subunit; *i.e.* that they act via a dominant-negative mechanism.

4.3.5 The AARS1 anticodon-binding domain is susceptible to dominant-negative mutations

R329H *AARS1* is a high-confidence pathogenic allele, as it has been identified in at least 46 individuals across 9 families with dominant peripheral neuropathy (Table 2.1). This includes a newly identified family comprising three individuals, all with dominant axonal Charcot-Marie-Tooth disease (Table 4.1). Therefore, characterizing R329H as a dominant-negative allele suggests that a dominant-negative effect is a relevant component of pathogenicity. If other *AARS1* alleles with less robust genetic evidence—such as G102R—also have a dominant-negative effect, this increases the likelihood that they too are pathogenic. To apply this dominant-negative assay to other *AARS1* alleles, we first focused on a set of variants at the R329 residue and the surrounding region. This includes R326W, a loss-of-function variant identified in a multi-generational family with CMT,[110] and R329S, a recently identified variant in a patient with axonal neuropathy (Table 4.1). Interestingly, previous work by McLaughlin et al. identified a high degree of cytosine methylation in this area, making these nucleotides susceptible to cytosine deamination. This study predicted numerous missense variants that could arise from such C→T changes, including R326W and R329S, which have now both identified in individuals with dominant peripheral neuropathy.[107] McLaughlin et al. also predicted that cytosine deamination could lead to the mutation R329C. Therefore, in addition to the now-confirmed patient alleles, we included R329C in our functional studies with the hypothesis that it will be ultimately be identified in CMT patients. The R326W, R329S, and R329C variants were

Table 4.1. Clinical information of four additional patients with *AARS1*-mediated dominant peripheral neuropathy.

Patient ID	Gender	Age	Mutation	Age of Onset	CMTNS	Vibration LL	Vibration UL	Cutaneous LL	Cutaneous UL	Ulnar DML (ms) <3.4	Ulnar NCV1 (m/s) >49	Ulnar NCV2 (m/s) >50	Ulnar CMAP (mV) >2.8	Median DML (ms) <3.5	Median NCV (m/s) >48	Median CMAP (mV) >3.5
75872-0001	Male	56	R329S	11	29 (Severe)	Absent toes, ankles, knees	Absent fingers, reduced wrists and elbows	Absent toes, ankles, knees	Absent fingers, wrists	4.4	24	23	0.323	NR	NR	NR
75292-0001	Male	76	R329H	56	11 (Low moderate)	Reduced toes, ankles, knees	Normal	Normal	Normal	3.2	39	42	9.3	5.4	39	5.4
75292-9001	Female	44	R329H	20	7 (Mild)	Reduced toes, ankles, knees	Normal	Normal	Normal	2.9	49	50	8	4.5	43	4.9
75292-0100	Male	72	R329H	25	11 (Low moderate)	Absent toes, reduced ankles	Normal	Normal	Normal	3.2	45	55	10.1	4.8	41	5.6

CMTNS=CMT Neuropathy Score, LL=lower limb, UL=upper limb, DML=distal motor latency (upper limit of normal is 3.4 milliseconds for the ulnar nerve and 3.5 milliseconds for the median nerve), NCV=nerve conduction velocity (lower limit of normal is 49 or 50 meters per second for the ulnar nerve and 48 meters per second for the median nerve), CMAP = compound muscle action potential (lower limit of normal is 2.8 millivolts for the ulnar nerve and 3.5 for the median nerve). NR=no recording.

introduced into the *AARS1* open reading frame with site-directional mutagenesis, then cloned into the galactose-inducible pAG425 expression vector and transformed into the ptetO7-*ALA1* yeast strain. G757* *AARS1*, R329H *AARS1*, and wild-type *AARS1* were included as controls. To demonstrate that only pathogenic *AARS1* alleles have a dominant toxic effect, G931S *AARS1*, a benign polymorphism found in the general population (with a gnomAD allele count of 2,147/282,842, including 20 homozygous individuals[334]), was included as a negative control. As previously reported,[110] R326W did not support yeast growth in the absence of *ALA1* (Figure 4.9A, top panel). Consistent with the functional importance of this region, R329S and R329C also did not support yeast growth (Figure 4.9A, top panel). G931S *AARS1* supported yeast growth comparable to yeast expressing wild-type *AARS1*. These variants were then tested in the presence of the wild-type *AARS1* allele expressed from p413. Similar to R329H and G103R, R326W, R329C, and R329S *AARS1* repressed yeast growth compared to G757*, wild-type, or G931S *AARS1* (Figure 4.9B, top panel). This shows that these three alleles also exert a dominant toxic effect. Notably, as was the case for G102R and R329H, this growth defect was improved when strains were plated on media with no doxycycline, which restores endogenous *ALA1* expression (Figure 4.9 A and B, bottom panels). This supporting the argument that the repressed growth phenotype is due to an alanine-tRNA charging defect.

To determine if the toxicity of these alleles depends on their ability to dimerize with wild-type AARS1, the dimer-disrupting Q855* variant was introduced *in cis* with R326W, R329S, and R329C. These double-mutant alleles were transformed into yeast expressing wild-type *AARS1* from p413, and plated on galactose and doxycycline to express the double-mutant *AARS1* allele and repress endogenous *ALA1*. Q855* rescued the yeast growth phenotype for all three variants (Figure 4.10B, top panel), indicating that R326W, R329S, and R329C exert their dominantly toxic effect through dimerization—*i.e.*, they are also dominant-negative alleles. In sum with G102R and R329H, these experiments present convincing evidence that dominant pathogenic *AARS1* alleles exert a dominant-negative effect.

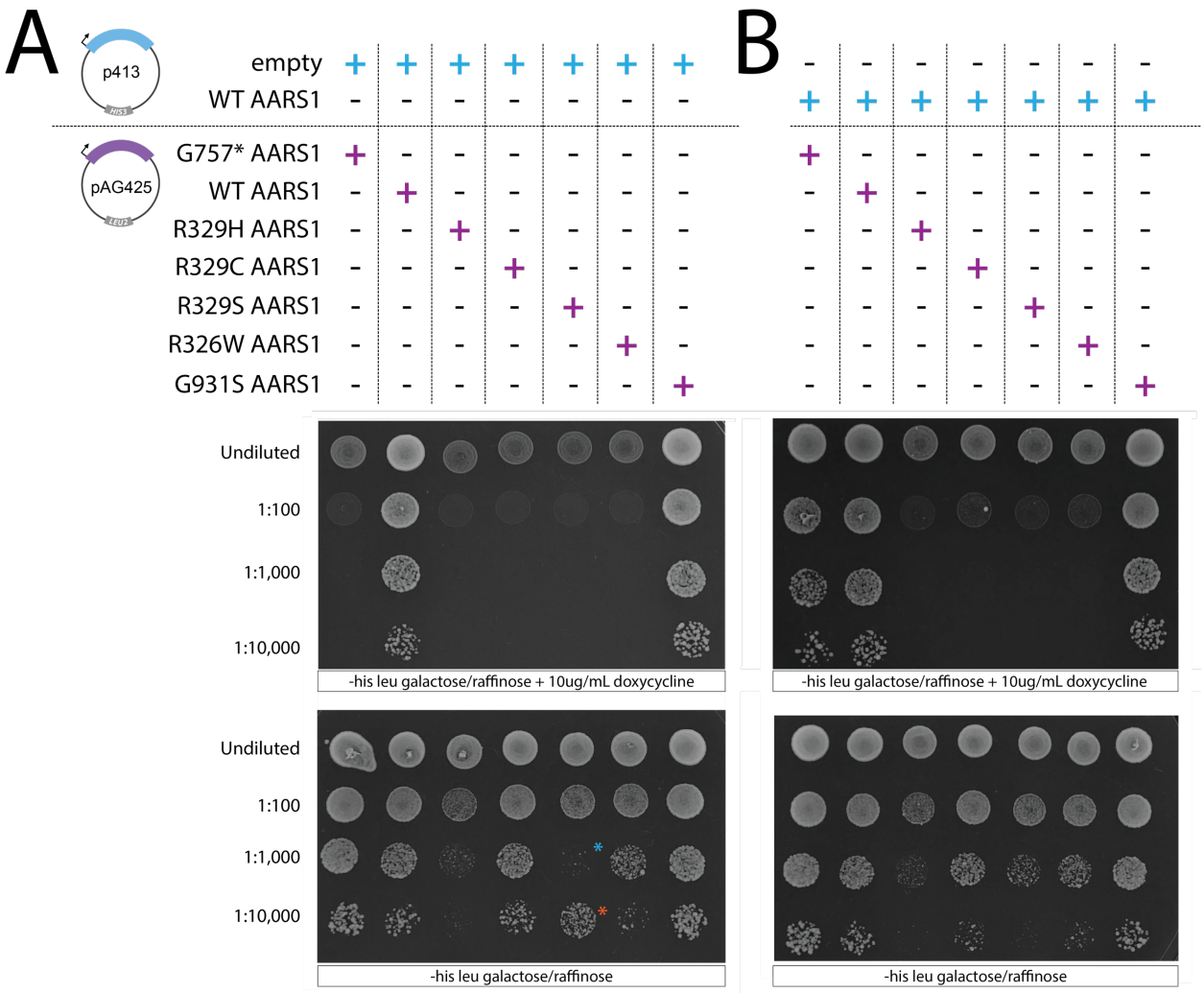


Figure 4.9. R326W, R329C, and R329S *AARS1* are dominantly toxic to yeast.

(A) Yeast strains expressing an empty 413 vector and wild-type or mutant AARS1 from pAG425. The blue asterisks indicates the spot is diluted at 1:10,000; the orange asterisk indicates the spot is diluted at 1:1,000 **(B)** Yeast strains expressing wild-type *AARS1* from p413, and wild-type or mutant *AARS1* from pAG425. For top panels of **(A)** and **(B)**, strains were spotted in a 1:10 dilution on galactose and raffinose media lacking histidine and leucine, with 10µg/mL doxycycline. For bottom panels of **(A)** and **(B)**, strains were spotted in a 1:10 dilution on galactose and raffinose media lacking histidine and leucine, with no added doxycycline to repress endogenous *ALA1*.

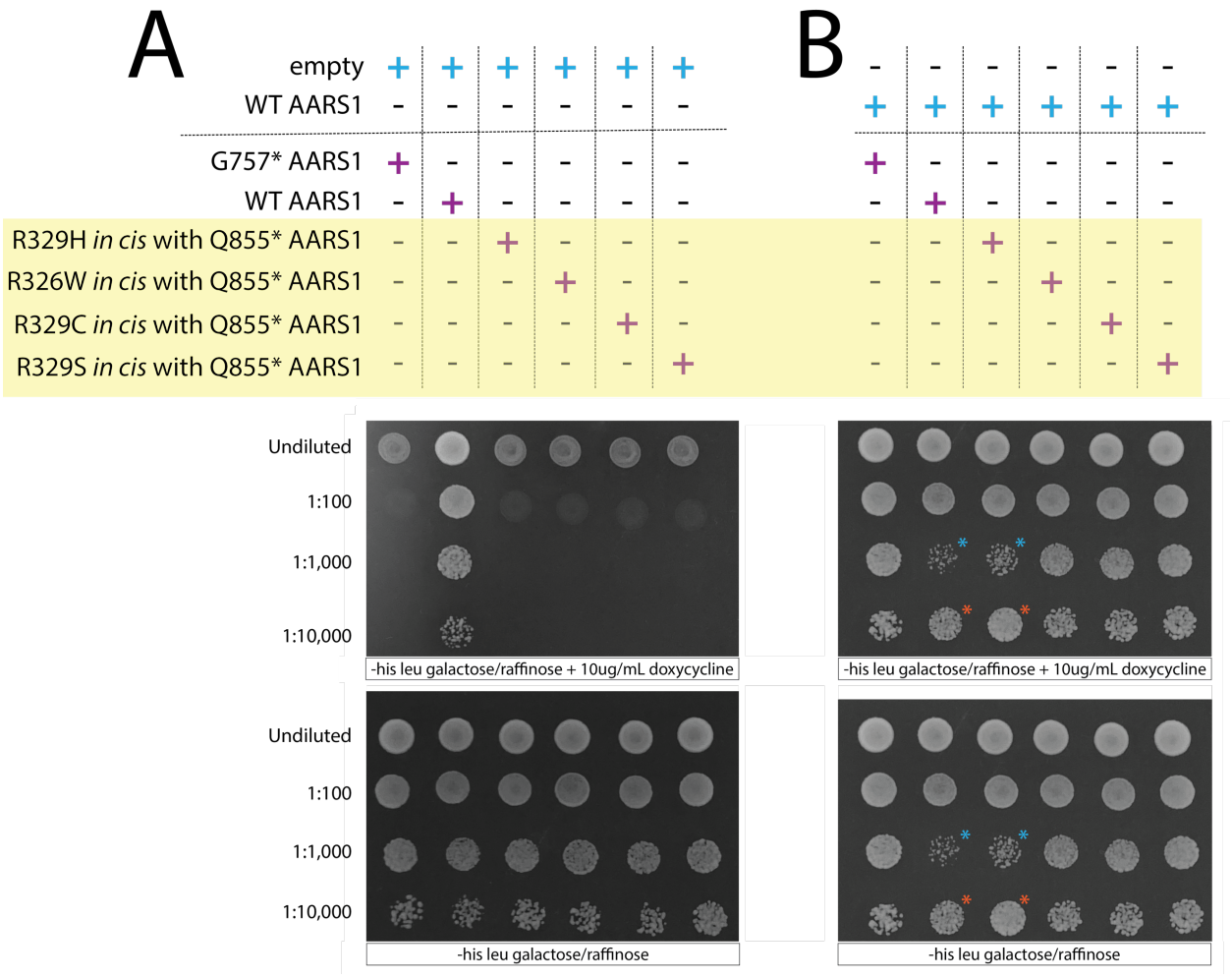


Figure 4.10. Reducing dimerization of R326W, R329C, or R329S AARS1 with wild-type AARS1 rescues yeast growth.

(A) Yeast strains expressing an empty 413 vector and wild-type or mutant AARS1 from pAG425. **(B)** Yeast strains expressing wild-type *AARS1* from p413, and wild-type or mutant *AARS1* from pAG425. Blue asterisks indicate a spot is diluted at 1:10,000; orange asterisks indicate a spot is diluted at 1:1,000. For top panels of **(A)** and **(B)**, strains were spotted in a 1:10 dilution on galactose and raffinose media lacking histidine and leucine, with 10µg/mL doxycycline. For bottom panels of **(A)** and **(B)**, strains were spotted in a 1:10 dilution on galactose and raffinose media lacking histidine and leucine, with no added doxycycline.

4.4 Discussion

Here, we present a humanized yeast assay to study the dominant toxicity of CMT-associated *AARS1* alleles. We demonstrate that multiple pathogenic, loss-of-function *AARS1* variants repress yeast growth when co-expressed with the wild-type *AARS1* allele, indicating that they are dominantly toxic. We also show that these variants retain the ability to dimerize with the wild-type AARS1 protein, and that disrupting this interaction by deleting critical dimerization residues from the mutant protein is sufficient to rescue the repressed yeast growth. This provides strong evidence that these pathogenic *AARS1* alleles can act via a dominant-negative mechanism to repress the activity of wild-type AARS1.

An important consideration when interpreting results from this model is that the allelic expression is intentionally skewed, with the pathogenic allele over-expressed relative to the wild-type one. This does not accurately reflect the approximately equal expression in human tissues; therefore, any dominant-negative effects in patients are likely to be much weaker than those demonstrated here, and less likely to have such significant consequences for cell viability. However, this would be more consistent with the patients' late-onset and tissue-restricted phenotype. A terminally differentiated peripheral neuron that must maintain local protein translation far from the soma may be particularly susceptible to even mild dominant-negative effects of an ARS mutation. To fully determine if a dominant-negative effect drives dominant *AARS1*-mediated peripheral neuropathy, a knock-in animal model (*e.g.*, mouse or worm) with an axonal pathology is required. Then, a dimer-disrupting variant such as Q855* can be introduced *in cis* with the pathogenic allele to determine if this ameliorates the neuronal phenotype.

It will also be important to adapt this dominant-negative yeast model to study other pathogenic variants in not only *AARS1,* but in *GARS1, HARS1, YARS1,* and *WARS1.* We hypothesize that these alleles will also produce dominant-negative effects, because these genes also encode homo-dimeric ARS enzymes (Table 1.1). Although there has been recent debate[30,348] as to whether AARS1 functions as a dimer or a monomer, the results of our study demonstrate that it functions as a dimer. The C-terminal dimerization domain (beginning at G757) that was previously reported to be dispensable for aminoacylation *in vitro*[30] is likely required for AARS1 stability *in*

vivo; truncating the protein at this residue leads to undetectable levels of AARS1 protein when expressed in yeast cells. Furthermore, the globular domain beginning at Q855*, while not required for protein stability in yeast or mammalian cells, is required for AARS1 dimerization and for *AARS1* complementation in yeast, suggesting that dimerization is required for AARS1 function *in vivo*. Finally, chemically crosslinked AARS1 from human fibroblast cells is detected at two molecular weights, one corresponding to a monomeric form and one corresponding to a dimeric form. In total, these data indicate that the dimerization domain is necessary for AARS1 function.

Although we hypothesize that dominant-negative effects will be a common theme for dominant pathogenic ARS variants, it is possible that this is only one component of the disease mechanism. For instance, there are rare exceptions to the loss-of-function pattern seen in patient alleles—the E337K *AARS1* variant increases rather than decreases AARS1 function, despite its proximity to dominant-negative anticodon-binding domain alleles.[110] It will be important to carefully reconsider the genetic evidence for the pathogenicity of E337K *AARS1*, as well as to assess it for a similar dominant-toxicity in yeast. If E337K is dominantly toxic in this system, comparing protein translation in yeast expressing E337K *AARS1* to protein translation in yeast expressing R329H *AARS1* may provide insights into the pathogenic mechanism of E337K, as discussed in Section 5.2.1.

It is important to point out that dominant-negative and gain-of-function effects are not mutually exclusive as causes for ARS-related neuropathy. Indeed, these two mechanisms may work in concert to exacerbate neuronal pathology. A recent study showed that the pathogenic *AARS1* variants N71Y, G102R, and R329H each cause a conformational change that enables binding to Neuropilin-1,[348] a widely expressed receptor that modulates a variety of signaling pathways, which is critical for neurovascular development.[362] However, although such an interaction might compound the damage in patient neurons, our yeast model proves that a neuronal-specific (or mammalian-specific) interaction is not required to make G102R or R329H toxic to cells.

To the best of our knowledge, there have been few yeast systems developed to test human pathogenic variants for a dominant-negative effect; notable examples include reporter assays to

test for dominant-negative p53 mutations[436,437] and enzymatic evaluation of yeast expressing dominant-negative UDP-galactose4-epimerase (GALE) alleles.[346,438] Here, we describe a tractable yeast model for rapidly evaluating patient variants in aminoacyl-tRNA synthetase genes for a dominant-negative effect. This system is likely applicable beyond aminoacyl-tRNA synthetases, and should be considered when evaluating dominant-negative patient variants in any essential, highly conserved human gene.

Chapter 5
Conclusions and Future Directions

5.1 Summary

Aminoacyl-tRNA synthetases (ARSs) are a family of enzymes that ligate tRNA to cognate amino acids, forming a critical substrate for protein translation. Research over the past two decades has implicated ARS-encoding genes in dominant and recessive human disease phenotypes. Due to their essential function, the complete loss of an ARS-encoding gene is incompatible with cellular life. However, bi-allelic ARS mutations that significantly decrease enzymatic function (but do not ablate it) cause a wide array of complex human phenotypes; partial loss of mitochondrial ARS function causes severe mitochondrial disorders, and partial loss of cytoplasmic ARS function causes multisystem disorders that can include central nervous system pathologies, liver failure, interstitial lung disease, and global developmental delay.[68] The full range of genotypes and phenotypes for these recessive disorders is not yet fully defined. It is also unclear how reduced ARS function impacts protein translation and cellular health, or why some tissues seem to be more sensitive to mutations in certain ARS genes than others (for example, why variants in *MARS1* predominantly cause lung phenotypes[210]).

In addition to recessive diseases, there is strong genetic evidence to implicate five ARSs—*AARS1, GARS1, HARS1, YARS1,* and *WARS1*—in dominant disease, which is restricted to axonal peripheral neuropathies. All five genes encode cytoplasmic, homodimeric enzymes (GARS1 also acts in the mitochondria).[13] Mutations in these genes cause motor and/or sensory neuron degeneration that leads to muscle wasting in the distal extremities, which usually begins in adolescence or adulthood. Moving forward, it will be important to determine if other ARS genes can also cause dominant neuropathy or if there is something unique about the five loci already

implicated. This will be critical for improving patient diagnosis and may also provide insight into common disease mechanisms. For example, if the remaining homodimeric ARS enzymes—but no monomeric ones—are found to cause dominant peripheral neuropathy, this would suggest dimerization is required for the molecular pathology and support a dominant-negative disease mechanism. In parallel to gene discovery, it will be important to build appropriate disease models to directly test mechanistic hypotheses *in vivo.* This dissertation work addressed several of these objectives; the results are summarized below.

5.1.1 Summary of Chapter 2

In Chapter 2, we evaluate novel ARS variants for a role in recessive or dominant disease. These variants are often discovered through diagnostic exome sequencing, and it can be difficult to obtain a complete family history and ascertain if the variant segregates with disease. In this context, determining the variant's impact on protein function contributes to determining pathogenicity. To assess patient variants in *GARS1, HARS1, MARS1, NARS1*, and *TARS1*, yeast complementation assays were performed using the human coding sequence of each gene. Here, the endogenous yeast ARS gene is effectively replaced with the human ARS. If the growth of yeast carrying the patient allele is reduced compared to the growth of yeast carrying the wild-type human ARS, this indicates that the patient variant impairs gene function.

Variants in *GARS1, HARS1, MARS1,* and *NARS1* were identified in individuals with dominant peripheral neuropathies. In yeast complementation assays, these variants reduced yeast viability, indicating that they impaired gene function. This is consistent with the pattern of impaired function seen across most dominant ARS alleles,[68,307] and supports the argument that these variants are pathogenic. This work expands the spectrum of pathogenic variants in *GARS1* and *HARS1*, and in the case of *NARS1*, contributes towards implicating an additional homodimeric ARS in dominant peripheral neuropathy. However, some of the loss-of-function variants—specifically, M236del *NARS1,* S356N *HARS1*, and A397T *MARS1*— have limited genetic evidence for pathogenicity, and require further investigation to confidently implicate them in disease.

In Chapter 2, we also investigate bi-allelic variants in *MARS1* and *TARS1*, which were identified in patients with multisystem recessive diseases. Only one of two *MARS1* alleles (detected *in trans*) significantly decreased yeast growth compared to wild-type. However, yeast complementation assays often cannot detect mild hypomorphic alleles. Moreover, because the patient phenotype was highly consistent with the established lung-predominant *MARS1* phenotype, it is likely that both *MARS1* variants are pathogenic in this individual. Similarly, only one of the four tested *TARS1* alleles reduced yeast growth. Since the individual is homozygous for this allele, it is likely that this deleterious variant is pathogenic. The other three *TARS1* alleles require additional experiments in a more sensitive assay to resolve their role in disease.

5.1.2 Summary of Chapter 3

In Chapter 3, we describe a model organism pipeline to predict the pathogenicity of novel variants in a homodimeric ARS that had not been associated with disease—TARS1. A panel of mutations in *TARS1* was designed to affect highly conserved residues, then tested for reduced function in a yeast complementation assay. Loss-of-function mutations were then prioritized for testing in *C. elegans*. Over-expression of G541R *tars-1* caused GABA-ergic neuron defects, similar to those caused by known pathogenic *HARS1* or *AARS1* alleles. However, these neuronal defects were not recapitulated when G541R was knocked into the endogenous *tars-1* locus.

To determine if G541R *TARS1* causes a peripheral neuropathy in a mammalian system, this variant was edited into the mouse *Tars1* gene. *Tars1*$^{G541R/+}$ mice were evaluated at 2 months, 5-6 months, and 1 year of age. Although there were behavioral indications of neuromuscular defects, these were not reproducible, and nerve conduction amplitudes and velocities were normal. Western blots from mouse brain lysate showed that Tars1 abundance was decreased in *Tars1*$^{G541R/+}$ mice compared to *Tars1*$^{+/+}$ littermates, suggesting that G541R Tars1 is not fully stable. This likely accounts for the absence of neuromuscular phenotypes in these mice. There are also broader limitations in modeling an axonal length-dependent phenotype in mice. Mice have shorter peripheral neurons than humans, and might require an over-expression of toxic protein to produce a phenotype. Overall, experiments assessing *TARS1* for a role in dominant peripheral neuropathy were inconclusive.

An additional loss-of-function allele, R433H *TARS1*, was identified in the initial yeast complementation assays. This allele partially supported yeast growth, indicating that it is hypomorphic. R433H was not dominantly toxic when over-expressed in *C. elegans* neurons, but did lead to gross phenotypic differences when edited into the endogenous *tars-1* locus. Homozygous R433H/R433H *tars-1* worms were significantly developmentally delayed and had locomotion defects. This indicated that R433H was a candidate for studying recessive phenotypes in a mammalian model, because it reduced protein function enough to produce a phenotype, but not enough to be lethal. This variant was generated in mouse *Tars1* using CRISPR-Cas9 editing. *Tars1*$^{R433H/R433H}$ mice were grossly normal, and were mated to mice that were heterozygous for a null *Tars1* allele (F538Kfs*4). One striking phenotype of *Tars1*$^{R433H/F538Kfs*4}$ mice was a high incidence of neonatal mortality. Histopathology of these pups found that they had no air in their lungs, a thinner dorsal epidermis, and follicular hypoplasia. *Tars1*$^{R433H/F538Kfs*4}$ mice that survived past this stage lived to adulthood, but were smaller than their *Tars1*$^{R433H/+}$ littermates. They also showed hair loss in consistent patterns on their heads and upper backs, beginning at 13 weeks old. This is particularly interesting in light of recent reports identifying bi-allelic *TARS1* mutations in patients with trichothiodystrophy, a disorder marked by brittle hair. This mouse model will be a valuable resource for future work studying the consequences of reduced *Tars* function across different cell types and tissues, particularly in the lung and hair.

5.1.3 Summary of Chapter 4

Implicating an additional homodimeric ARS, like TARS1, in dominant peripheral neuropathy is one way to strengthen the argument for a dominant-negative mechanism. Another is to test a dominant-negative effect directly, as demonstrated in Chapter 4. Here, a standard yeast complementation assay was adapted to introduce both mutant and wild-type ARS into the cell and test their combined effect. This Chapter focused on two dominant pathogenic variants in *AARS1*, G102R and R329H. Both variants repressed yeast growth when co-expressed with wild-type *AARS1*. Both variants also retained the ability to interact with wild-type AARS1, as determined by co-immunoprecipitation studies in HEK293T cells. To reduce this interaction, a

truncation mutation, Q855*, was designed in the dimerization domain. This mutation failed to support yeast growth, but still produced a stably expressed protein. It also significantly reduced dimerization with wild-type AARS. When this mutation was modeled *in cis* with G102R or R329H, the repressed yeast growth phenotype was fully rescued, demonstrating that the dimerization between the loss-of-function protein and the wild-type protein was responsible for the reduced growth. This provides direct evidence that G102R and R329H can act as dominant-negative alleles. This assay was expanded to assess additional variants in the anti-codon binding domain (R326W, R329S, and R329C), all of which similarly behaved as dominant-negative alleles.

In summary, this dissertation work made several important contributions to the field of ARS-mediated disease. Through assessing patient variants for loss-of-function effects, and through designing a model organism pipeline to implicate novel ARS in mammalian phenotypes, this work expanded the known locus, allelic, and phenotypic heterogeneity of ARS-mediated disease. It also directly tested the hypothesis that pathogenic ARS alleles act through a dominant-negative mechanism, demonstrating that pathogenic *AARS1* variants are dominant-negatives in a yeast cell. The development of this assay creates substantial opportunities to explore a dominant-negative effect as a common mechanism of pathogenic ARS alleles, as discussed below.

The author is the primary contributor to this Chapter. Matthew Pun and Molly Kuo assisted with the computational analysis of the yeast proteome.

5.2 Future Directions

Among the many remaining questions pertaining to ARS-mediated disease, there are two of particular interest: 1) What is the mechanism of dominant ARS-mediated peripheral neuropathy, and 2) How does reduced function of a ubiquitously expressed housekeeping gene affect certain tissues more than others? The experiments proposed in this Section aim to address these questions.

We propose that dominant ARS variants act through a dominant-negative mechanism, which reduces the collective amount of functional ARS protein in the cell. The large majority of pathogenic ARS variants are missense mutations that have been shown to decrease protein function, but do not impair protein stability. (The only ARS variant with multiple lines of evidence demonstrating that function is not reduced is discussed below in Section 5.2.1.) Importantly, the only five ARS genes with strong genetic evidence of dominant pathogenicity all function as homodimers. These lines of evidence are all consistent with a mechanism in which a mutant subunit of the dimer compromises the activity of the wild-type subunit. In this model, only the dimers comprising two wild-type subunits can charge enough tRNAs to meet the protein synthesis needs of the cell, which is insufficient in sensitive cells like peripheral neurons. In Chapter 4, we present evidence that five *AARS1* alleles have a dominant-negative effect in yeast. This is a promising start to testing a broadly applicable dominant-negative mechanism of disease. To strengthen this argument, additional pathogenic variants in *AARS1*, as well as *GARS1, YARS1, HARS1,* and *WARS1* need to be tested in this system. Demonstrating that pathogenic alleles across these five genes can have a dominant-negative effect in yeast would provide strong evidence that this is a common characteristic of disease alleles and as such, would suggest that a dominant-negative effect is relevant to the mechanism. Furthermore, if a dominant-negative effect that leads to reduced tRNA charging is a common mechanism, then similar loss-of-function alleles in other cytoplasmic, homodimeric ARS are also candidates for dominant peripheral neuropathy. Defining this mechanism will allow us to predict the allelic spectrum of dominant pathogenic alleles in these genes.

ARS enzymes are ubiquitously expressed house-keeping proteins required in every cell. It is currently unknown why mutations in these genes can give rise to more severe phenotypes in some tissues than others. In the case of dominant disease, the phenotype is restricted to the peripheral nervous system. This may reflect the fact that these neurons are uniquely long, terminally differentiated, and perform protein synthesis at synapses far from the soma.[439] As such, they might be uniquely sensitive to a decrease in ARS function caused by a dominant-negative allele. In the case of recessive disorders, significantly reduced function of ARS activity broadly causes neurological phenotypes. However, some tissues are particularly sensitive to decreased function of particular ARS genes (for example, the lung is most predominantly

affected in patients with *MARS1* and *FARSB* mutations[145,210]). We propose that this may be due to the production of critical proteins in these tissues that have a high requirement for the amino acid corresponding to the affected synthetase. In this model, decreased availability of the required charged tRNA may have an outsized impact on the synthesis of those critical proteins. Testing this hypothesis requires a model of ARS-mediated recessive phenotypes with specifically affected tissues, such as the compound heterozygous *Tars1* mouse described in Chapter 3.

5.2.1 Extended analysis of dominant-negative AARS1 alleles

We have developed a yeast complementation assay that can be used to study both loss-of-function and dominant-negative characteristics of *AARS1* variants. This presents the opportunity to comprehensively evaluate all known pathogenic *AARS1* alleles and to define the downstream consequences of these alleles for cellular health. These studies should be pursued with three major aims: (1) test additional dominant *AARS1* alleles for a dominant-negative effect, (2) delineate the differences between loss-of-function dominantly toxic and loss-of-function non-toxic *AARS1* alleles, and (3) define the impact of hypomorphic and dominantly toxic *AARS1* variants on protein translation.

First, additional AARS1 variants found in patients with dominant peripheral neuropathy should be assessed for a dominant-negative effect. These include N71Y and S627L, two variants that have been previously shown to reduce enzymatic function.[107,110] This is an also an opportunity to test two variants that have not yet been functionally evaluated, D893N[112] and E688G.[111] These additional variants segregate with dominant peripheral neuropathy in small pedigrees and affect highly conserved residues (E668 is conserved between humans and bacteria, and D893 is conserved between humans and fruit flies).[112,111] (Incidentally, using human AARS1 for this yeast assay allows residues that are not conserved to yeast, like D893, to be evaluated). If any of these additional *AARS1* alleles are dominantly toxic when co-expressed with wild-type *AARS1,* they can then be placed *in cis* with Q855* to determine if decreasing dimerization improves yeast growth (this test does not apply to D893N, which would be cut off by the Q855*

truncation). Identifying additional dominant-negative *AARS1* alleles would provide strong evidence of a loss-of-function disease mechanism.

Another possibility is that these variants may have no effect, or increase enzymatic activity like E337K *AARS1*. E337K *AARS1* was identified in a small pedigree with five genotyped affected individuals and only one genotyped, unaffected individual.[110] All affected individuals were heterozygous for E337K *AARS1*, while the one unaffected individual did not carry the mutation.[110] The E337 residue is conserved between humans and fruit flies,[110] and is absent from gnomAD.[334]. E337K increases yeast growth in a complementation assay and increases *in vitro* aminoacylation activity, acting as a hypermorph.[110] E337K is the only neuropathy-associated ARS variant known to cause a gain of enzymatic function; as an outlier, it will be important to continue to strengthen the genetic argument for pathogenicity by identifying additional patients with this variant.

To functionally investigate E337K *AARS1* further, it will be informative to co-express it with wild-type *AARS1* in yeast to determine if it causes a dominantly toxic effect. If so, there are at least two possible explanations. The first possibility is that exceeding cellular demands for changed $tRNA^{Ala}$ is actually detrimental to protein translation. The availability of tRNA for a given codon determines ribosome speed.[440] Codons corresponding to highly abundant tRNAs are translated faster than codons corresponding to lowly abundant tRNAs (rare codons).[440,441] Local pauses or slow-downs in ribosomal speed can provide a chance for the polypeptide that is being translated to fold properly.[441,442] These slow-downs are often found when translating across the domain boundaries of multi-domain proteins, allowing time for the independent folding of the previously translated domain.[441] Silent mutations that alter translational speed by swapping a rare codon for a common one, or vice versa, can impact the folding and function of disease-associated proteins.[443] If E337K *AARS1* increases the proportion of charged $tRNA^{Ala}$ in the tRNA pool, it may disrupt these carefully coordinated translation dynamics by increasing the translation speed through alanine codons. This could disrupt the timing of co-translational protein folding and, for proteins with sensitive folding dynamics, may cause them to misfold and be targeted for degradation. Alternately, E337K *AARS1* could cause a broader increase in protein misfolding, trigger protein aggregations and/or inducing the unfolded protein response pathway

in the endoplasmic reticulum, which would trigger PERK to phosphorylate eIF2α and induce a global repression of translation initiation through preventing the recycling of eIF2.[351]

Alternately, E337K could cause aberrant interactions with some component of the protein translation machinery, resulting in similar decreases in protein translation. This interaction could be increased binding to $tRNA^{Ala}$, which could decrease the availability of free tRNA for wild-type AARS1; *in vitro* aminoacylation studies found that E337K AARS1 exhibited a 5-fold decrease in K_m for tRNA relative to wild-type AARS1, indicating increased affinity for tRNA.[110] However, this did not result in increased turnover (k_{cat} for E337K was 0.4, compared to 0.6 for wild-type).[110] This data indicate that E337K could be dominantly toxic through altering tRNA binding dynamics. Another possibility is that E337K AARS1 causes aberrant interactions with ribosomes or other components of the translation machinery, leading to impaired protein synthesis similar to dominant-negative alleles. To further investigate these possibilities, co-immunoprecipitation studies should be performed with E337K AARS1 to investigate protein binding partners using mass spectrometry, or tRNA binding partners using either Northern blots hybridization or RNA-seq methods suitable for detecting tRNA.[444]

Second, the yeast assay developed in this thesis should be applied to compare loss-of-function toxic *AARS1* alleles and loss-of-function non-toxic *AARS1* alleles. It is currently unclear what distinguishes these two types of variants. The heterozygous parents of individuals with recessive disease have not been reported to have a peripheral neuropathy, although most have not been clinically evaluated, and there may be undetected late-onset or sub-clinical phenotypes. One possibility is that recessive missense variants are destabilizing, which would mean that an otherwise dominantly toxic protein does not accumulate to sufficient levels to have a toxic effect. Indeed, decreased protein stability has been demonstrated for some variants—individuals who are compound heterozygous for I699T and C901Y *AARS1*, or for T726A and T756I *AARS1* show a 60-80% reduction in total AARS1 protein.[105] The G931D *AARS1* variant also reduces AARS1 levels; individuals with the G931D/+ *AARS1* genotype express less AARS1 protein than +/+ individuals, and individuals with the G931D/Y690Lfs*3 *AARS1* genotype express less AARS1 protein than +/Y690Lfs*3 individuals.[103] Another possibility is that recessive *AARS1* variants

reduce dimerization with the opposite subunit. In this scenario, they might only cause a very mild dominant negative effect in heterozygous carriers, or no dominant-negative effect at all.

To test these two hypotheses, a panel of recessive missense *AARS1* alleles should be expressed in yeast to evaluate the total amount of AARS1 protein generated by each isolated allele. Any stable recessive variants should then be co-expressed with the wild-type allele in a dominant-negative assay, with the hypothesis that only dominant—but not recessive—missense alleles will have a dominant-negative effect. If recessive variants do act as dominant-negative alleles, it is possible that: (1) the variant may be stably expressed in yeast cells but not in human cells, (2) the variant may only cause weakly dominant-negative effects in a human peripheral neuron that is not sufficient to cause peripheral neuropathy, or (3) individuals heterozygous for this variant do develop neuropathy that is late-onset, mild, and/or goes undetected. However, if stably expressed recessive variants do not act as dominant-negative alleles, they should be further investigated to determine if they reduce dimerization. Here, the most sensitive assay to use is ultracentrifugation, which has previously been used to study dimerization of GARS[300] and HARS.[176,350]

Finally, generating appropriate models of both dominant and recessive *AARS1* alleles in yeast will allow us to define the impact of these variants on protein translation, as discussed below. Ideally, these studies should be performed with a hypomorphic recessive allele, a dominant-negative allele like R329H (co-expressed with wild-type), and E337K (co-expressed with wild-type). We would expect that a recessive allele and a dominant-negative allele would have similar effects on protein translation. This would also be an opportunity to test the hypotheses proposed for E337K by determining if it affects protein translation, or acts through some other mechanism.

Loss-of-function pathogenic *AARS1* alleles may impair protein synthesis by increasing the pool of uncharged $tRNA^{Ala}$ in the cell. This is recognized by GCN2, stimulating its kinase function to phosphorylate eIF2α, which then prevents eIF2B from recycling GDP-eIF2 into GTP-eIF2.[3] This prevents eIF2 from delivering initiator $tRNA^{Met}$ to the ribosome to begin protein translation.[3] To determine whether this pathway is activated in yeast expressing pathogenic *AARS1* alleles, yeast lysate should be evaluated via Western blot for an increase in eIF2α phosphorylation. If yeast

expressing the pathogenic *AARS1* alleles have increased eIF2α phosphorylation, this indicates that there is likely reduced protein synthesis in the cell. Here, we would expect that a hypomorphic *AARS1* allele and a dominant-negative *AARS1* allele would both increase eIF2α phosphorylation. If E337K causes widespread protein misfolding, as proposed above, it may also trigger eIF2α phosphorylation through the unfolded protein response pathway.[351] Alternately, if E337K acts by causing localized protein misfolding, sequestering tRNA or aberrantly binding to other components of translation machinery, we would not expect to see an increase in eIF2α.

Subsequent effects on global protein translation can be tested by treating yeast with puromycin, an aminonucleoside that can be incorporated into nascent peptides.[445] After treatment, yeast are lysed and puromycin incorporation is measured with immunoblotting.[445] Here, we would expect that a hypomorphic, recessive *AARS1* allele and a dominant-negative *AARS1* allele would both show a large decrease in puromycin incorporation. If E337K acts to repress protein synthesis (either by triggering the unfolded protein response, sequestering tRNA, or aberrantly binding to translation machinery), we would also expect it to show decreased puromycin incorporation. These observations, in tandem with the eIF2α phosphorylation, would show that the dominant-negative *AARS1* allele reduces protein translation similar to a recessive hypomorphic *AARS1* allele. These observations would also indicate that, although E337K is not a dominant-negative allele, it causes similar defects in protein translation, which would help identify a unifying cellular pathology.

If *AARS1* alleles do not show global reductions in protein synthesis, this may indicate that there are amino-acid specific defects in protein translation. In theory, reducing the amount of amino acid that is ligated to tRNAs should have an effect similar to starving the cell for that amino acid. Here, studies that decrease arginine availability[304] or inhibit histidine biosynthesis[446] have demonstrated that these amino acid restrictions cause ribosomes to stall at histidine or arginine codons, respectively. Therefore, we might expect cells expressing pathogenic loss-of-function *AARS1* alleles to stall at alanine codons, which could be detected with ribosome sequencing. Here, we would expect that yeast expressing a hypomorphic recessive *AARS1* allele and a dominant-negative *AARS1* allele to show increased pile-up of ribosomes at alanine codons, relative to yeast expressing wild-type *AARS1*. Ribosome profiling may also be a high-resolution

approach to discern the mechanism of E337K. For example, if E337K decreases the amount of charged $tRNA^{Ala}$ in the cell by inappropriately sequestering uncharged $tRNA^{Ala}$, we would expect to see an increase in ribosome stalling similar to the loss-of-function alleles. On the other hand, if E337K overproduces charged $tRNA^{Ala}$ and interferes with ribosome kinetics, we would expect to see a *decrease* in ribosome abundance at alanine codons compared to the wild-type allele. In this case, the identity of transcripts with this decreased ribosome abundance should be further scrutinized to determine if there is a corresponding decrease in protein abundance. If *AARS1* alleles do not cause detectable ribosome stalling, this may reflect either a global repression of protein synthesis through eIF2α phosphorylation, or rapid turnover of stalled ribosomes that cannot be detected by ribosome profiling. In this case, such experiments may need to be performed in a yeast background that lacks factors that can dissociate stalled ribosomes (Dom34, Hbs1, and Rli1).[446]

If hypomorphic and dominant-negative *AARS1* alleles cause ribosome stalling at alanine codons, we would expect yeast proteins with a high alanine content to be most affected. These proteins are primarily ribosomal proteins and cell wall proteins (Table 5.1). In lieu of commercially available antibodies for yeast cell wall proteins, these genes could be endogenously tagged to facilitate immunostaining and determine whether expressing *AARS1* alleles decreases their abundance. Alternately, the fitness of the cell wall could be evaluated using stains such as calcofluor white, which produces greater signal when there is cell wall stress, or aniline blue and FITC-ConA, which stain for 1,3-β-glucan and mannoproteins, respectively.[447] Here, we would expect a recessive *AARS1* allele and a dominant-negative *AARS1* to decrease cell wall integrity, which could contribute to decreased cell viability. In the case of E337K, if it acts to decrease the amount of charged $tRNA^{Ala}$ in the cell, we would also expect to see impaired cell wall integrity. However, if it triggers misfolding of specific proteins, or global protein misfolding, we would not necessarily expect this to impact on cell wall proteins.

5.2.2 Identifying dominant-negative alleles in GARS1, HARS1, WARS1, and YARS1.

Table 5.1. Alanine-rich proteins in *S. cerevisiae*.

Name	% alanine
CWP2 Cell wall protein*	26
HOR7 Hypo-osmolarity responsive protein	24
RLA3 60S acidic ribosomal protein P1-beta	23
PAU7 Seripauperin-7*	22
CCW12 Covalently-linked cell wall protein 12*	22
TIR3 Cell wall protein TIR3*	22
DDR2 Protein DDR2	21
FIT3 Facilitator of iron transport 3*	21
RLA4 60S acidic ribosomal protein P2-beta	21
TIR1 Cold shock-induced protein TIR1*	21
RLA2 60S acidic ribosomal protein P2-alpha	21
TIP1 Temperature shock-inducible protein 1*	20
TIR2 Cold shock-induced protein*	20
TOM6 Mitochondrial import receptor subunit	20
RLA1 60S acidic ribosomal protein P1-alpha	19
HS150 Cell wall mannoprotein HSP150*	18
PIR3 Cell wall mannoprotein PIR3*	18
ATP5E ATP synthase subunit epsilon, mitochondrial	18
PAU5 Seripauperin-5*	17
DAN1 Cell wall protein DAN1*	17
PAU17 Seripauperin-17*	17
RL13B 60S ribosomal protein L13-B	17
LSO2 Protein LSO2	16
PAU16 Seripauperin-16*	16
PAU19 Seripauperin-19*	16
PAU3 Seripauperin-3*	16

*Cell wall proteins

The humanized yeast assay to detect dominant-negative alleles should next be expanded to assess pathogenic alleles in other dominant ARS disease genes. Here, the initial focus should be on variants with strong genetic evidence of pathogenicity, such as G240R[160,448] and G526R[449] in *GARS1,* T132I[174] and D364Y[174] in *HARS1*, D314G[265] and H257R[264] in *WARS1,* and G41R and E196K in *YARS1*.[379275] For studies of *HARS1, WARS1,* and *YARS1*, yeast strains with a doxycycline-repressible promoter driving *HTS1, WRS1,* and *TYS1* are available commercially from the Hughes Collection.[435] Although such a strain does not exist for *GARS1* studies, another graduate student in the Antonellis group, Sheila Marte, is currently working to replace the endogenous *GRS1* promoter with a doxycycline-repressible promoter.

The above neuropathy-associated ARS variants are loss-of-function alleles in homodimeric enzymes; therefore, we expect them to act as dominant-negative alleles when co-expressed with the respective wild-type gene, similar to the dominant-negative variants in *AARS1*. Interestingly, there is some preliminary evidence to suggest that G41R and E196K *YARS1* show properties of dominant-negatives alleles when tested in yeast. When each variant was modeled in the yeast ortholog *TYS1* and over-expressed in yeast with wild-type *TYS1*, they repressed yeast growth.[275] Similarly, yeast expressing wild-type *TYS1* and G41R *YARS1* grew slower than yeast expressing wild-type *TYS1* and wild-type *YARS1*, or yeast expressing *TYS1* and E196K *YARS1*.[275] This could indicate that G41R *YARS1* may also repress the endogenous *TYS1,* acting similar to R329H *AARS1* (as discussed in Chapter 4).

To conclusively determine that dominant toxicity is a result of a dominant-negative effect, as opposed to some other gain of function, it is necessary to reduce the dimerization of the mutant allele and determine if this rescues the phenotype. In the cases of *GARS1* and *YARS1*, dimer-reducing alleles have already been identified. The T209K *GARS1* variant was identified as a recessive allele in zebrafish *gars*, and was shown to significantly reduce the percentage of dimeric gars in native western blots. Additionally, P167T *YARS1*, which causes a severe multisystem disease in the homozygous state, was shown to reduce binding to wild-type YARS1 in co-immunoprecipitation assays.[273] Identifying dimer-reducing variants in *HARS1* and *WARS1* may be more challenging. Here, the published crystal structures of human WARS1[60] and human HARS1[41] will be valuable resources for identifying residues to target that could disrupt

dimerization. In these structures, the dimer interface is internal, and does not lend itself to large domain deletions. Alanine scanning may be the best approach, followed by yeast complementation assays and immunoblotting to identify loss-of-function, stably expressed alleles. Once these alleles have been identified, they can be tested for reduced binding to the wild-type protein using the co-immunoprecipitation approach described in Chapter 4.

If multiple pathogenic variants across multiple ARS genes are shown to act as dominant-negative alleles, this would provide strong evidence that this is a relevant disease mechanism. Importantly, benign polymorphisms selected from gnomAD should be included in assessments for each of these genes, to demonstrate the specificity of a dominant-negative effect. These data would indicate that reduced ARS function is likely a central component of disease pathology, which would support previous observations of impaired protein translation in neuronal models of dominant ARS disease.[306,350] Alternately, yeast may be more sensitive to reduced function of some ARS than others, and it may be difficult to detect dominant-negative effects for all variants in all ARS genes. It is also possible that some pathogenic ARS variants may have an alternative mechanism of action; if any variants cause dominant toxicity that is not rescued by reducing dimerization, this could provide an opportunity to explore other gain-of-function hypotheses, such as interference with other components of the translation machinery.

5.2.3 Defining a panel of dominant-negative GARS1 alleles

One goal of the Antonellis group is to use high-throughput techniques to identify all loss-of-function variants in pathogenic ARS genes. This will circumvent the need to individually test patient variants for loss-of-function effects in yeast, which is not a feasible approach to assess all possible variants of uncertain significance. In collaboration with Jacob Kitzman's group, a postdoctoral fellow in the Antonellis laboratory, Megan Forrest, is performing saturation genome editing on *GARS1 in vitro* to identify all variants that do not support cell growth. This work will provide novel insights into *GARS1* biology, as well as help separate out benign variants and deleterious variants. However, as discussed previously, we do not expect all loss-of-function variants to be dominantly toxic, and as such, it will be difficult to differentiate between candidates for recessive disease and candidates for dominant disease.

Here, the yeast dominant toxicity assay can play a significant role. Any loss-of-function variants identified in *GARS1* saturation genome editing should first be assessed for an effect on protein stability, using a flow-sorting method such as VAMP-seq[450] that is amenable to testing a large library of variants. Then, a smaller panel of variants that are loss-of-function but stably expressed can be tested in the yeast dominant toxicity assay. Although this assay is low-throughput, this assay does have the advantage of being relatively inexpensive and quick. If mutagenesis and cloning were performed in a high-throughput fashion, then over 100 variants could be tested in 1-2 months. The benefits of overlapping a yeast dominant toxicity assay with the saturation mutagenesis efforts are two-fold: (1) *GARS1* variants that are loss-of-function but dominantly toxic will be pre-emptively defined, providing support for pathogenicity if they are identified in patients with dominant peripheral neuropathy, and (2) distinguishing loss-of-function, non-dominant variants from loss-of-function dominant variants will yield greater insight into the mechanism of *GARS1* pathogenicity. If both types of variants are mapped onto the GARS1 crystal structure, this could also provide insights into the structural features that play a role in disease.

5.2.4 Predicting dominant-negative alleles in additional homodimeric, cytoplasmic ARS

Once the yeast dominant-negative assay has been appropriately validated for a series of pathogenic alleles in ARS genes already implicated in dominant neuropathy, this assay can then be used to predict and assess variants in ARS genes that have not yet been implicated in dominant disease, particularly other the five remaining homodimeric, cytoplasmic ARS (*TARS1, CARS1, KARS1, DARS1,* and *SARS1*). This could produce preliminary evidence that they are candidate genes for dominant peripheral neuropathy. This assay could be incorporated into the model organism prediction pipeline described in Chapter 2, to help prioritize candidate mutations for testing in worm and mouse.

The above efforts should begin with *TARS1*; the appropriate yeast strain is available from the Hughes Collection[435] and the crystal structure of TARS1 is available[56] to inform careful design of loss-of-function mutations, as well as mutations that could decrease dimerization (a critical

part of demonstrating a dominant-negative effect, as discussed above). Currently, another graduate student in the Antonellis group, Allison Cale, is developing these tools. After designing *TARS1* variants at conserved residues and testing them for impaired function, it will be important to assess the variant's effect on protein stability. Then, the yeast dominant-negative assay can be used to identify the best candidates to test further, which may improve the likelihood that the selected variant is dominantly toxic to neurons in a multicellular organism.

All of the remaining homodimeric ARS enzymes should also be evaluated in this system. Here, it would be helpful to begin with the ARS loci that include peripheral neuropathy as part of a recessive pathology. This might indicate that peripheral nerves are sensitive to decreased function of that ARS. This applies to patients with bi-allelic variants in cysteinyl-tRNA synthetase (*CARS1*)[126] and lysyl-tRNA synthetase (*KARS1*).[195] The remaining cytoplasmic, homodimeric enzymes are aspartyl-tRNA synthetase (DARS1) and seryl-tRNA synthetase (SARS1), which could be targeted next. For *CARS1, KARS1, DARS1,* and *SARS1*, yeast strains with a doxycycline-repressible promoter driving the respective ARS ortholog are commercially available[435]; however, any other methods of repressing endogenous yeast ARS function while over-expressing both mutant and wild-type human ARS would also be valid approaches.

5.2.5 Detecting ARS dominant toxicity in cultured mammalian cells

As the dominant-negative assay expands to additional ARS, it would be valuable to develop an orthogonal approach in human cell lines. This would obviate the need for developing and validating numerous yeast strains. A cell line should be selected with a single copy of the wild-type ARS gene. HAP1 cells would be the most feasible and applicable across different ARS genes. (Alternately CRISPR-Cas9 editing could be used to reduce the number of functional ARS copies in multi-ploidal cell lines, including neuroblastoma cell lines that can be differentiated to harbor neurite projections.) Then, wild-type or mutant ARS expressed by a doxycycline-inducible promoter could be transfected into the cell population, with a drug resistance cassette or GFP marker to select for the transfected population. Application of doxycycline to the cells would induce expression of wild-type or mutant ARS, and dominant toxicity would be measured by determining cell viability. This approach should first be validated with known pathogenic

dominant-negative alleles like R329H *AARS1*, but then could be expanded as a predictive assay for novel ARS loci and alleles. Moreover, because cell lines are more amenable to high-throughput analysis of ARS loci, many variants could be assessed simultaneously. In this scenario, comparing the frequency of alleles before and after doxycycline induction would allow identification of the alleles that were dominantly toxic and reduced cell viability. In combination with saturation genome editing, this approach could be a high-throughput assay to differentiate between loss-of-function, dominantly toxic alleles and loss-of-function, non-toxic alleles, as discussed above for *GARS1*.

5.2.6 Defining the requirement for threonyl-tRNA synthetase in the mouse lung and hair follicle

The experiments proposed above would significantly advance our understanding of the allelic spectrum of dominant ARS-mediated disease, as well as the mechanism of disease. We predict that mild to moderate impairment of protein synthesis, caused by reduction in any population of charged tRNA, specifically impacts the peripheral nervous system due to the long, terminally differentiated axons that perform local protein synthesis at synapses.[439] This is in contrast to our proposed mechanism of ARS-mediated recessive disease, in which a greater reduction of ARS function causes severe impairment of protein synthesis, affecting numerous tissues across the body that have high protein translation demands. However, there are still outstanding questions as to why pathogenic variants in some ARS genes affect certain tissues more than others. For example, it is unclear why the lung is particularly sensitive to mutations in *MARS1*[210], or why the liver is particularly sensitive to mutations in *LARS*.[196] One hypothesis is that there are critical proteins in these tissues with a high requirement for a certain amino acid, and that such proteins are therefore particularly sensitive to reduced function of the corresponding ARS.

The phenotypes of the *Tars*$^{R433H/F538Kfs*4}$ mouse (Section 3.3.12) may provide some clues. One of the most striking phenotypes was premature death at P0. Histopathology of these mice showed that there was no air present in the alveoli, indicating that mice either died *in utero* or upon birth. There was also an absence of PAS+ club cells in the bronchioles, indicating a lack of glycoproteins and/or mucins. Intriguingly, these proteins are threonine-rich; the backbone of mucin proteins is composed of tandem proline, threonine, and/or serine repeats (PTS domains).

These PTS domains are densely O-glycosylated, which contribute to the structure and function of mucous membranes.[412] We hypothesize that reduced TARS1 function preferentially impacts the production of these mucin proteins, impairing the formation or maintenance of the mucosal linings in the lungs.

Initial research into hypothesis should include immunohistochemistry staining for a series of mucin proteins expressed in the lung, including mucin-4, mucin-20, and endomucin (see Table 3.2). If there is reduced mucin protein expression, protein mass spectrometry of lung tissue should be performed and the threonine content of misregulated proteins should be analyzed. If there is a significant depletion of threonine-rich proteins, this would indicate an amino acid specific defect in protein synthesis. If there is no significant depletion for threonine-rich proteins, this may indicate a global impairment in protein synthesis that happens to include mucin proteins. In either case, this analysis may provide insight into the pathways dysregulated in the lung tissue of these mice.

It will also be important to determine if the absent PAS+ stain reflects the absence of club cells, or merely a lack of PAS+ material within them. This could be done by performing immunohistochemistry for SCGB1A1, the primary product of club cells (which is only 7% threonine, and therefore unlikely to be targeted by a threonine-specific mechanism). Club cells are important in maintaining lung homeostasis and protection against infection. They are also stem cells.[451] Club cells arise from the primary stem cell population, basal cells; then, club cells can differentiate into either ciliated cells or the mucin-producing goblet cells.[451] If club cell abundance is reduced in *Tars*$^{R433H/F538Kfs*4}$ mice, this could indicate an impairment in stem cell proliferation or differentiation. Indeed, stem cells may be particularly sensitive to reductions in ARS function—a recent investigation of the microcephaly phenotype in recessive *NARS1*-mediated disease demonstrated that in brain organoids derived from patient cells, radial glial cells showed reduced viability and proliferation.[221] More broadly, undifferentiated stem cells have been found to maintain a low basal rate of translation, which increases upon differentiation.[452,453,454,455] For example, in quiescent skeletal muscle stem cells, eIF2α phosphorylation maintains stem cell state; loss of phosphorylation by replacing the phosphorylated eIF2α serine residue with alanine causes differentiation and activation of the

myogenic program.[456] One explanation for a possible lack of club cells in *Tars*$^{R433H/F538Kfs*4}$ mice is that the uncharged levels of tRNA^{Thr} in basal stem cells maintain a state of global translation repression through eIF2α phosphorylation, preventing differentiation into club cells.

Another indication of impaired stem cell function in *Tars*$^{R433H/F538Kfs*4}$ mice may be found in their hair phenotype. 71% of adult *Tars*$^{R433H/F538Kfs*4}$ mice lose hair on their heads and upper back by 23 weeks of age (Figure 3.18); however, significant work is needed to further define this phenotype. One explanation is that the reservoir of hair stem cells is affected by impaired *Tars* function, and may be unable to differentiate and generate new hair follicles. A study of protein synthesis in mouse hair follicles found that translation was highly dynamic across the stages of the hair follicle life cycle.[457] The authors used OP-puromycin incorporation to quantify the rate of nascent protein synthesis across bulge stem cells and differentiated cells. They found that protein synthesis increased in the transition from the quiescence stage (telogen) to growth stage (anagen), and was highest in differentiated cell populations in the growth stage. This data suggests there is a correlation between hair stem cell activation and increased protein translation rate. Here, as with basal cells in the lung, reduced Tars function may impair the ability of bulge stem cells to exit quiescence by preventing an increase in global protein translation needed to fully execute the differentiation program. In both cases, immunostaining of different cell populations in both the lung and the hair follicle may provide more information on whether there are increased populations of stem cells and/or decreased populations of differentiated cells. It would also be informative to measure OP-puromycin incorporation in differentiating lung or hair cells and determine if these cells show reduced global protein synthesis compared to their control littermates.

5.3 Concluding Remarks

The work presented in this dissertation addressed three aims: 1) expanding the allelic, locus, and phenotypic heterogeneity of ARS-mediated disorders, 2) predicting the pathogenicity of ARS loci that have not yet been implicated in disease, and 3) testing high-confidence pathogenic ARS alleles for a dominant-negative effect. Here, we have characterized newly identified patient alleles in *GARS1, HARS1, MARS1, TARS1,* and *NARS1*, and provided an assessment of their

pathogenicity. We have also developed a pipeline of model organism assays to predict the pathogenicity of dominant and recessive *TARS1* alleles. Through these efforts, we have produced new models of *TARS1*-mediated recessive phenotypes that will be assets in investigating the effects of reduced *TARS1* activity on cellular function. Finally, we have pioneered a yeast assay to assess human ARS alleles for a dominant-negative effect, and demonstrated that pathogenic *AARS1* variants are dominant-negative alleles. This work strengthens the evidence for a common, loss-of-function mechanism of ARS-mediated disease.

Appendix A

Oligonucleotide Sequences

Table A.1 Gateway cloning primers for yeast complementation constructs. Primers are listed 5' to 3'.

Primer	Sequence
MARS1 ORF GW F	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGAAGGAGATAGAACCATGAGACTGTTCGTGAGTGATGG
MARS1 ORF GW R	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTACTTTTTCTTCTTGCCTTTAGG
HARS1 ORF GW F	GGGGACAAGTTTGTACAAAAAAGCAGGCTCAATGGCAGAGCGTGCGGCGCTGG
HARS1 ORF GW R	GGGGACCACTTTGTACAAGAAAGCTGGGTTTCAGCAGATGCAGAGGGGCTGG
NARS1 ORF GW F	GGGGACAAGTTTGTACAAAAAAGCAGGCTCTATGGTGCTAGCAGAGCTGTACG
NARS1 ORF GW R	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTATGGCGTGCAACGCTGGACA
TARS1 ORF GW F	GGGGACAAGTTTGTACAAAAAAGCAGGCTCAATGTTTGAGGAGAAGGCCAGC
TARS1 ORF GW R	GGGGACCACTTTGTACAAGAAAGCTGGGTTTTAAAATTCTTCTTCTGC
scALA1 6xHis GW F	GGAGAAGCTTTCCATTCATCATCACCATCACCATTAAGACCCAGCTTTCT
scALA1 6xHis GW R	AGAAAGCTGGGTCTTAATGGTGATGGTGATGATGAATGGAAAGCTTCTCC

Table A.2 Mutagenesis primers for yeast complementation assays. Primers are listed 5' to 3'.

Primer	Sequence
G327R *GARS1* Mut F	GCTGCTGCCCAGATTAGAAATTCTTTTAGAA
G327R *GARS1* Mut R	TTCTAAAAGAATTTCTAATCTGGGCAGCAGC
V155G *HARS1* Mut F	ACCACATAGCAAAGGGATATCGGCGGGATAA
V155G *HARS1* Mut R	TTATCCCGCCGATATCCCTTTGCTATGTGGT
Y330C *HARS1* Mut F	CTCGAGGGCTGGATTGCTACACTGGGGTGAT
Y330C *HARS1* Mut R	ATCACCCCAGTGTAGCAATCCAGCCCTCGAG
S356N *HARS1* Mut F	CCCTGGGTGTGGGCAATGTGGCTGCTGGAG
S356N *HARS1* Mut R	CTCCAGCAGCCACATTGCCCACACCCAGGG
A397T *MARS1* Mut F	TGTGCTCGCTTCCTGACTGACCGCTTCGTGG
A397T *MARS1* Mut R	CCACGAAGCGGTCAGTCAGGAAGCGAGCACA
R618C *MARS1* Mut F	CCTGCTGACATCTGGTGCTTCTATCTGCTGT
R618C *MARS1* Mut R	ACAGCAGATAGAAGCACCAGATGTCAGCAGG
Y307C *MARS1* Mut F	AGTGGAACACCCTCTGTCTGTGTGGGACAGA
Y307C *MARS1* Mut R	TCTGTCCCACACAGACAGAGGGTGTTCCACT
S461F *NARS1* Mut F	CCCGTCTTACTGAATTTGTCGACGTGTTGAT
S461F *NARS1* Mut R	ATCAACACGTCGACAAATTCAGTAAGACGGG
C342Y *NARS1* Mut F	CACGTGGAAGCTGAGTATCCTTTCCTGACTTT
C342Y *NARS1* Mut R	AAAGTCAGGAAAGGATACTCAGCTTCCACGTG
ΔM236 *NARS1* Mut F	CAACAACAGACACATGATCCGAGGAGAAAA
ΔM236 *NARS1* Mut R	TTTTCTCCTCGGATCATGTGTCTGTTGTTG
R131H *TARS1* Mut F	TGTGGGACCTGGACCACCCTCTGGAAGAAGA
R131H *TARS1* Mut R	TCTTCTTCCAGAGGGTGGTCCAGGTCCCACA
V372I *TARS1* Mut F	AGAGGATTCCAGGAGATAGTCACCCCAAACA
V372I *TARS1* Mut R	TGTTTGGGGTGACTATCTCCTGGAATCCTCT
R619C *TARS1* Mut F	TTTTGGCTGTCCCCTTGCCAGGTAATGGTAG
R619C *TARS1* Mut R	CTACCATTACCTGGCAAGGGGACAGCCAAAA
Q639P *TARS1* Mut F	CCCAAAAGGTACGACCACAATTCCACGATGC
Q639P *TARS1* Mut R	GCATCGTGGAATTGTGGTCGTACCTTTTGGG
R663Q *TARS1* Mut F	TGAATAAAAAGATTCAAAATGCACAGTTAGC
R663Q *TARS1* Mut R	GCTAACTGTGCATTTTGAATCTTTTTATTCA
G541R *TARS1* Mut F	GATGGAGCTTTCTATCGCCCAAAGATTGACATAC
G541R *TARS1* Mut R	GTATGTCAATCTTTGGGCGATAGAAAGCTCCATC
R433H *TARS1* Mut F	CGAGAACTGCCTCTGCACCTAGCTGATTTTGGG
R433H *TARS1* Mut R	CCCAAAATCAGCTAGGTGCAGAGGCAGTTCTCG
N412Y *TARS1* Mut F	CCTGAAACCCATGTACTGCCCAGGACACTGC
N412Y *TARS1* Mut R	GCAGTGTCCTGGGCAGTACATGGGTTTCAGG
G757* *AARS1* Mut F	ATTGTGGCTGTCACATGACTGGTGCCGAGGCCCAG
G757* *AARS1* Mut R	CTGGGCCTCGGCACCAGTCATGTGACAGCCACAAT
R329H *AARS1* Mut F	TGAGACGGATTCTCCACCGAGCTGTCCGATA

R329H *AARS1* Mut R	TATCGGACAGCTCGGTGGAGAATCCGTCTCA
G102R *AARS1* Mut F	TTCTTCGAGATGCTGCGCTCTTGGTCTTTTG
G102R *AARS1* Mut R	CAAAAGACCAAGAGCGCAGCATCTCGAAGAA
R326W *AARS1* Mut F	GGATATGTGTTGAGATGGATTCTCCGCCGAG
R326W *AARS1* Mut R	CTCGGCGGAGAATCCATCTCAACACATATCC
R329S *AARS1* Mut F	TTGAGACGGATTCTCAGCCGAGCTGTCCGAT
R329S *AARS1* Mut R	ATCGGACAGCTCGGCTGAGAATCCGTCTCAA
R329C *AARS1* Mut F	TTGAGACGGATTCTCTGCCGAGCTGTCCGAT
R329C *AARS1* Mut R	ATCGGACAGCTCGGCAGAGAATCCGTCTCAA
Q855* *AARS1* Mut F	GACAGCAACCCCAACTAGCCTCCTGTCATCC
Q855* *AARS1* Mut R	GGATGACAGGAGGCTAGTTGGGGTTGCTGTC
ΔKNVGCLQ *AARS1* Mut F	GCACAGGCCACAGGCGAGGCGCTGCAGCTG
ΔKNVGCLQ *AARS1* Mut R	CAGCTGCAGCGCCTCGCCTGTGGCCTGTGC
ΔNVG *AARS1* Mut F	CAGGCCACAGGCAAGTGCCTGCAGGAGGCG
ΔNVG *AARS1* Mut R	CGCCTCCTGCAGGCACTTGCCTGTGGCCTG
ΔC947 *AARS1* Mut F	GGCAAGAACGTTGGCCTGCAGGAGGCGCTG
ΔC947 *AARS1* Mut R	CAGCGCCTCCTGCAGGCCAACGTTCTTGCC
ΔQE *AARS1* Mut F	AACGTTGGCTGCCTGGCGCTGCAGCTGGCC
ΔQE *AARS1* Mut R	GGCCAGCTGCAGCGCCAGGCAGCCAACGTT

Table A.3 Gateway cloning primers for *C. elegans* expression constructs. Primers are listed 5' to 3'.

Primer	Sequence
c.e. tars-1 GW F	GGGGACAAGTTTGTACAAAAAAGCAGGCTCAAAAAATGCGATT GAACTGTTTCC
c.e. tars-1 GW R	GGGGACCACTTTGTACAAGAAAGCTGGGTTTTATGCCCATTCCT CAGACTTTTCC

Table A.4 Mutagenesis primers for *tars-1* cDNA. Primers are listed 5' to 3'.

Primer	Sequence
c.e. tars-1 G540R Mut F	AGAAATGAAATGTCTAGAGCTTTGACTGGACTT
c.e. tars-1 G540R Mut R	AAGTCCAGTCAAAGCTCTAGACATTTCATTTCT
c.e. tars-1 N411Y Mut F	GGTTTGAAGCCGATGTACTGCCCGGGGCATT
c.e. tars-1 N411Y Mut R	AATGCCCCGGGCAGTACATCGGCTTCAAACC
c.e. tars-1 R432H Mut F	ATGAGCTTCCATTCCATTTCGCTGATTTTGGA
c.e. tars-1 R432H Mut R	TCCAAAATCAGCGAAATGGAATGGAAGCTCAT

Table A.5 Single guide RNA sequences for *C. elegans* gene editing. Primers are listed 5' to 3'.

Primer	Sequence
G540R sgRNA	CCGGGAGATGGAGCATTCTACGG
R432H sgRNA	CCATTCCGATTCGCTGATTTTGG

Table A.6 Single stranded oligonucleotide repair templates for *C. elegans* gene editing. Primers are listed 5' to 3'.

Primer	Sequence
G540R ssODN	TCAGGCAGAAAATGGGTGTTGAACCCGGGAGATGGAG CATTCTACCGGCCGAAAATCGATATCACCATTCAAGAT GCTCTCAAGAGAAA
R432H ssODN	GTCTCATGTTCGGACACATGCCACACACCTACAATGAG CTCCCGTTTCATTTTGCAGACTTTGGAGTTTTGCACAGA AATGAAATGTCTGGTGCTTTGACTG

Table A.7 PCR primers for *tars-1* G540R locus and *tars-1* R432H locus (*C. elegans*). Primers are listed 5' to 3'.

Primer	Sequence
G540R *tars-1* F	CGTCAAGATCAAATTTCCGAGG
G540R *tars-1* R	CGAGTCAGTCAGAGACTATGCA
R432H *tars-1* F	ATCGGCGTTCTGGTATCC
R432H *tars-1* R	CCAATGCGCTCAACGCTT

Table A.8 PCR primers for mouse *Tars1* G540R and F538Kfs*4 locus, and *Tars1* R432H locus. Primers are listed 5' to 3'.

Primer	Sequence
G540R and F538Kfs*4 *Tars* F	GTGTCGAGTGAGGGAATGTTTTCTGTA
G540R and F538Kfs*4 *Tars* R	TTAAACCTGATGGGCAACTGAAA
R432H *Tars* F	CTTCTATTGCTCTGTGAGGC
R432H *Tars* R	TATCTCCCTTCACCTGTAACC

Appendix B

Full Western Blot Images

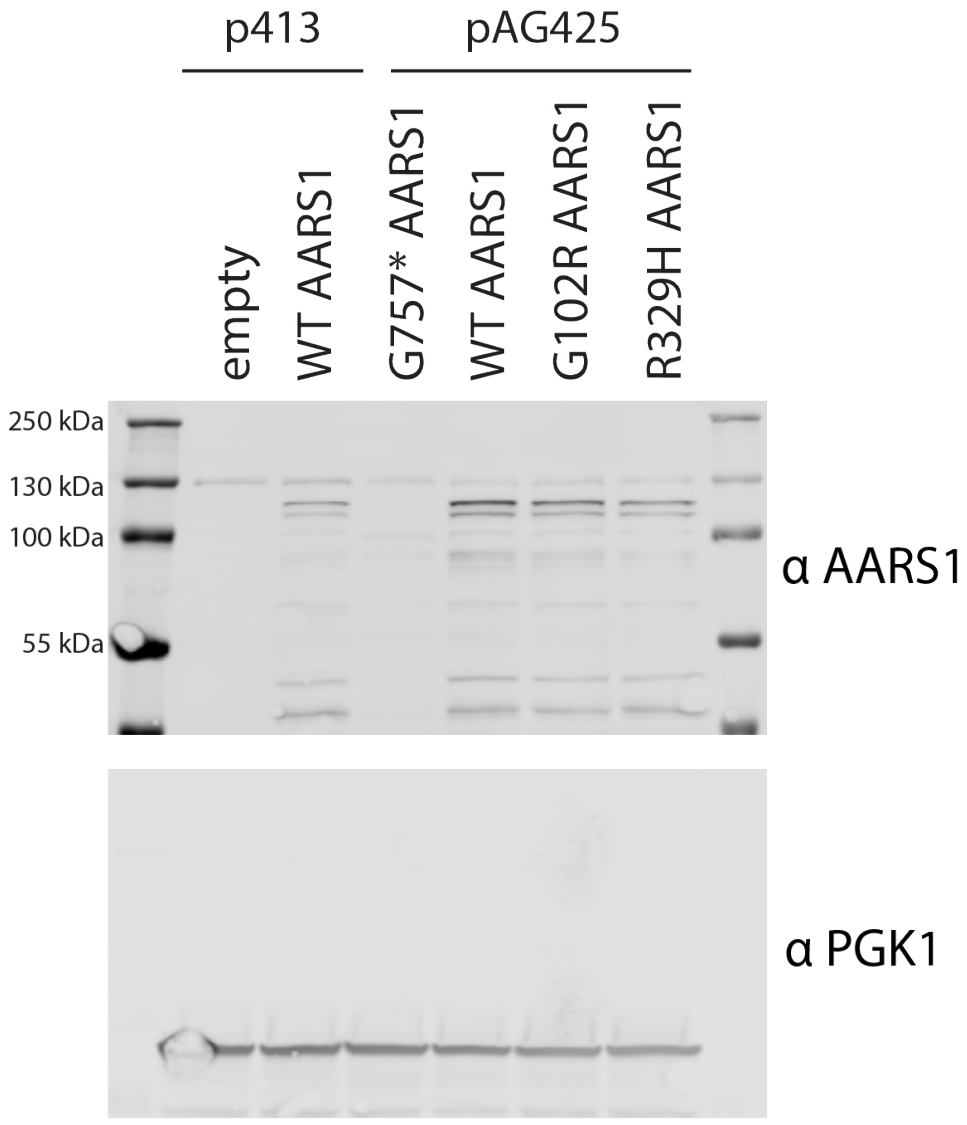


Figure B.1 Full image of western blot presented in Figure 4.1

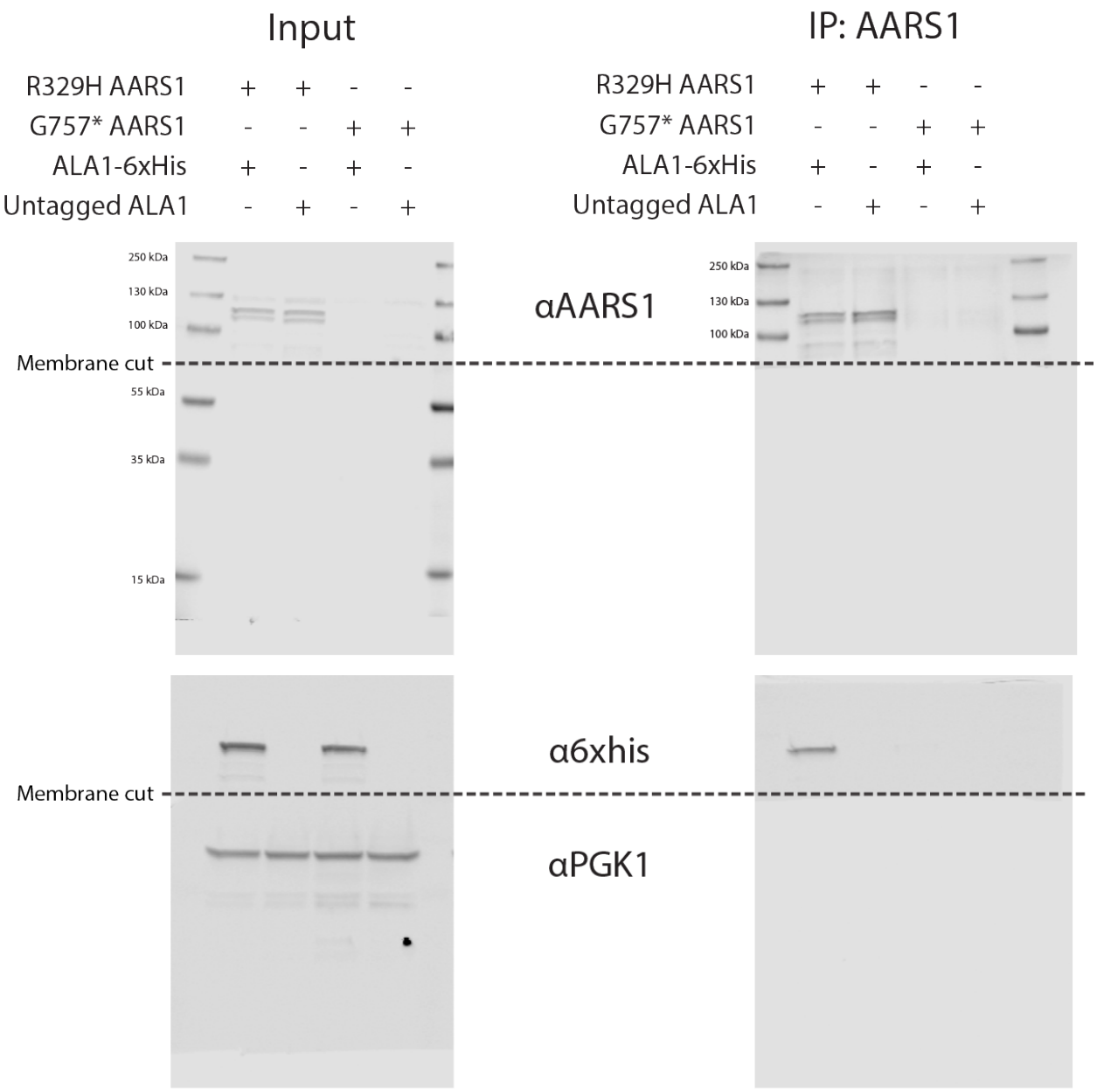
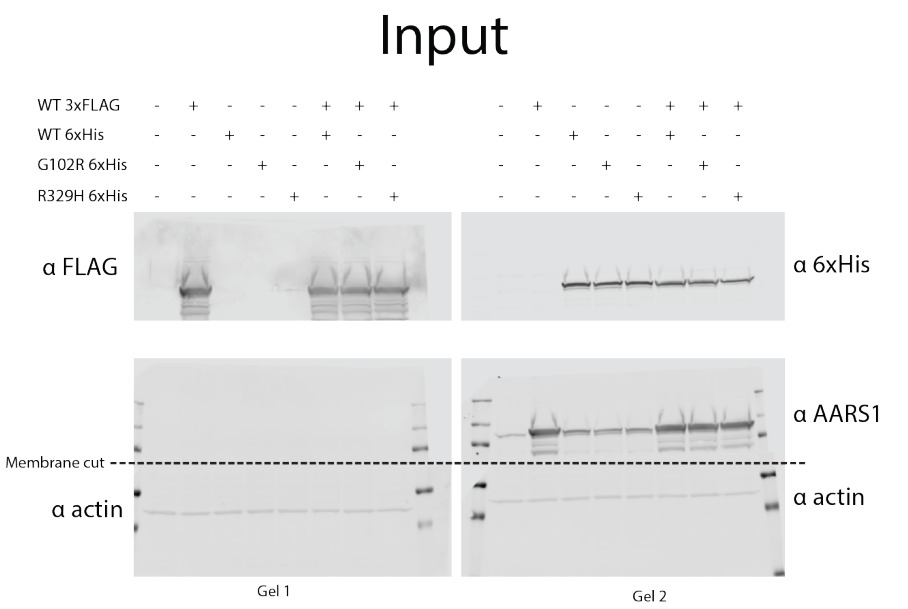
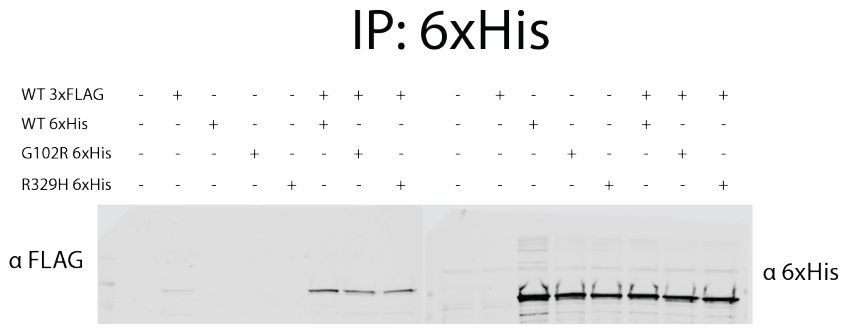


Figure B.2 Full image of western blot presented in Figure 4.3





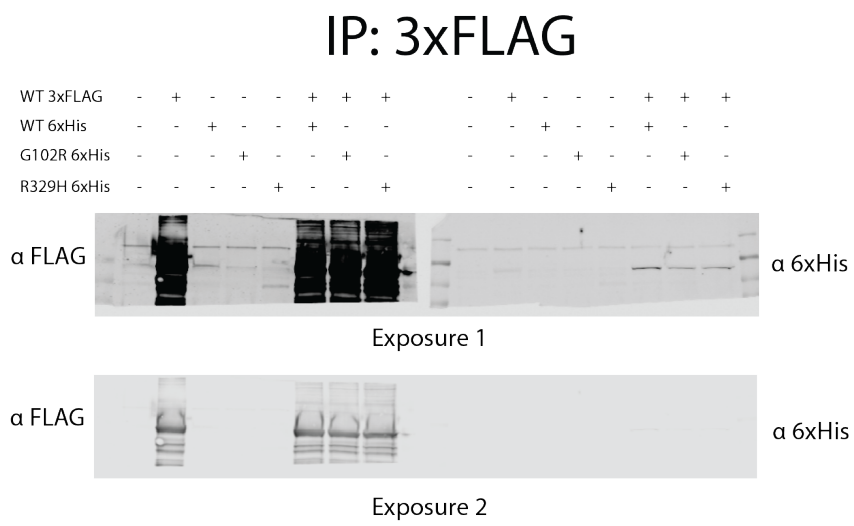


Figure B.3 Full image of western blot presented in Figure 4.4

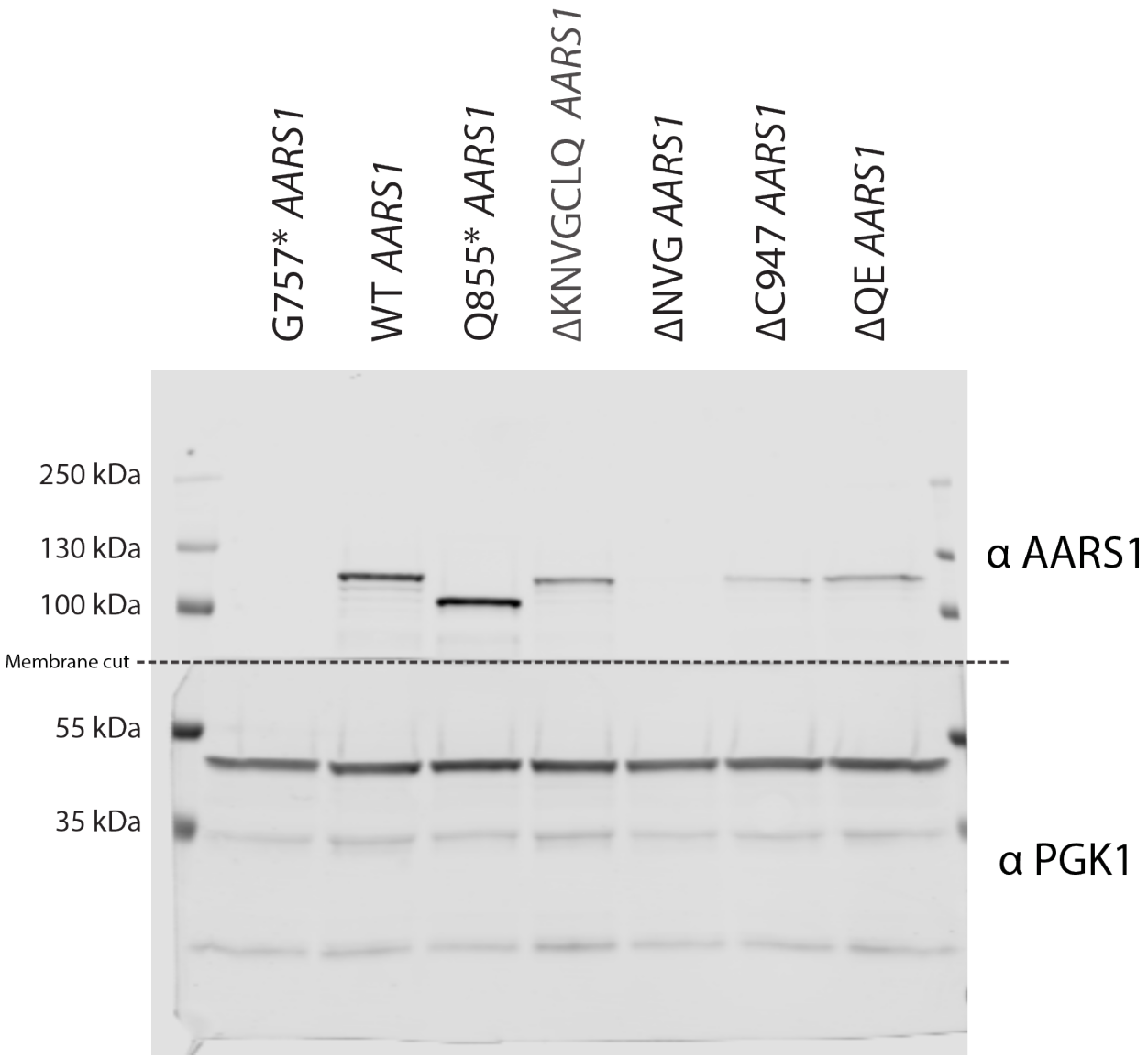
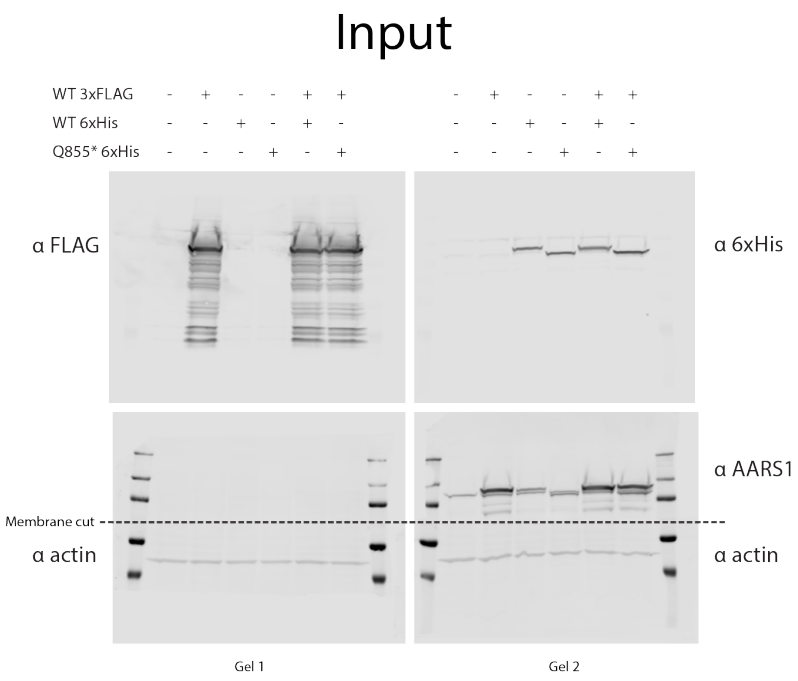
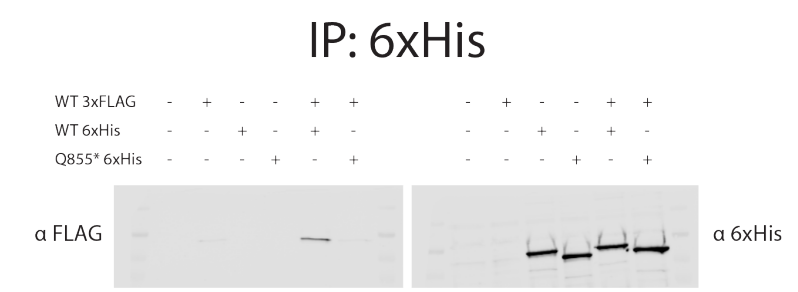


Figure B.4 Full image of western blot presented in Figure 4.6





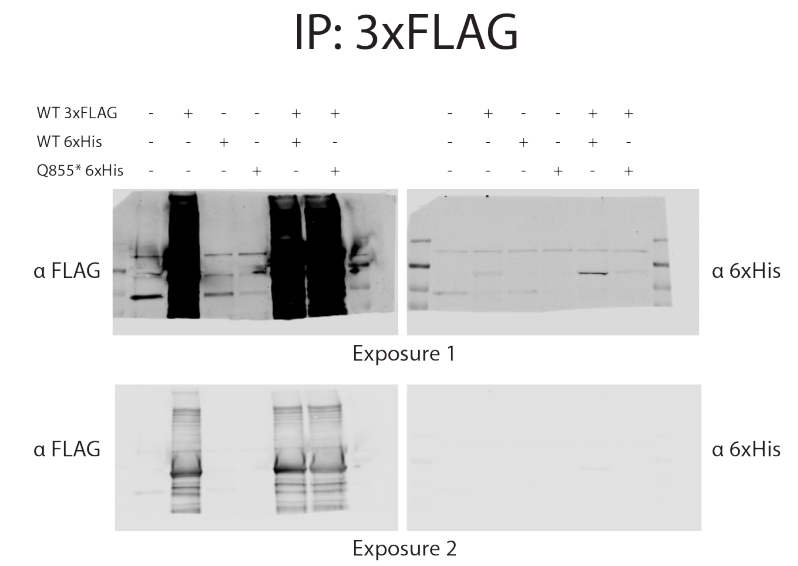


Figure B.5 Full image of western blot presented in Figure 4.7

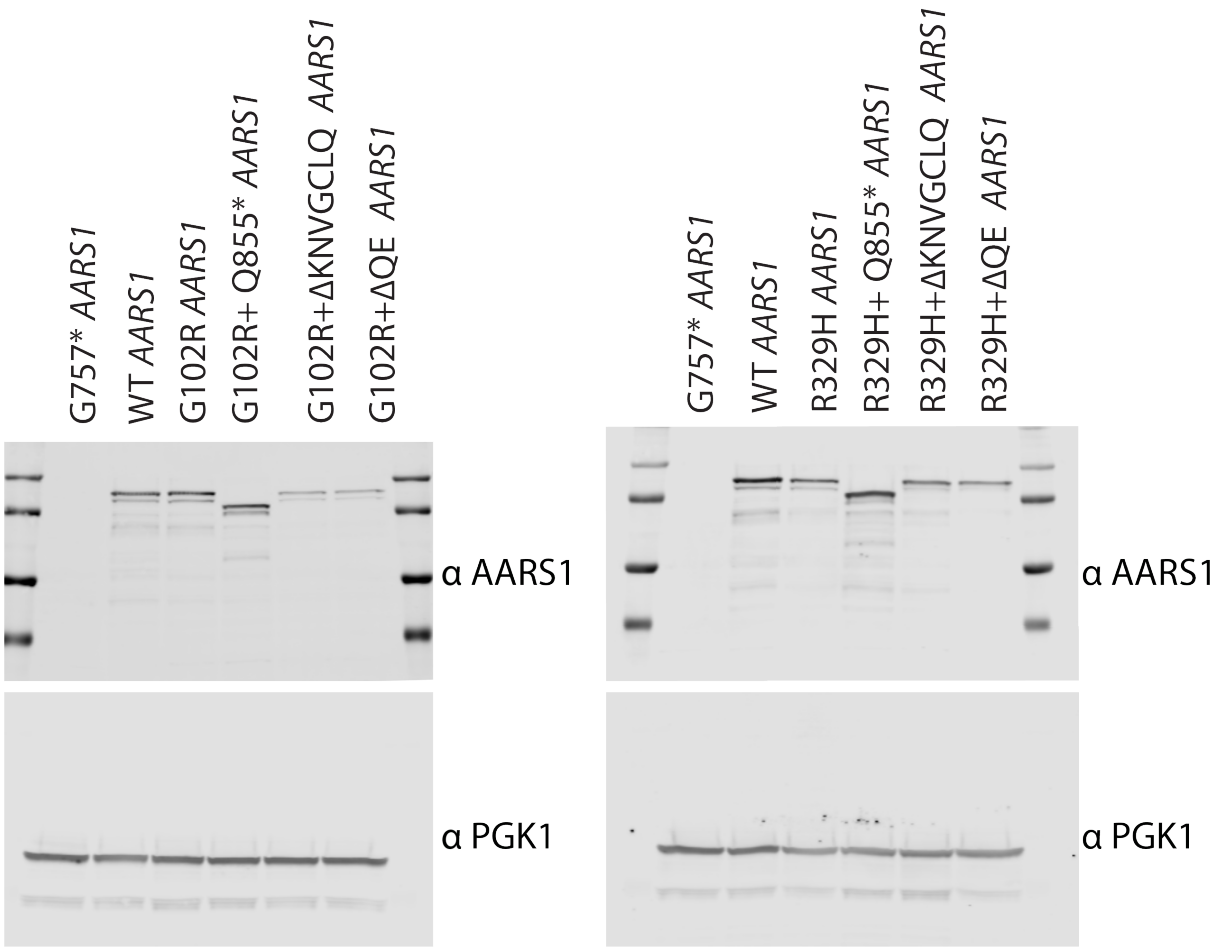


Figure B.6 Full image of western blot presented in Figure 4.8

References

1. Alberts, B. *et al.* From RNA to Protein. *Molecular Biology of the Cell. 4th edition* (2002).

2. Hinnebusch, A. G. The Scanning Mechanism of Eukaryotic Translation Initiation. *Annu. Rev. Biochem.* **83**, 779–812 (2014).

3. Wek, R. C. Role of eIF2α Kinases in Translational Control and Adaptation to Cellular Stress. *Cold Spring Harb Perspect Biol* **10**, a032870 (2018).

4. Hinnebusch, A. G. & Lorsch, J. R. The Mechanism of Eukaryotic Translation Initiation: New Insights and Challenges. *Cold Spring Harb Perspect Biol* **4**, a011544 (2012).

5. Dever, T. E. & Green, R. The elongation, termination, and recycling phases of translation in eukaryotes. *Cold Spring Harb Perspect Biol* **4**, a013706 (2012).

6. Proud, C. G. Phosphorylation and Signal Transduction Pathways in Translational Control. *Cold Spring Harb Perspect Biol* **11**, a033050 (2019).

7. Vattem, K. M. & Wek, R. C. Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. *Proc Natl Acad Sci U S A* **101**, 11269–11274 (2004).

8. Harding, H. P. *et al.* An Integrated Stress Response Regulates Amino Acid Metabolism and Resistance to Oxidative Stress. *Molecular Cell* **11**, 619–633 (2003).

9. Venturi, G. & Montanaro, L. How Altered Ribosome Production Can Cause or Contribute to Human Disease: The Spectrum of Ribosomopathies. *Cells* **9**, 2300 (2020).

10. McLachlan, F., Sires, A. M. & Abbott, C. M. The role of translation elongation factor eEF1 subunits in neurodevelopmental disorders. *Human Mutation* **40**, 131–141 (2019).

11. Bugiani, M., Boor, I., Powers, J. M., Scheper, G. C. & van der Knaap, M. S. Leukoencephalopathy with vanishing white matter: a review. *J Neuropathol Exp Neurol* **69**, 987–996 (2010).

12. Suzuki, T. The expanding world of tRNA modifications and their disease relevance. *Nat Rev Mol Cell Biol* **22**, 375–392 (2021).

13. Antonellis, A. & Green, E. D. The role of aminoacyl-tRNA synthetases in genetic diseases. *Annual review of genomics and human genetics* **9**, 87–107 (2008).

14. Tolkunova, E., Park, H., Xia, J., King, M. P. & Davidson, E. The human lysyl-tRNA synthetase gene encodes both the cytoplasmic and mitochondrial enzymes by means of an unusual alternative splicing of the primary transcript. *J Biol Chem* **275**, 35063–35069 (2000).

15. Alexandrova, J., Paulus, C., Rudinger-Thirion, J., Jossinet, F. & Frugier, M. Elaborate uORF/IRES features control expression and localization of human glycyl-tRNA synthetase. *RNA biology* **12**, 1301–1313 (2015).

16. Echevarría, L. *et al.* Glutamyl-tRNAGln amidotransferase is essential for mammalian mitochondrial translation in vivo. *Biochem J* **460**, 91–101 (2014).

17. Pouplana, L. R. de & Schimmel, P. Two Classes of tRNA Synthetases Suggested by Sterically Compatible Dockings on tRNA Acceptor Stem. *Cell* **104**, 191–193 (2001).

18. Fraser, T. H. & Rich, A. Amino acids are not all initially attached to the same position on transfer RNA molecules. *Proc Natl Acad Sci U S A* **72**, 3044–3048 (1975).

19. Yadavalli, S. S. & Ibba, M. Chapter 1 - Quality control in aminoacyl-tRNA synthesis: Its role in translational fidelity. in *Advances in Protein Chemistry and Structural Biology* (ed. Marintchev, A.) vol. 86 1–43 (Academic Press, 2012).

20. Hausmann, C. D. & Ibba, M. Aminoacyl-tRNA synthetase complexes: molecular multitasking revealed. *FEMS Microbiol Rev* **32**, 705–721 (2008).

21. Kyriacou, S. V. & Deutscher, M. P. An Important Role for the Multienzyme Aminoacyl-tRNA Synthetase Complex in Mammalian Translation and Cell Growth. *Mol Cell* **29**, 419–427 (2008).

22. Ray, P. S., Arif, A. & Fox, P. L. Macromolecular complexes as depots for releasable regulatory proteins. *Trends Biochem Sci* **32**, 158–164 (2007).

23. Han, J. M. *et al.* Leucyl-tRNA synthetase is an intracellular leucine sensor for the mTORC1-signaling pathway. *Cell* **149**, 410–424 (2012).

24. Kang, T. *et al.* AIMP3/p18 controls translational initiation by mediating the delivery of charged initiator tRNA to initiation complex. *J Mol Biol* **423**, 475–481 (2012).

25. Lo, W.-S. *et al.* Human tRNA synthetase catalytic nulls with diverse functions. *Science* **345**, 328–332 (2014).

26. Guo, M. & Schimmel, P. Essential nontranslational functions of tRNA synthetases. *Nature chemical biology* **9**, 145–153 (2013).

27. Witt, L. J., Curran, J. J. & Strek, M. E. The Diagnosis and Treatment of Antisynthetase Syndrome. *Clin Pulm Med* **23**, 218–226 (2016).

28. Rubio Gomez, M. A. & Ibba, M. Aminoacyl-tRNA synthetases. *RNA* **26**, 910–936 (2020).

29. Khan, K., Baleanu-Gogonea, C., Willard, B., Gogonea, V. & Fox, P. L. 3-Dimensional architecture of the human multi-tRNA synthetase complex. *Nucleic Acids Research* **48**, 8740–8754 (2020).

30. Sun, L., Song, Y., Blocquel, D., Yang, X.-L. & Schimmel, P. Two crystal structures reveal design for repurposing the C-Ala domain of human AlaRS. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 14300–14305 (2016).

31. Euro, L. *et al.* Structural modeling of tissue-specific mitochondrial alanyl-tRNA synthetase (AARS2) defects predicts differential effects on aminoacylation. *Front Genet* **6**, 21 (2015).

32. Pan, F. *et al.* Multiple molecular forms of cysteinyl-tRNA synthetase from rat liver: purification and subunit structure. *Biochim Biophys Acta* **452**, 271–283 (1976).

33. Motorin, Y. & Waller, J. P. Purification and properties of cysteinyl-tRNA synthetase from rabbit liver. *Biochimie* **80**, 579–590 (1998).

34. Newberry, K. J., Hou, Y.-M. & Perona, J. J. Structural origins of amino acid selection without editing by cysteinyl-tRNA synthetase. *EMBO J* **21**, 2778–2787 (2002).

35. Kim, K. R. *et al.* Crystal structure of human cytosolic aspartyl-tRNA synthetase, a component of multi-tRNA synthetase complex. *Proteins* **81**, 1840–1846 (2013).

36. Bonnefond, L. *et al.* Toward the full set of human mitochondrial aminoacyl-tRNA synthetases: characterization of AspRS and TyrRS. *Biochemistry* **44**, 4805–4816 (2005).

37. Kaiser, E. *et al.* The Human EPRS Locus (Formerly the QARS Locus): A Gene Encoding a Class I and a Class II Aminoacyl-tRNA Synthetase. *Genomics* **19**, 280–290 (1994).

38. Sissler, M., González-Serrano, L. E. & Westhof, E. Recent Advances in Mitochondrial Aminoacyl-tRNA Synthetases and Disease. *Trends in molecular medicine* **23**, 693–708 (2017).

39. Rodova, M., Ankilova, V. & Safro, M. G. Human phenylalanyl-tRNA synthetase: cloning, characterization of the deduced amino acid sequences in terms of the structural domains and coordinately regulated expression of the alpha and beta subunits in chronic myeloid leukemia cells. *Biochem Biophys Res Commun* **255**, 765–773 (1999).

40. Qin, X., Deng, X., Chen, L. & Xie, W. Crystal Structure of the Wild-Type Human GlyRS Bound with tRNA(Gly) in a Productive Conformation. *Journal of Molecular Biology* **428**, 3603–3614 (2016).

41. Koh, C. Y., Wetzel, A. B., de van der Schueren, W. J. & Hol, W. G. J. Comparison of histidine recognition in human and trypanosomatid histidyl-tRNA synthetases. *Biochimie* **106**, 111–120 (2014).

42. Pierce, S. B. *et al.* Mutations in mitochondrial histidyl tRNA synthetase HARS2 cause ovarian dysgenesis and sensorineural hearing loss of Perrault syndrome. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 6543–6548 (2011).

43. Nureki, O. *et al.* Enzyme structure with two catalytic sites for double-sieve selection of substrate. *Science* **280**, 578–582 (1998).

44. Guo, M., Ignatov, M., Musier-Forsyth, K., Schimmel, P. & Yang, X.-L. Crystal structure of tetrameric form of human lysyl-tRNA synthetase: Implications for multisynthetase complex formation. *PNAS* **105**, 2331–2336 (2008).

45. Hei, Z., Wu, S., Liu, Z., Wang, J. & Fang, P. Retractile lysyl-tRNA synthetase-AIMP2 assembly in the human multi-aminoacyl-tRNA synthetase complex. *J Biol Chem* **294**, 4775–4783 (2019).

46. Cusack, S., Yaremchuk, A. & Tukalo, M. The 2 Å crystal structure of leucyl-tRNA synthetase and its complex with a leucyl-adenylate analogue. *EMBO J* **19**, 2351–2361 (2000).

47. Kellermann, O., Viel, C. & Waller, J. P. Methionyl-tRNA synthetase from sheep mammary gland. Purification of a fully active monomeric enzyme derived from high-molecular-weight complexes by controlled proteolysis. *Eur J Biochem* **88**, 197–204 (1978).

48. Mechulam, Y. *et al.* Crystal structure of Escherichia coli methionyl-tRNA synthetase highlights species-specific features. *J Mol Biol* **294**, 1287–1297 (1999).

49. Park, J. S. *et al.* Unique N-terminal extension domain of human asparaginyl-tRNA synthetase elicits CCR3-mediated chemokine activity. *Int J Biol Macromol* **120**, 835–845 (2018).

50. Simon, M. *et al.* Mutations of human NARS2, encoding the mitochondrial asparaginyl-tRNA synthetase, cause nonsyndromic deafness and Leigh syndrome. *PLoS genetics* **11**, e1005097 (2015).

51. Ciara, E. *et al.* Clinical and molecular characteristics of newly reported mitochondrial disease entity caused by biallelic PARS2 mutations. *J Hum Genet* **63**, 473–485 (2018).

52. Ognjenović, J. *et al.* The crystal structure of human GlnRS provides basis for the development of neurological disorders. *Nucleic Acids Res* **44**, 3420–3431 (2016).

53. Kim, H. S., Cha, S. Y., Jo, C. H., Han, A. & Hwang, K. Y. The crystal structure of arginyl-tRNA synthetase from Homo sapiens. *FEBS Letters* **588**, 2328–2334 (2014).

54. Xu, X. *et al.* Unique domain appended to vertebrate tRNA synthetase is essential for vascular development. *Nat Commun* **3**, 681 (2012).

55. Chimnaronk, S., Gravers Jeppesen, M., Suzuki, T., Nyborg, J. & Watanabe, K. Dual-mode recognition of noncanonical tRNAsSer by seryl-tRNA synthetase in mammalian mitochondria. *EMBO J* **24**, 3369–3379 (2005).

56. Fang, P. *et al.* Structural basis for full-spectrum inhibition of translational functions on a tRNA synthetase. *Nature Communications* **6**, 6402 (2015).

57. Wang, Y. *et al.* A Human Disease-causing Point Mutation in Mitochondrial Threonyl-tRNA Synthetase Induces Both Structural and Functional Defects. *The Journal of biological chemistry* **291**, 6507–6520 (2016).

58. Negrutskii, B. S., Shalak, V. F., Kerjan, P., El'skaya, A. V. & Mirande, M. Functional interaction of mammalian valyl-tRNA synthetase with elongation factor EF-1alpha in the complex with EF-1H. *J Biol Chem* **274**, 4545–4550 (1999).

59. Kohda, D., Yokoyama, S. & Miyazawa, T. Thermostable valyl-tRNA, isoleucyl-tRNA and methionyl-tRNA synthetases from an extreme thermophile Thermus thermophilus HB8: protein structure and Zn2+ binding. *FEBS Lett* **174**, 20–23 (1984).

60. Shen, N., Guo, L., Yang, B., Jin, Y. & Ding, J. Structure of human tryptophanyl-tRNA synthetase in complex with tRNATrp reveals the molecular basis of tRNA recognition and specificity. *Nucleic Acids Res* **34**, 3246–3258 (2006).

61. Doublié, S., Bricogne, G., Gilmore, C. & Carter Jr, C. W. Tryptophanyl-tRNA synthetase crystal structure reveals an unexpected homology to tyrosyl-tRNA synthetase. *Structure* **3**, 17–31 (1995).

62. Austin, J. & First, E. A. Catalysis of tyrosyl-adenylate formation by the human tyrosyl-tRNA synthetase. *J Biol Chem* **277**, 14812–14820 (2002).

63. Greenberg, Y. *et al.* The novel fragment of tyrosyl tRNA synthetase, mini-TyrRS, is secreted to induce an angiogenic response in endothelial cells. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **22**, 1597–1605 (2008).

64. Ahn, Y. H. *et al.* Secreted tryptophanyl-tRNA synthetase as a primary defence system against infection. *Nature microbiology* **2**, 16191 (2016).

65. Wakasugi, K. & Schimmel, P. Highly Differentiated Motifs Responsible for Two Cytokine Activities of a Split Human tRNA Synthetase*. *Journal of Biological Chemistry* **274**, 23155–23159 (1999).

66. Jobin, P. G. *et al.* Moonlighting matrix metalloproteinase substrates: Enhancement of proinflammatory functions of extracellular tyrosyl-tRNA synthetase upon cleavage. *J Biol Chem* **295**, 2186–2202 (2020).

67. Mukhopadhyay, R., Jia, J., Arif, A., Ray, P. S. & Fox, P. L. The GAIT system: a gatekeeper of inflammatory gene expression. *Trends Biochem Sci* **34**, 324–331 (2009).

68. Meyer-Schuman, R. & Antonellis, A. Emerging mechanisms of aminoacyl-tRNA synthetase mutations in recessive and dominant human disease. *Human molecular genetics* **26**, R114–R127 (2017).

69. Seburn, K. L., Nangle, L. A., Cox, G. A., Schimmel, P. & Burgess, R. W. An active dominant mutation of glycyl-tRNA synthetase causes neuropathy in a Charcot-Marie-Tooth 2D mouse model. *Neuron* **51**, 715–726 (2006).

70. Dallabona, C. *et al.* Novel (ovario) leukodystrophy related to AARS2 mutations. *Neurology* **82**, 2063–2071 (2014).

71. Coughlin, C. R. *et al.* Mutations in the mitochondrial cysteinyl-tRNA synthase gene, CARS2, lead to a severe epileptic encephalopathy and complex movement disorder. *Journal of medical genetics* **52**, 532–540 (2015).

72. Sun, C. *et al.* Loss-of-function mutations in Lysyl-tRNA synthetase cause various leukoencephalopathy phenotypes. *Neurology Genetics* **5**, (2019).

73. Theisen, B. E. *et al.* Deficiency of WARS2, encoding mitochondrial tryptophanyl tRNA synthetase, causes severe infantile onset leukoencephalopathy. *Am J Med Genet A* **173**, 2505–2510 (2017).

74. Raviglione, F. *et al.* Clinical findings in a patient with FARS2 mutations and early-infantile-encephalopathy with epilepsy. *American Journal of Medical Genetics Part A* **170**, 3004–3007 (2016).

75. Scheper, G. C. *et al.* Mitochondrial aspartyl-tRNA synthetase deficiency causes leukoencephalopathy with brain stem and spinal cord involvement and lactate elevation. *Nat Genet* **39**, 534–539 (2007).

76. Steenweg, M. E. *et al.* Leukoencephalopathy with thalamus and brainstem involvement and high lactate 'LTBL' caused by EARS2 mutations. *Brain : a journal of neurology* **135**, 1387–1394 (2012).

77. Sommerville, E. W. *et al.* Clinical Features, Molecular Heterogeneity, and Prognostic Implications in YARS2-Related Mitochondrial Myopathy. *JAMA neurology* (2017).

78. Sommerville, E. W. *et al.* Instability of the mitochondrial alanyl-tRNA synthetase underlies fatal infantile-onset cardiomyopathy. *Hum Mol Genet* **28**, 258–268 (2019).

79. Riley, L. G. *et al.* Mutation of the mitochondrial tyrosyl-tRNA synthetase gene, YARS2, causes myopathy, lactic acidosis, and sideroblastic anemia–MLASA syndrome. *American journal of human genetics* **87**, 52–59 (2010).

80. Baertling, F. *et al.* Neonatal encephalocardiomyopathy caused by mutations in VARS2. *Metabolic Brain Disease* 1–4 (2017).

81. Verrigni, D. *et al.* Novel mutations in KARS cause hypertrophic cardiomyopathy and combined mitochondrial respiratory chain defect. *Clin Genet* **91**, 918–923 (2017).

82. Elo, J. M. *et al.* Mitochondrial phenylalanyl-tRNA synthetase mutations underlie fatal infantile Alpers encephalopathy. *Human molecular genetics* **21**, 4521–4529 (2012).

83. Walker, M. A. *et al.* Novel compound heterozygous mutations expand the recognized phenotypes of FARS2-linked disease. *J Child Neurol* **31**, 1127–1137 (2016).

84. Danhauser, K. *et al.* EARS2 mutations cause fatal neonatal lactic acidosis, recurrent hypoglycemia and agenesis of corpus callosum. *Metabolic Brain Disease* **31**, 717–721 (2016).

85. Riley, L. G. *et al.* LARS2 Variants Associated with Hydrops, Lactic Acidosis, Sideroblastic Anemia, and Multisystem Failure. *JIMD Reports* **28**, 49–57 (2016).

86. Lühl, S. *et al.* Novel homozygous RARS2 mutation in two siblings without pontocerebellar hypoplasia - further expansion of the phenotypic spectrum. *Orphanet journal of rare diseases* **11**, 140 (2016).

87. Wortmann, S. B. *et al.* Biallelic variants in WARS2 encoding mitochondrial tryptophanyl-tRNA synthase in six individuals with mitochondrial encephalopathy. *Hum Mutat* **38**, 1786–1795 (2017).

88. Ardissone, A. *et al.* A Novel Homozygous YARS2 Mutation in Two Italian Siblings and a Review of Literature. *JIMD Reports* **20**, 95–101 (2015).

89. Mizuguchi, T. *et al.* PARS2 and NARS2 mutations in infantile-onset neurodegenerative disorder. *Journal of Human Genetics* 1–5 (2017).

90. Hallmann, K. *et al.* A homozygous splice-site mutation in CARS2 is associated with progressive myoclonic epilepsy. *Neurology* **83**, 2183–2187 (2014).

91. Webb, B. D. *et al.* Novel, compound heterozygous, single-nucleotide variants in MARS2 associated with developmental delay, poor growth, and sensorineural hearing loss. *Human mutation* **36**, 587–592 (2015).

92. Vanlander, A. V. *et al.* Two Siblings with Homozygous Pathogenic Splice-Site Variant in Mitochondrial Asparaginyl–tRNA Synthetase (NARS2). *Human mutation* **36**, 222–231 (2015).

93. Alkhateeb, A. M., Aburahma, S. K., Habbab, W. & Thompson, I. R. Novel mutations in WWOX, RARS2, and C10orf2 genes in consanguineous Arab families with intellectual disability. *Metabolic Brain Disease* **31**, 901–907 (2016).

94. Lee, J.-M. *et al.* Two Korean siblings with recently described ovarioleukodystrophy related to AARS2 mutations. *European journal of neurology* **24**, e21–e22 (2017).

95. Hamatani, M. *et al.* The first Japanese case of leukodystrophy with ovarian failure arising from novel compound heterozygous AARS2 mutations. *Journal of Human Genetics* **61**, 899–902 (2016).

96. Pierce, S. B. *et al.* Mutations in LARS2, encoding mitochondrial leucyl-tRNA synthetase, lead to premature ovarian failure and hearing loss in Perrault syndrome. *American journal of human genetics* **92**, 614–620 (2013).

97. Moosa, S. *et al.* Confirmation of CAGSSS syndrome as a distinct entity in a Danish patient with a novel homozygous mutation in IARS2. *American Journal of Medical Genetics Part A* **173**, 1102–1108 (2017).

98. 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. *Human mutation* **36**, 281–281 (2015).

99. Takezawa, Y. *et al.* Novel IARS2 mutations in Japanese siblings with CAGSSS, Leigh, and West syndrome. *Brain Dev* **40**, 934–938 (2018).

100. Santos-Cortez, R. L. P. *et al.* Mutations in KARS, Encoding Lysyl-tRNA Synthetase, Cause Autosomal-Recessive Nonsyndromic Hearing Impairment DFNB89. *The American Journal of Human Genetics* **93**, 132–140 (2013).

101. Riley, L. G. *et al.* The expanding LARS2 phenotypic spectrum: HLASA, Perrault syndrome with leukodystrophy, and mitochondrial myopathy. *Hum Mutat* **41**, 1425–1434 (2020).

102. Simons, C. *et al.* Loss-of-function alanyl-tRNA synthetase mutations cause an autosomal-recessive early-onset epileptic encephalopathy with persistent myelination defect. *American journal of human genetics* **96**, 675–681 (2015).

103. Nakayama, T. *et al.* Deficient activity of alanyl-tRNA synthetase underlies an autosomal recessive syndrome of progressive microcephaly, hypomyelination, and epileptic encephalopathy. *Human mutation* (2017).

104. Marten, L. M. *et al.* Recurrent acute liver failure in alanyl-tRNA synthetase-1 (AARS1) deficiency. *Mol Genet Metab Rep* **25**, (2020).

105. Botta, E. *et al.* Protein instability associated with AARS1 and MARS1 mutations causes Trichothiodystrophy. *Hum Mol Genet* ddab123 (2021) doi:10.1093/hmg/ddab123.

106. Latour, P. *et al.* A Major Determinant for Binding and Aminoacylation of tRNAAla in Cytoplasmic Alanyl-tRNA Synthetase Is Mutated in Dominant Axonal Charcot-Marie-Tooth Disease. *The American Journal of Human Genetics* **86**, 77–82 (2010).

107. 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). *Human mutation* **33**, 244–253 (2012).

108. Lin, K.-P. *et al.* The mutational spectrum in a cohort of Charcot-Marie-Tooth disease type 2 among the Han Chinese in Taiwan. *PLoS ONE* **6**, e29393 (2011).

109. Motley, W. W. *et al.* A novel AARS mutation in a family with dominant myeloneuropathy. *Neurology* **84**, 2040–2047 (2015).

110. Weterman, M. A. J. *et al.* Hypermorphic and hypomorphic AARS alleles in patients with CMT2N expand clinical and molecular heterogeneities. *Hum. Mol. Genet.* **27**, 4036–4050 (2018).

111. Bansagi, B. *et al.* Genotype/phenotype correlations in AARS-related neuropathy in a cohort of patients from the United Kingdom and Ireland. *J Neurol* **262**, 1899–1908 (2015).

112. Zhao, Z. *et al.* Alanyl-tRNA synthetase mutation in a family with dominant distal hereditary motor neuropathy. *Neurology* **78**, 1644–1649 (2012).

113. Lynch, D. S. *et al.* Analysis of Mutations in AARS2 in a Series of CSF1R-Negative Patients With Adult-Onset Leukoencephalopathy With Axonal Spheroids and Pigmented Glia. *JAMA neurology* **73**, 1433–1439 (2016).

114. Szpisjak, L. *et al.* Novel AARS2 gene mutation producing leukodystrophy: a case report. *Journal of Human Genetics* **62**, 329–333 (2017).

115. Wang, D., Yu, M., Zhang, W., Wang, Z. & Yuan, Y. AARS2 Compound Heterozygous Variants in a Case of Adult-Onset Leukoencephalopathy With Axonal Spheroids and Pigmented Glia. *J. Neuropathol. Exp. Neurol.* **77**, 997–1000 (2018).

116. Tang, Y. *et al.* AARS2 leukoencephalopathy: A new variant of mitochondrial encephalomyopathy. *Molecular Genetics & Genomic Medicine* **7**, e00582 (2019).

117. Uzun, G. Adult-onset leukodystrophy with homozygous AARS2 mutation located in the aminoacylation domain. *Neurology India; Mumbai* **67**, 871–872 (2019).

118. Felhi, R. *et al.* Mutations in aARS genes revealed by targeted next-generation sequencing in patients with mitochondrial diseases. *Mol Biol Rep* **47**, 3779–3787 (2020).

119. Dong, Q. *et al.* An adolescence-onset male leukoencephalopathy with remarkable cerebellar atrophy and novel compound heterozygous AARS2 gene mutations: a case report. *J. Hum. Genet.* **63**, 841–846 (2018).

120. Sun, J., Quan, C., Luo, S.-S., Zhou, L. & Zhao, C.-B. Leukodystrophy without Ovarian Failure Caused by Compound Heterozygous Alanyl-tRNA Synthetase 2 Mutations. *Chin Med J (Engl)* **130**, 3021–3022 (2017).

121. Taylor, R. W. *et al.* Use of whole-exome sequencing to determine the genetic basis of multiple mitochondrial respiratory chain complex deficiencies. *JAMA* **312**, 68–77 (2014).

122. Peragallo, J. H., Keller, S., van der Knaap, M. S., Soares, B. P. & Shankar, S. P. Retinopathy and optic atrophy: expanding the phenotypic spectrum of pathogenic variants in AARS2 gene. *Ophthalmic Genet* **39**, 99–102 (2018).

123. Kuo, M. E., Antonellis, A. & Shakkottai, V. G. Alanyl-tRNA Synthetase 2 (AARS2)-Related Ataxia Without Leukoencephalopathy. *Cerebellum* **19**, 154–160 (2020).

124. Zhou, Y. *et al.* Novel alanyl-tRNA synthetase 2 (AARS2) homozygous mutation in a consanguineous Chinese family with premature ovarian insufficiency. *Fertility and Sterility* **112**, 569-576.e2 (2019).

125. Kiraly-Borri, C. *et al.* Siblings with lethal primary pulmonary hypoplasia and compound heterozygous variants in the AARS2 gene: further delineation of the phenotypic spectrum. *Cold Spring Harb Mol Case Stud* **5**, (2019).

126. 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. *Am. J. Hum. Genet.* **104**, 520–529 (2019).

127. Samanta, D., Gokden, M. & Willis, E. Clinicopathologic Findings of CARS2 Mutation. *Pediatric Neurology* **87**, 65–69 (2018).

128. Taft, R. J. *et al.* Mutations in DARS Cause Hypomyelination with Brain Stem and Spinal Cord Involvement and Leg Spasticity. *The American Journal of Human Genetics* **92**, 774–780 (2013).

129. Ong, M. *et al.* Genotype-phenotype variability of DARS mutation - case reports of a trio of siblings. *EJMCR* 110–115 (2020) doi:10.24911/ejmcr/1731551044010.

130. Wolf, N. I. *et al.* DARS-associated leukoencephalopathy can mimic a steroid-responsive neuroinflammatory disorder. *Neurology* **84**, 226–230 (2015).

131. Köhler, C. *et al.* Early-onset leukoencephalopathy due to a homozygous missense mutation in the DARS2 gene. *Molecular and Cellular Probes* **29**, 319–322 (2015).

132. Shimojima, K. *et al.* A novel DARS2 mutation in a Japanese patient with leukoencephalopathy with brainstem and spinal cord involvement but no lactate elevation. *Hum Genome Var* **4**, 17051 (2017).

133. Moore, S. A., Kumar, N. & Gavrilova, R. H. Leukoencephalopathy with brain stem and spinal cord involvement (and high lactate): raising the bar for diagnosis. *J Neurol* **259**, 2494–2497 (2012).

134. Lin, J. *et al.* Leukoencephalopathy With Brainstem and Spinal Cord Involvement and Normal Lactate: A New Mutation in the DARS2 Gene. *Journal of Child Neurology* **25**, 1425–1428 (2010).

135. Labauge, P., Dorboz, I., Eymard-Pierre, E., Dereeper, O. & Boespflug-Tanguy, O. Clinically asymptomatic adult patient with extensive LBSL MRI pattern and DARS2 mutations. *J Neurol* **258**, 335–337 (2011).

136. Cheng, F.-B. *et al.* Adult-onset leukoencephalopathy with brain stem and spinal cord involvement in Chinese Han population: a case report and literature review. *Neurol India* **61**, 161–163 (2013).

137. Uluc, K. *et al.* Leukoencephalopathy with brain stem and spinal cord involvement and high lactate: A genetically proven case with distinct MRI findings. *Journal of the Neurological Sciences* **273**, 118–122 (2008).

138. Mendes, M. I. *et al.* Bi-allelic Mutations in EPRS, Encoding the Glutamyl-Prolyl-Aminoacyl-tRNA Synthetase, Cause a Hypomyelinating Leukodystrophy. *Am J Hum Genet* **102**, 676–684 (2018).

139. Taskin, B. D. *et al.* Early-Onset Mild Type Leukoencephalopathy Caused by a Homozygous EARS2 Mutation. *Journal of Child Neurology* **31**, 938–941 (2016).

140. Biancheri, R. *et al.* Expanding the Clinical and Magnetic Resonance Spectrum of Leukoencephalopathy with Thalamus and Brainstem Involvement and High Lactate (LTBL) in a Patient Harboring a Novel EARS2 Mutation. in *JIMD Reports, Volume 23* 85–89 (Springer Berlin Heidelberg, 2015).

141. Oliveira, R. *et al.* Lethal Neonatal LTBL Associated with Biallelic EARS2 Variants: Case Report and Review of the Reported Neuroradiological Features. *JIMD Reports* **33**, 61–68 (2017).

142. Prasun, P., Mintz, C., Cork, E., Naidich, T. P. & Webb, B. D. Broad spectrum of clinical presentation in EARS2 beyond typical "leukoencephalopathy with thalamus and brain stem involvement". *Journal of the Neurological Sciences* **406**, 116448 (2019).

143. Monfrini, E. *et al.* Late-onset leukoencephalopathy in a patient with recessive EARS2 mutations. *Neurol Genet* **6**, e488 (2020).

144. Krenke, K. *et al.* FARSA mutations mimic phenylalanyl-tRNA synthetase deficiency caused by FARSB defects. *Clinical Genetics* **96**, 468–472 (2019).

145. Antonellis, A. *et al.* Compound heterozygosity for loss-of-function FARSB variants in a patient with classic features of recessive aminoacyl-tRNA synthetase-related disease. *Hum Mutat* **39**, 834–840 (2018).

146. Zadjali, F. *et al.* Homozygosity for FARSB mutation leads to Phe-tRNA synthetase-related disease of growth restriction, brain calcification, and interstitial lung disease. *Human Mutation* **39**, 1355–1359 (2018).

147. Xu, Z. *et al.* Bi-allelic Mutations in Phe-tRNA Synthetase Associated with a Multi-system Pulmonary Disease Support Non-translational Function. *Am J Hum Genet* **103**, 100–114 (2018).

148. Vernon, H. J., McClellan, R., Batista, D. A. S. & Naidu, S. Mutations in FARS2 and non-fatal mitochondrial dysfunction in two siblings. *American Journal of Medical Genetics Part A* **167A**, 1147–1151 (2015).

149. Cho, J. S. *et al.* FARS2 mutation and epilepsy: Possible link with early-onset epileptic encephalopathy. *Epilepsy Research* **129**, 118–124 (2017).

150. Almannai, M. *et al.* FARS2 deficiency; new cases, review of clinical, biochemical, and molecular spectra, and variants interpretation based on structural, functional, and evolutionary significance. *Mol Genet Metab* **125**, 281–291 (2018).

151. Hotait, M. *et al.* FARS2 Mutations: More Than Two Phenotypes? A Case Report. *Front Genet* **11**, 787 (2020).

152. Chen, Z. & Zhang, Y. A patient with juvenile-onset refractory status epilepticus caused by two novel compound heterozygous mutations in FARS2 gene. *Int J Neurosci* **129**, 1094–1097 (2019).

153. Sahai, S. K. *et al.* FARS2 mutations presenting with pure spastic paraplegia and lesions of the dentate nuclei. *Ann Clin Transl Neurol* **5**, 1128–1133 (2018).

154. Yang, Y. *et al.* A Newly Identified Missense Mutation in FARS2 Causes Autosomal-Recessive Spastic Paraplegia. *Human mutation* **37**, 165–169 (2016).

155. Forman, E. B., Gorman, K. M., Ennis, S. & King, M. D. FARS2 Causing Complex Hereditary Spastic Paraplegia With Dysphonia: Expanding the Disease Spectrum. *J Child Neurol* **34**, 621 (2019).

156. Barcia, G. *et al.* Novel FARS2 variants in patients with early onset encephalopathy with or without epilepsy associated with long survival. *Eur J Hum Genet* **29**, 533–538 (2021).

157. McMillan, H. J. *et al.* Compound heterozygous mutations in glycyl-tRNA synthetase are a proposed cause of systemic mitochondrial disease. *BMC medical genetics* **15**, 36 (2014).

158. Nafisinia, M. *et al.* Compound heterozygous mutations in glycyl-tRNA synthetase (GARS) cause mitochondrial respiratory chain dysfunction. *PLoS One* **12**, (2017).

159. Oprescu, S. N. *et al.* Compound heterozygosity for loss-of-function GARS variants results in a multisystem developmental syndrome that includes severe growth retardation. *Human mutation* **38**, 1412–1420 (2017).

160. Antonellis, A. *et al.* Glycyl tRNA synthetase mutations in Charcot-Marie-Tooth disease type 2D and distal spinal muscular atrophy type V. *The American Journal of Human Genetics* **72**, 1293–1299 (2003).

161. Rohkamm, B. *et al.* Further evidence for genetic heterogeneity of distal HMN type V, CMT2 with predominant hand involvement and Silver syndrome. *Journal of the neurological sciences* **263**, 100–106 (2007).

162. James, P. A. *et al.* Severe childhood SMA and axonal CMT due to anticodon binding domain mutations in the GARS gene. *Neurology* **67**, 1710–1712 (2006).

163. Sivakumar, K. *et al.* Phenotypic spectrum of disorders associated with glycyl-tRNA synthetase mutations. *Brain : a journal of neurology* **128**, 2304–2314 (2005).

164. Del Bo, R. *et al.* Coexistence of CMT-2D and distal SMA-V phenotypes in an Italian family with a GARS gene mutation. *Neurology* **66**, 752–754 (2006).

165. Eskuri, J. M., Stanley, C. M., Moore, S. A. & Mathews, K. D. Infantile onset CMT2D/dSMA V in monozygotic twins due to a mutation in the anticodon-binding domain of GARS. *Journal of the peripheral nervous system : JPNS* **17**, 132–134 (2012).

166. Chung, P. *et al.* Glycyl tRNA Synthetase (GARS) Gene Variant Causes Distal Hereditary Motor Neuropathy V. *Case Rep Pediatr* **2018**, 8516285 (2018).

167. Yu, X. *et al.* A Novel Mutation of GARS in a Chinese Family With Distal Hereditary Motor Neuropathy Type V. *Front Neurol* **9**, 571 (2018).

168. Nan, H. *et al.* Novel GARS mutation presenting as autosomal dominant intermediate Charcot-Marie-Tooth disease. *J Peripher Nerv Syst* **24**, 156–160 (2019).

169. Yalcouyé, A. *et al.* A novel mutation in the GARS gene in a Malian family with Charcot-Marie-Tooth disease. *Mol Genet Genomic Med* **7**, e00782 (2019).

170. Morelli, K. H. *et al.* Allele-specific RNA interference prevents neuropathy in Charcot-Marie-Tooth disease type 2D mouse models. *J. Clin. Invest.* **129**, 5568–5583 (2019).

171. Markovitz, R. *et al.* GARS-related disease in infantile spinal muscular atrophy: Implications for diagnosis and treatment. *American Journal of Medical Genetics Part A* **182**, 1167–1176 (2020).

172. Galatolo, D. *et al.* Bi-allelic mutations in HARS1 severely impair histidyl-tRNA synthetase expression and enzymatic activity causing a novel multisystem ataxic syndrome. *Hum Mutat* **41**, 1232–1237 (2020).

173. Puffenberger, E. G. *et al.* Genetic mapping and exome sequencing identify variants associated with five novel diseases. *PLoS ONE* **7**, e28936 (2012).

174. Safka Brozkova, D. *et al.* Loss of function mutations in HARS cause a spectrum of inherited peripheral neuropathies. *Brain : a journal of neurology* **138**, 2161–2172 (2015).

175. Vester, A. *et al.* A Loss-of-Function Variant in the Human Histidyl-tRNA Synthetase (HARS) Gene is Neurotoxic In Vivo. *Human mutation* **34**, 191–199 (2013).

176. Abbott, J. A. *et al.* Substrate interaction defects in histidyl-tRNA synthetase linked to dominant axonal peripheral neuropathy. *Hum. Mutat.* **39**, 415–432 (2018).

177. Yu, J., Jiang, W., Cao, L., Na, X. & Yang, J. Two novel likely pathogenic variants of HARS2 identified in a Chinese family with sensorineural hearing loss. *Hereditas* **157**, 47 (2020).

178. Karstensen, H. G. *et al.* Novel HARS2 missense variants identified in individuals with sensorineural hearing impairment and Perrault syndrome. *Eur J Med Genet* **63**, 103733 (2020).

179. Zou, S. *et al.* Whole-exome sequencing identifies rare pathogenic and candidate variants in sporadic Chinese Han deaf patients. *Clin Genet* **97**, 352–356 (2020).

180. Demain, L. A. M. *et al.* A recurrent missense variant in HARS2 results in variable sensorineural hearing loss in three unrelated families. *J Hum Genet* **65**, 305–311 (2020).

181. Orenstein, N. *et al.* Bi-allelic IARS mutations in a child with intra-uterine growth retardation, neonatal cholestasis, and mild developmental delay. *Clinical Genetics* **91**, 913–917 (2017).

182. Kopajtich, R. *et al.* Biallelic IARS Mutations Cause Growth Retardation with Prenatal Onset, Intellectual Disability, Muscular Hypotonia, and Infantile Hepatopathy. *The American Journal of Human Genetics* **99**, 414–422 (2016).

183. Fagbemi, A. *et al.* Refractory very early-onset inflammatory bowel disease associated with cytosolic isoleucyl-tRNA synthetase deficiency: A case report. *World J Gastroenterol* **26**, 1841–1846 (2020).

184. Smigiel, R. *et al.* New evidence for association of recessive IARS gene mutations with hepatopathy, hypotonia, intellectual disability and growth retardation. *Clinical Genetics* **92**, 671–673 (2017).

185. Vona, B. *et al.* Expanding the clinical phenotype of IARS2-related mitochondrial disease. *BMC Med Genet* **19**, (2018).

186. McMillan, H. J. *et al.* Congenital Visual Impairment and Progressive Microcephaly Due to Lysyl-Transfer Ribonucleic Acid (RNA) Synthetase (KARS) Mutations: The

Expanding Phenotype of Aminoacyl-Transfer RNA Synthetase Mutations in Human Disease. *Journal of Child Neurology* **30**, 1037–1043 (2015).

187. Zhou, X.-L. *et al.* Mutations in KARS cause early-onset hearing loss and leukoencephalopathy: Potential pathogenic mechanism. *Hum Mutat* **38**, 1740–1750 (2017).

188. Scheidecker, S. *et al.* Mutations in KARS cause a severe neurological and neurosensory disease with optic neuropathy. *Hum Mutat* **40**, 1826–1840 (2019).

189. Itoh, M. *et al.* Biallelic KARS pathogenic variants cause an early-onset progressive leukodystrophy. *Brain* **142**, 560–573 (2019).

190. Vargas, A. *et al.* Progressive Early-Onset Leukodystrophy Related to Biallelic Variants in the KARS Gene: The First Case Described in Latin America. *Genes (Basel)* **11**, 1437 (2020).

191. Ardissone, A. *et al.* KARS-related diseases: progressive leukoencephalopathy with brainstem and spinal cord calcifications as new phenotype and a review of literature. *Orphanet J Rare Dis* **13**, 45 (2018).

192. Cappuccio, G. *et al.* Bi-allelic KARS1 pathogenic variants affecting functions of cytosolic and mitochondrial isoforms are associated with a progressive and multisystem disease. *Human Mutation* **42**, 745–761 (2021).

193. Peluso, F. *et al.* Leopard-like retinopathy and severe early-onset portal hypertension expand the phenotype of KARS1-related syndrome: a case report. *BMC Med Genomics* **14**, 25 (2021).

194. Lin, S.-J. *et al.* Biallelic variants in KARS1 are associated with neurodevelopmental disorders and hearing loss recapitulated by the knockout zebrafish. *Genet Med* 1–11 (2021) doi:10.1038/s41436-021-01239-1.

195. McLaughlin, H. M. *et al.* Compound Heterozygosity for Loss-of-Function Lysyl-tRNA Synthetase Mutations in a Patient with Peripheral Neuropathy. *The American Journal of Human Genetics* **87**, 560–566 (2010).

196. Casey, J. P. *et al.* Identification of a mutation in LARS as a novel cause of infantile hepatopathy. *Molecular genetics and metabolism* **106**, 351–358 (2012).

197. Casey, J. P. *et al.* Clinical and genetic characterisation of infantile liver failure syndrome type 1, due to recessive mutations in LARS. *Journal of inherited metabolic disease* **38**, 1085–1092 (2015).

198. Lenz, D. *et al.* Genotypic diversity and phenotypic spectrum of infantile liver failure syndrome type 1 due to variants in LARS1. *Genet Med* **22**, 1863–1873 (2020).

199. Hirata, K. *et al.* Severe course with lethal hepatocellular injury and skeletal muscular dysgenesis in a neonate with infantile liver failure syndrome type 1 caused by novel LARS1 mutations. *Am J Med Genet A* **185**, 866–870 (2021).

200. Peroutka, C. *et al.* Severe Neonatal Manifestations of Infantile Liver Failure Syndrome Type 1 Caused by Cytosolic Leucine-tRNA Synthetase Deficiency. *JIMD Rep* **45**, 71–76 (2019).

201. Tabolacci, E. *et al.* Infantile Liver Failure Syndrome 1 associated with a novel variant of the LARS1 gene: Clinical, genetic, and functional characterization. *Clin Genet* **99**, 601–603 (2021).

202. Soldà, G. *et al.* First independent replication of the involvement of LARS2 in Perrault syndrome by whole-exome sequencing of an Italian family. *Journal of Human Genetics* **61**, 295–300 (2016).

203. Kosaki, R., Horikawa, R., Fujii, E. & Kosaki, K. Biallelic mutations in LARS2 can cause Perrault syndrome type 2 with neurologic symptoms. *American Journal of Medical Genetics Part A* **176**, 404–408 (2018).

204. Al-Jaroudi, D., Enabi, S. & AlThagafi, M. S. Perrault syndrome with amenorrhea, infertility, Tarlov cyst, and degenerative disc. *Gynecol Endocrinol* **35**, 1037–1039 (2019).

205. Carminho-Rodrigues, M. T. *et al.* LARS2-Perrault syndrome: a new case report and literature review. *BMC Med Genet* **21**, 109 (2020).

206. Zafar, S. *et al.* Novel Mutations in CLPP, LARS2, CDH23, and COL4A5 Identified in Familial Cases of Prelingual Hearing Loss. *Genes (Basel)* **11**, 978 (2020).

207. van der Knaap, M. S. *et al.* Biallelic variants in LARS2 and KARS cause deafness and (ovario)leukodystrophy. *Neurology* **92**, e1225–e1237 (2019).

208. Tucker, E. J. *et al.* Genomic sequencing highlights the diverse molecular causes of Perrault syndrome: a peroxisomal disorder (PEX6), metabolic disorders (CLPP, GGPS1), and mtDNA maintenance/translation disorders (LARS2, TFAM). *Hum Genet* **139**, 1325–1343 (2020).

209. Pan, Z. *et al.* Perrault syndrome: Clinical report and retrospective analysis. *Mol Genet Genomic Med* **8**, e1445 (2020).

210. van Meel, E. *et al.* Rare recessive loss-of-function methionyl-tRNA synthetase mutations presenting as a multi-organ phenotype. *BMC medical genetics* **14**, 106 (2013).

211. Hadchouel, A. *et al.* Biallelic Mutations of Methionyl-tRNA Synthetase Cause a Specific Type of Pulmonary Alveolar Proteinosis Prevalent on Réunion Island. *The American Journal of Human Genetics* **96**, 826–831 (2015).

212. Lenz, D. *et al.* Rescue of respiratory failure in pulmonary alveolar proteinosis due to pathogenic MARS1 variants. *Pediatr Pulmonol* **55**, 3057–3066 (2020).

213. Rips, J. *et al.* MARS variant associated with both recessive interstitial lung and liver disease and dominant Charcot-Marie-Tooth disease. *Eur J Med Genet* **61**, 616–620 (2018).

214. Alzaid, M., Alshamrani, A., Harbi, A. S. A., Alenzi, A. & Mohamed, S. Methionyl-tRNA synthetase novel mutation causes pulmonary alveolar proteinosis. *Saudi Med J* **40**, 195–198 (2019).

215. Sun, Y. *et al.* Mutations in methionyl-tRNA synthetase gene in a Chinese family with interstitial lung and liver disease, postnatal growth failure and anemia. *Journal of Human Genetics* **48**, 337–651 (2017).

216. Gonzalez, M. *et al.* Exome sequencing identifies a significant variant in methionyl-tRNA synthetase (MARS) in a family with late-onset CMT2. *Journal of neurology, neurosurgery, and psychiatry* **84**, 1247–1249 (2013).

217. Hyun, Y. S. *et al.* Rare variants in methionyl- and tyrosyl-tRNA synthetase genes in late-onset autosomal dominant Charcot-Marie-Tooth neuropathy. *Clinical Genetics* **86**, 592–594 (2014).

218. Sagi-Dain, L. *et al.* Whole-exome sequencing reveals a novel missense mutation in the MARS gene related to a rare Charcot-Marie-Tooth neuropathy type 2U. *Journal of the Peripheral Nervous System* **23**, 138–142 (2018).

219. Gillespie, M. K. *et al.* A Novel Mutation in MARS in a Patient with Charcot-Marie-Tooth Disease, Axonal, Type 2U with Congenital Onset. *J Neuromuscul Dis* **6**, 333–339.

220. 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 Biology* **10**, e1001288 (2012).

221. Wang, L. *et al.* Loss of NARS1 impairs progenitor proliferation in cortical brain organoids and leads to microcephaly. *Nat Commun* **11**, (2020).

222. Manole, A. *et al.* De Novo and Bi-allelic Pathogenic Variants in NARS1 Cause Neurodevelopmental Delay Due to Toxic Gain-of-Function and Partial Loss-of-Function Effects. *Am J Hum Genet* **107**, 311–324 (2020).

223. 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. *Molecular genetics & genomic medicine* **3**, 59–68 (2015).

224. Sofou, K., Kollberg, G., Hedberg-Oldfors, C. & Oldfors, A. The phenotypic variability and natural history of NARS2 associated disease. *European Journal of Paediatric Neurology* **31**, 31–37 (2021).

225. Palombo, F. *et al.* Autozygosity-driven genetic diagnosis in consanguineous families from Italy and the Greater Middle East. *Hum Genet* **139**, 1429–1441 (2020).

226. Seaver, L. H. *et al.* Lethal NARS2-Related Disorder Associated With Rapidly Progressive Intractable Epilepsy and Global Brain Atrophy. *Pediatr Neurol* **89**, 26–30 (2018).

227. Lee, J. S. *et al.* Genetic heterogeneity in Leigh syndrome: Highlighting treatable and novel genetic causes. *Clinical Genetics* **97**, 586–594 (2020).

228. Al Balushi, A. *et al.* Phenotypes and genotypes of mitochondrial aminoacyl-tRNA synthetase deficiencies from a single neurometabolic clinic. *JIMD Rep* **51**, 3–10 (2019).

229. Yin, X. *et al.* The genotypic and phenotypic spectrum of PARS2-related infantile-onset encephalopathy. *J Hum Genet* **63**, 971–980 (2018).

230. Zhang, X. *et al.* Mutations in QARS, encoding glutaminyl-tRNA synthetase, cause progressive microcephaly, cerebral-cerebellar atrophy, and intractable seizures. *American journal of human genetics* **94**, 547–558 (2014).

231. Salvarinova, R. *et al.* Expansion of the QARS deficiency phenotype with report of a family with isolated supratentorial brain abnormalities. *neurogenetics* **16**, 145–149 (2015).

232. Datta, A. *et al.* Case Report. *J Child Neurol* **32**, 403–407 (2017).

233. Leshinsky-Silver, E. *et al.* Severe growth deficiency, microcephaly, intellectual disability, and characteristic facial features are due to a homozygous QARS mutation. *Neurogenetics* **18**, 141–146 (2017).

234. Wolf, N. I. *et al.* Mutations in RARS cause hypomyelination. *Annals of Neurology* **76**, 134–139 (2014).

235. Nafisinia, M. *et al.* Mutations in RARS cause a hypomyelination disorder akin to Pelizaeus-Merzbacher disease. *Eur J Hum Genet* **25**, 1134–1141 (2017).

236. Mendes, M. I. *et al.* RARS1-related hypomyelinating leukodystrophy: Expanding the spectrum. *Ann Clin Transl Neurol* **7**, 83–93 (2019).

237. Rezaei, Z. *et al.* Hypomyelinating Leukodystrophy with Spinal Cord Involvement Caused by a Novel Variant in RARS: Report of Two Unrelated Patients. *Neuropediatrics* **50**, 130–134 (2019).

238. Li, Z. *et al.* A novel mutation in the promoter of RARS2 causes pontocerebellar hypoplasia in two siblings. *Journal of Human Genetics* **60**, 363–369 (2015).

239. Ngoh, A. *et al.* RARS2 mutations in a sibship with infantile spasms. *Epilepsia* **57**, e97–e102 (2016).

240. Cassandrini, D. *et al.* Pontocerebellar hypoplasia type 6 caused by mutations in RARS2: definition of the clinical spectrum and molecular findings in five patients. *Journal of inherited metabolic disease* **36**, 43–53 (2013).

241. Edvardson, S. *et al.* Deleterious mutation in the mitochondrial arginyl-transfer RNA synthetase gene is associated with pontocerebellar hypoplasia. *The American Journal of Human Genetics* **81**, 857–862 (2007).

242. Rankin, J. *et al.* Pontocerebellar hypoplasia type 6: A British case with PEHO-like features. *American Journal of Medical Genetics Part A* **152A**, 2079–2084 (2010).

243. Namavar, Y. *et al.* Clinical, neuroradiological and genetic findings in pontocerebellar hypoplasia. *Brain : a journal of neurology* **134**, 143–156 (2011).

244. Glamuzina, E. *et al.* Further delineation of pontocerebellar hypoplasia type 6 due to mutations in the gene encoding mitochondrial arginyl-tRNA synthetase, RARS2. *Journal of inherited metabolic disease* **35**, 459–467 (2012).

245. Nishri, D. *et al.* RARS2 mutations cause early onset epileptic encephalopathy without ponto-cerebellar hypoplasia. *European Journal of Paediatric Neurology* **20**, 412–417 (2016).

246. van Dijk, T. *et al.* RARS2 Mutations: Is Pontocerebellar Hypoplasia Type 6 a Mitochondrial Encephalopathy? *JIMD Reports* **33**, 87–92 (2017).

247. Nevanlinna, V. *et al.* A patient with pontocerebellar hypoplasia type 6: Novel RARS2 mutations, comparison to previously published patients and clinical distinction from PEHO syndrome. *European Journal of Medical Genetics* **63**, 103766 (2020).

248. Musante, L. *et al.* Mutations of the aminoacyl-tRNA-synthetases SARS and WARS2 are implicated in the etiology of autosomal recessive intellectual disability. *Human mutation* **38**, 621–636 (2017).

249. Belostotsky, R. *et al.* Mutations in the mitochondrial seryl-tRNA synthetase cause hyperuricemia, pulmonary hypertension, renal failure in infancy and alkalosis, HUPRA syndrome. *American journal of human genetics* **88**, 193–200 (2011).

250. Rivera, H. *et al.* A new mutation in the gene encoding mitochondrial seryl-tRNA synthetase as a cause of HUPRA syndrome. *BMC nephrology* **14**, 195 (2013).

251. Zhou, Y. *et al.* Novel SARS2 variants identified in a Chinese girl with HUPRA syndrome. *Mol Genet Genomic Med* **9**, e1650 (2021).

252. Linnankivi, T., Neupane, N., Richter, U., Isohanni, P. & Tyynismaa, H. Splicing Defect in Mitochondrial Seryl-tRNA Synthetase Gene Causes Progressive Spastic Paresis Instead of HUPRA Syndrome. *Human mutation* **37**, 884–888 (2016).

253. Theil, A. F. *et al.* Bi-allelic TARS Mutations Are Associated with Brittle Hair Phenotype. *Am J Hum Genet* **105**, 434–440 (2019).

254. Diodato, D. *et al.* VARS2 and TARS2 mutations in patients with mitochondrial encephalomyopathies. *Human mutation* **35**, 983–989 (2014).

255. Li, X. *et al.* Novel compound heterozygous TARS2 variants in a Chinese family with mitochondrial encephalomyopathy: a case report. *BMC Med Genet* **21**, 217 (2020).

256. Karaca, E. *et al.* Genes that Affect Brain Structure and Function Identified by Rare Variant Analyses of Mendelian Neurologic Disease. *Neuron* **88**, 499–513 (2015).

257. Friedman, J. *et al.* Biallelic mutations in valyl-tRNA synthetase gene VARS are associated with a progressive neurodevelopmental epileptic encephalopathy. *Nat Commun* **10**, 707 (2019).

258. Siekierska, A. *et al.* Biallelic VARS variants cause developmental encephalopathy with microcephaly that is recapitulated in vars knockout zebrafish. *Nat Commun* **10**, 708 (2019).

259. Stephen, J. *et al.* Loss of function mutations in VARS encoding cytoplasmic valyl-tRNA synthetase cause microcephaly, seizures, and progressive cerebral atrophy. *Hum Genet* **137**, 293–303 (2018).

260. Ruzman, L. *et al.* A novel VARS2 gene variant in a patient with epileptic encephalopathy. *Ups J Med Sci* **124**, 273–277 (2019).

261. Bruni, F. *et al.* Clinical, biochemical, and genetic features associated with VARS2-related mitochondrial disease. *Hum Mutat* **39**, 563–578 (2018).

262. Chin, H.-L. *et al.* A combination of two novel VARS2 variants causes a mitochondrial disorder associated with failure to thrive and pulmonary hypertension. *J Mol Med (Berl)* **97**, 1557–1566 (2019).

263. Pereira, S. *et al.* Mitochondrial Encephalopathy: First Portuguese Report of a VARS2 Causative Variant. *JIMD Rep* **42**, 113–119 (2018).

264. Tsai, P.-C. *et al.* A recurrent WARS mutation is a novel cause of autosomal dominant distal hereditary motor neuropathy. *Brain : a journal of neurology* (2017).

265. Wang, B. *et al.* A novel WARS mutation (p.Asp314Gly) identified in a Chinese distal hereditary motor neuropathy family. *Clinical Genetics* **96**, 176–182 (2019).

266. Maffezzini, C. *et al.* Mutations in the mitochondrial tryptophanyl-tRNA synthetase cause growth retardation and progressive leukoencephalopathy. *Mol Genet Genomic Med* **7**, e654 (2019).

267. Virdee, M., Swarnalingam, E., Kozenko, M., Tarnopolsky, M. & Jones, K. Expanding the Phenotype: Neurodevelopmental Disorder, Mitochondrial, With Abnormal Movements and Lactic Acidosis, With or Without Seizures (NEMMLAS) due to WARS2 Biallelic Variants, Encoding Mitochondrial Tryptophanyl-tRNA Synthase. *J Child Neurol* **34**, 778–781 (2019).

268. Vantroys, E. *et al.* Severe hepatopathy and neurological deterioration after start of valproate treatment in a 6-year-old child with mitochondrial tryptophanyl-tRNA synthetase deficiency. *Orphanet J Rare Dis* **13**, 80 (2018).

269. Burke, E. A. *et al.* Biallelic mutations in mitochondrial tryptophanyl-tRNA synthetase cause Levodopa-responsive infantile-onset Parkinsonism. *Clin Genet* **93**, 712–718 (2018).

270. Hübers, A., Huppertz, H.-J., Wortmann, S. B. & Kassubek, J. Mutation of the WARS2 Gene as the Cause of a Severe Hyperkinetic Movement Disorder. *Mov Disord Clin Pract* **7**, 88–90 (2020).

271. 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* **173**, 126–134 (2016).

272. Tracewska-Siemiątkowska, A. *et al.* An Expanded Multi-Organ Disease Phenotype Associated with Mutations in YARS. *Genes (Basel)* **8**, 381 (2017).

273. Williams, K. B. *et al.* Homozygosity for a mutation affecting the catalytic domain of tyrosyl-tRNA synthetase (YARS) causes multisystem disease. *Hum Mol Genet* **28**, 525–538 (2019).

274. Zeiad, R. K. H. M. *et al.* A Novel Homozygous Missense Mutation in the YARS Gene: Expanding the Phenotype of YARS Multisystem Disease. *J Endocr Soc* **5**, bvaa196 (2021).

275. Jordanova, A. *et al.* Disrupted function and axonal distribution of mutant tyrosyl-tRNA synthetase in dominant intermediate Charcot-Marie-Tooth neuropathy. *Nature Genetics* **38**, 197–202 (2006).

276. Gonzaga-Jauregui, C. *et al.* Exome Sequence Analysis Suggests that Genetic Burden Contributes to Phenotypic Variability and Complex Neuropathy. *Cell Rep* **12**, 1169–1183 (2015).

277. Sasarman, F., Nishimura, T., Thiffault, I. & Shoubridge, E. A. A novel mutation in YARS2 causes myopathy with lactic acidosis and sideroblastic anemia. *Human mutation* **33**, 1201–1206 (2012).

278. Nakajima, J. *et al.* A novel homozygous YARS2 mutation causes severe myopathy, lactic acidosis, and sideroblastic anemia 2. *Journal of Human Genetics* **59**, 229–232 (2014).

279. Shahni, R. *et al.* A distinct mitochondrial myopathy, lactic acidosis and sideroblastic anemia (MLASA) phenotype associates with YARS2 mutations. *American Journal of Medical Genetics Part A* **161A**, 2334–2338 (2013).

280. Smith, F. *et al.* Sideroblastic anemia with myopathy secondary to novel, pathogenic missense variants in the YARS2 gene. *Haematologica* **103**, e564–e566 (2018).

281. Riley, L. G. *et al.* The phenotypic spectrum of germline YARS2 variants: from isolated sideroblastic anemia to mitochondrial myopathy, lactic acidosis and sideroblastic anemia 2. *Haematologica* **103**, 2008–2015 (2018).

282. Ishimura, R. *et al.* RNA function. Ribosome stalling induced by mutation of a CNS-specific tRNA causes neurodegeneration. *Science* **345**, 455–459 (2014).

283. Skre, H. Genetic and clinical aspects of Charcot-Marie-Tooth's disease. *Clinical Genetics* **6**, 98–118 (1974).

284. Braathen, G. J., Sand, J. C., Lobato, A., Høyer, H. & Russell, M. B. Genetic epidemiology of Charcot-Marie-Tooth in the general population. *Eur J Neurol* **18**, 39–48 (2011).

285. Shy, M. E., Lupski, J. R., Chance, P. F., Klein, C. J. & Dyck, P. J. Hereditary Motor and Sensory Neuropathies: An Overview of Clinical, Genetic, Electrophysiologic, and Pathologic Features. 36.

286. Antonellis, A., Goldfarb, L. G. & Sivakumar, K. GARS1-Associated Axonal Neuropathy. in *GeneReviews®* (eds. Adam, M. P. et al.) (University of Washington, Seattle, 1993).

287. Stavrou, M., Sargiannidou, I., Christofi, T. & Kleopa, K. A. Genetic mechanisms of peripheral nerve disease. *Neuroscience Letters* **742**, 135357 (2021).

288. Pareyson, D., Scaioli, V. & Laurà, M. Clinical and electrophysiological aspects of Charcot-Marie-Tooth disease. *Neuromolecular Med* **8**, 3–22 (2006).

289. Züchner, S. *et al.* Mutations in the mitochondrial GTPase mitofusin 2 cause Charcot-Marie-Tooth neuropathy type 2A. *Nature Genetics* **36**, 449–451 (2004).

290. Filadi, R., Pendin, D. & Pizzo, P. Mitofusin 2: from functions to disease. *Cell Death Dis* **9**, 1–13 (2018).

291. Verhoeven, K. *et al.* Mutations in the small GTP-ase late endosomal protein RAB7 cause Charcot-Marie-Tooth type 2B neuropathy. *Am J Hum Genet* **72**, 722–727 (2003).

292. Cherry, S. *et al.* Charcot-Marie-Tooth 2B mutations in rab7 cause dosage-dependent neurodegeneration due to partial loss of function. *eLife* **2**, e01064 (2013).

293. Cosker, K. E. & Segal, R. A. Neuronal signaling through endocytosis. *Cold Spring Harb Perspect Biol* **6**, a020669 (2014).

294. Cioni, J.-M. *et al.* Late Endosomes Act as mRNA Translation Platforms and Sustain Mitochondria in Axons. *Cell* **176**, 56-72.e15 (2019).

295. Beijer, D. & Baets, J. The expanding genetic landscape of hereditary motor neuropathies. *Brain* **143**, 3540–3563 (2020).

296. Sambuughin, N. *et al.* Autosomal dominant distal spinal muscular atrophy type V (dSMA-V) and Charcot-Marie-Tooth disease type 2D (CMT2D) segregate within a single large kindred and map to a refined region on chromosome 7p15. *Journal of the Neurological Sciences* **161**, 23–28 (1998).

297. Francklyn, C. S., First, E. A., Perona, J. J. & Hou, Y.-M. Methods for Kinetic and Thermodynamic Analysis of Aminoacyl-tRNA Synthetases. *Methods* **44**, 100–118 (2008).

298. Chang, G. G., Pan, F., Yeh, C. & Huang, T. M. Colorimetric assay for aminoacyl-tRNA synthetases. *Anal Biochem* **130**, 171–176 (1983).

299. Boeke, J. D., Trueheart, J., Natsoulis, G. & Fink, G. R. 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. *Methods in enzymology* **154**, 164–175 (1987).

300. Nangle, L. A., Zhang, W., Xie, W., Yang, X.-L. & Schimmel, P. Charcot-Marie-Tooth disease-associated mutant tRNA synthetases linked to altered dimer interface and neurite distribution defect. *Proceedings of the National Academy of Sciences* **104**, 11239–11244 (2007).

301. Antonellis, A. *et al.* Functional analyses of glycyl-tRNA synthetase mutations suggest a key role for tRNA-charging enzymes in peripheral axons. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **26**, 10397–10406 (2006).

302. Götz, A. *et al.* Exome sequencing identifies mitochondrial alanyl-tRNA synthetase mutations in infantile mitochondrial cardiomyopathy. *American journal of human genetics* **88**, 635–642 (2011).

303. Jiang, P. *et al.* The exome sequencing identified the mutation in YARS2 encoding the mitochondrial tyrosyl-tRNA synthetase as a nuclear modifier for the phenotypic manifestation of Leber's hereditary optic neuropathy-associated mitochondrial DNA mutation. *Human molecular genetics* **25**, 584–596 (2016).

304. Darnell, A. M., Subramaniam, A. R. & O'Shea, E. K. Translational control through differential ribosome pausing during amino acid limitation in mammalian cells. *Mol Cell* **71**, 229-243.e11 (2018).

305. Ishimura, R., Nagy, G., Dotu, I., Chuang, J. H. & Ackerman, S. L. Activation of GCN2 kinase by ribosome stalling links translation elongation with translation initiation. *eLife* **5**, e14295 (2016).

306. Niehues, S. *et al.* Impaired protein translation in Drosophila models for Charcot–Marie–Tooth neuropathy caused by mutant tRNA synthetases. *Nature Communications* **6**, 7520–13 (2015).

307. Oprescu, S. N., Griffin, L. B., Beg, A. A. & Antonellis, A. Predicting the pathogenicity of aminoacyl-tRNA synthetase mutations. *Methods* **113**, 139–151 (2017).

308. Yu, M. *et al.* Amino acids stimulate glycyl-tRNA synthetase nuclear localization for mammalian target of rapamycin expression in bovine mammary epithelial cells. *Journal of Cellular Physiology* **234**, 7608–7621 (2019).

309. Park, S.-R. *et al.* A novel endogenous damage signal, glycyl tRNA synthetase, activates multiple beneficial functions of mesenchymal stem cells. *Cell Death Differ* **25**, 2023–2036 (2018).

310. Mo, Z. *et al.* Neddylation requires glycyl-tRNA synthetase to protect activated E2. *Nature structural & molecular biology* **23**, 730–737 (2016).

311. Park, M. C. *et al.* Secreted human glycyl-tRNA synthetase implicated in defense against ERK-activated tumorigenesis. *Proceedings of the National Academy of Sciences of the United States of America* **109**, E640-7 (2012).

312. Johanson, K., Hoang, T., Sheth, M. & Hyman, L. E. GRS1, a yeast tRNA synthetase with a role in mRNA 3' end formation. *Journal of Biological Chemistry* **278**, 35923–35930 (2003).

313. Ni, R. & Luo, L. A noncanonical function of histidyl-tRNA synthetase: inhibition of vascular hyperbranching during zebrafish development. *FEBS Open Bio* **8**, 722–731 (2018).

314. Levi, O. & Arava, Y. mRNA association by aminoacyl tRNA synthetase occurs at a putative anticodon mimic and autoregulates translation in response to tRNA levels. *PLoS Biol* **17**, e3000274 (2019).

315. Zhou, J. J. *et al.* Secreted histidyl-tRNA synthetase splice variants elaborate major epitopes for autoantibodies in inflammatory myositis. *The Journal of biological chemistry* **289**, 19269–19275 (2014).

316. Raben, N. *et al.* A motif in human histidyl-tRNA synthetase which is shared among several aminoacyl-tRNA synthetases is a coiled-coil that is essential for enzymatic activity and contains the major autoantigenic epitope. *Journal of Biological Chemistry* **269**, 24277–24283 (1994).

317. Howard, O. M. Z. *et al.* Histidyl-tRNA synthetase and asparaginyl-tRNA synthetase, autoantigens in myositis, activate chemokine receptors on T lymphocytes and immature dendritic cells. *The Journal of experimental medicine* **196**, 781–791 (2002).

318. Yeung, M. L. *et al.* Human tryptophanyl-tRNA synthetase is an IFN-γ–inducible entry factor for Enterovirus. *J Clin Invest* **128**, 5163–5177.

319. Miyanokoshi, M., Yokosawa, T. & Wakasugi, K. Tryptophanyl-tRNA synthetase mediates high-affinity tryptophan uptake into human cells. *J Biol Chem* **293**, 8428–8438 (2018).

320. Lee, H.-C. *et al.* Released Tryptophanyl-tRNA Synthetase Stimulates Innate Immune Responses against Viral Infection. *J Virol* **93**, e01291-18 (2019).

321. Otani, A. *et al.* A fragment of human TrpRS as a potent antagonist of ocular angiogenesis. *Proceedings of the National Academy of Sciences* **99**, 178–183 (2002).

322. Zeng, R. *et al.* Effect of mini-tyrosyl-tRNA synthetase/mini-tryptophanyl-tRNA synthetase on ischemic angiogenesis in rats: proliferation and migration of endothelial cells. *Heart and vessels* **26**, 69–80 (2011).

323. Zeng, R. *et al.* Inhibition of mini-TyrRS-induced angiogenesis response in endothelial cells by VE-cadherin-dependent mini-TrpRS. *Heart and vessels* **27**, 193–201 (2012).

324. Zeng, R. *et al.* Effect of Mini-Tyrosyl-tRNA Synthetase/Mini-Tryptophanyl-tRNA Synthetase on Angiogenesis in Rhesus Monkeys after Acute Myocardial Infarction. *Cardiovascular therapeutics* **34**, 4–12 (2016).

325. Sajish, M. *et al.* Trp-tRNA synthetase bridges DNA-PKcs to PARP-1 to link IFN-γ and p53 signaling. *Nature chemical biology* **8**, 547–554 (2012).

326. Kanaji, T. *et al.* Tyrosyl-tRNA synthetase stimulates thrombopoietin-independent hematopoiesis accelerating recovery from thrombocytopenia. *Proc Natl Acad Sci U S A* **115**, E8228–E8235 (2018).

327. Cao, X. *et al.* Acetylation promotes TyrRS nuclear translocation to prevent oxidative damage. *Proceedings of the National Academy of Sciences* **114**, 687–692 (2017).

328. Fu, G., Xu, T., Shi, Y., Wei, N. & Yang, X.-L. tRNA-controlled nuclear import of a human tRNA synthetase. *The Journal of biological chemistry* **287**, 9330–9334 (2012).

329. Wei, N. *et al.* Oxidative stress diverts tRNA synthetase to nucleus for protection against DNA damage. *Molecular Cell* **56**, 323–332 (2014).

330. Cheng, G. *et al.* Effect of mini-tyrosyl-tRNA synthetase on ischemic angiogenesis, leukocyte recruitment, and vascular permeability. *American journal of physiology. Regulatory, integrative and comparative physiology* **295**, R1138-46 (2008).

331. Chihara, T., Luginbuhl, D. & Luo, L. Cytoplasmic and mitochondrial protein translation in axonal and dendritic terminal arborization. *Nature neuroscience* **10**, 828–837 (2007).

332. Malissovas, N., Griffin, L. B., Antonellis, A. & Beis, D. Dimerization is required for GARS-mediated neurotoxicity in dominant CMT disease. *Human molecular genetics* **25**, 1528–1542 (2016).

333. Motley, W. W. *et al.* Charcot-Marie-Tooth-linked mutant GARS is toxic to peripheral neurons independent of wild-type GARS levels. *PLoS genetics* **7**, e1002399 (2011).

334. Karczewski, K. J. *et al.* The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434–443 (2020).

335. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature Publishing Group* **536**, 285–291 (2016).

336. Herskowitz, I. Functional inactivation of genes by dominant negative mutations. *Nature* **329**, 219–222 (1987).

337. Willis, A., Jung, E. J., Wakefield, T. & Chen, X. Mutant p53 exerts a dominant negative effect by preventing wild-type p53 from binding to the promoter of its target genes. *Oncogene* **23**, 2330–2338 (2004).

338. Barren, B. & Artemyev, N. O. Mechanisms of dominant negative G-protein alpha subunits. *J Neurosci Res* **85**, 3505–3514 (2007).

339. Doisne, N. *et al.* In vivo Dominant-Negative Effect of an SCN5A Brugada Syndrome Variant. *Front Physiol* **12**, 661413 (2021).

340. Shalaby, F. Y. *et al.* Dominant-negative KvLQT1 mutations underlie the LQT1 form of long QT syndrome. *Circulation* **96**, 1733–1736 (1997).

341. McLean, W. H. I. & Moore, C. B. T. Keratin disorders: from gene to therapy. *Hum Mol Genet* **20**, R189-197 (2011).

342. Veitia, R. A., Caburet, S. & Birchler, J. A. Mechanisms of Mendelian dominance. *Clinical Genetics* **93**, 419–428 (2018).

343. Friesen, R. H. E. & Lee, J. C. The Negative Dominant Effects of T340M Mutation on Mammalian Pyruvate Kinase*. *Journal of Biological Chemistry* **273**, 14772–14779 (1998).

344. Zhou, R., Yang, G. & Shi, Y. Dominant negative effect of the loss-of-function γ-secretase mutants on the wild-type enzyme through heterooligomerization. *Proc Natl Acad Sci U S A* **114**, 12731–12736 (2017).

345. Wada, H., Yeh, E. T. & Kamitani, T. A dominant-negative UBC12 mutant sequesters NEDD8 and inhibits NEDD8 conjugation in vivo. *J Biol Chem* **275**, 17008–17015 (2000).

346. Elsevier, J. P. & Fridovich-Keil, J. L. The Q188R mutation in human galactose-1-phosphate uridylyltransferase acts as a partial dominant negative. *J Biol Chem* **271**, 32002–32007 (1996).

347. Achilli, F. *et al.* An ENU-induced mutation in mouse glycyl-tRNA synthetase (GARS) causes peripheral sensory and motor phenotypes creating a model of Charcot-Marie-Tooth type 2D peripheral neuropathy. *Disease models & mechanisms* **2**, 359–373 (2009).

348. Sun, L. *et al.* CMT2N-causing aminoacylation domain mutants enable Nrp1 interaction with AlaRS. *Proc Natl Acad Sci U S A* **118**, (2021).

349. Lee, D. C. *et al.* A recurrent GARS mutation causes distal hereditary motor neuropathy. *J Peripher Nerv Syst* **24**, 320–323 (2019).

350. Mullen, P. *et al.* Neuropathy-associated histidyl-tRNA synthetase variants attenuate protein synthesis in vitro and disrupt axon outgrowth in developing zebrafish. *FEBS J.* (2020) doi:10.1111/febs.15449.

351. Bond, S., Lopez-Lloreda, C., Gannon, P. J., Akay-Espinoza, C. & Jordan-Sciutto, K. L. The Integrated Stress Response and Phosphorylated Eukaryotic Initiation Factor 2α in Neurodegeneration. *Journal of Neuropathology & Experimental Neurology* **79**, 123–143 (2020).

352. Meggouh, F., Bienfait, H. M. E., Weterman, M. A. J., de Visser, M. & Baas, F. Charcot-Marie-Tooth disease due to a de novo mutation of the RAB7 gene. *Neurology* **67**, 1476–1478 (2006).

353. Fritsch, A. *et al.* Dominant-negative effects of COL7A1 mutations can be rescued by controlled overexpression of normal collagen VII. *J Biol Chem* **284**, 30248–30256 (2009).

354. Zariñán, T. *et al.* Dominant negative effects of human follicle-stimulating hormone receptor expression-deficient mutants on wild-type receptor cell surface expression. Rescue of oligomerization-dependent defective receptor expression by using cognate decoys. *Mol Cell Endocrinol* **321**, 112–122 (2010).

355. Blocquel, D. *et al.* Alternative stable conformation capable of protein misinteraction links tRNA synthetase to peripheral neuropathy. *Nucleic acids research* (2017).

356. He, W. *et al.* CMT2D neuropathy is linked to the neomorphic binding activity of glycyl-tRNA synthetase. *Nature* **526**, 710–714 (2015).

357. Blocquel, D. *et al.* CMT disease severity correlates with mutation-induced open conformation of histidyl-tRNA synthetase, not aminoacylation loss, in patient cells. *Proc Natl Acad Sci U S A* **116**, 19440–19448 (2019).

358. Neufeld, G. *et al.* The neuropilins: multifunctional semaphorin and VEGF receptors that modulate axon guidance and angiogenesis. *Trends in cardiovascular medicine* **12**, 13–19 (2002).

359. Rosenstein, J. M., Krum, J. M. & Ruhrberg, C. VEGF in the nervous system. *Organogenesis* **6**, 107–114 (2010).

360. He, W. *et al.* CMT2D neuropathy is linked to the neomorphic binding activity of glycyl-tRNA synthetase. *Nature* **526**, 710–714 (2015).

361. Sleigh, J. N. *et al.* Neuropilin 1 sequestration by neuropathogenic mutant glycyl-tRNA synthetase is permissive to vascular homeostasis. *Sci Rep* **7**, 9216 (2017).

362. Raimondi, C., Brash, J. T., Fantin, A. & Ruhrberg, C. NRP1 function and targeting in neurovascular development and eye disease. *Prog Retin Eye Res* **52**, 64–83 (2016).

363. Lindsay, R. M. Role of neurotrophins and trk receptors in the development and maintenance of sensory neurons: an overview. *Philos Trans R Soc Lond B Biol Sci* **351**, 365–373 (1996).

364. Sleigh, J. N. *et al.* Trk receptor signaling and sensory neuron fate are perturbed in human neuropathy caused by Gars mutations. *Proceedings of the National Academy of Sciences of the United States of America* 201614557 (2017).

365. Zhang, Y. *et al.* HDAC-6 interacts with and deacetylates tubulin and microtubules in vivo. *EMBO J* **22**, 1168–1179 (2003).

366. Chakraborti, S., Natarajan, K., Curiel, J., Janke, C. & Liu, J. The emerging role of the tubulin code: From the tubulin molecule to neuronal function and disease. *Cytoskeleton* **73**, 521–550 (2016).

367. Reed, N. A. *et al.* Microtubule acetylation promotes kinesin-1 binding and transport. *Curr Biol* **16**, 2166–2172 (2006).

368. Benoy, V. *et al.* HDAC6 is a therapeutic target in mutant GARS-induced Charcot-Marie-Tooth disease. *Brain* **141**, 673–687 (2018).

369. Mo, Z. *et al.* Aberrant GlyRS-HDAC6 interaction linked to axonal transport deficits in Charcot-Marie-Tooth neuropathy. *Nat Commun* **9**, 1007 (2018).

370. Mo, Z. *et al.* Aberrant GlyRS-HDAC6 interaction linked to axonal transport deficits in Charcot-Marie-Tooth neuropathy. *Nat Commun* **9**, 1007 (2018).

371. Griffin, L. B. *et al.* Impaired function is a common feature of neuropathy-associated glycyl-tRNA synthetase mutations. *Human mutation* **35**, 1363–1371 (2014).

372. Picci, C. *et al.* HDAC6 inhibition promotes α-tubulin acetylation and ameliorates CMT2A peripheral neuropathy in mice. *Experimental Neurology* **328**, 113281 (2020).

373. d'Ydewalle, C. *et al.* HDAC6 inhibitors reverse axonal loss in a mouse model of mutant HSPB1-induced Charcot-Marie-Tooth disease. *Nat Med* **17**, 968–974 (2011).

374. Krukowski, K. *et al.* HDAC6 inhibition effectively reverses chemotherapy-induced peripheral neuropathy. *Pain* **158**, 1126–1137 (2017).

375. Lund, E. & Dahlberg, J. E. Proofreading and aminoacylation of tRNAs before export from the nucleus. *Science* **282**, 2082–2085 (1998).

376. Bervoets, S. *et al.* Transcriptional dysregulation by a nucleus-localized aminoacyl-tRNA synthetase associated with Charcot-Marie-Tooth neuropathy. *Nat Commun* **10**, 5045 (2019).

377. Kuo, M. E. & Antonellis, A. Ubiquitously Expressed Proteins and Restricted Phenotypes: Exploring Cell-Specific Sensitivities to Impaired tRNA Charging. *Trends in Genetics* **36**, 105–117 (2020).

378. Jarvik, G. P. & Browning, B. L. Consideration of Cosegregation in the Pathogenicity Classification of Genomic Variants. *The American Journal of Human Genetics* **98**, 1077–1081 (2016).

379. Jordanova, A. *et al.* Dominant intermediate Charcot-Marie-Tooth type C maps to chromosome 1p34-p35. *The American Journal of Human Genetics* **73**, 1423–1430 (2003).

380. Ionasescu, V. V. *et al.* A Dejerine-Sottas neuropathy family with a gene mapped on chromosome 8. *Muscle & Nerve* **19**, 319–323 (1996).

381. Christodoulou, K. *et al.* Mapping of a distal form of spinal muscular atrophy with upper limb predominance to chromosome 7p. *Human Molecular Genetics* **4**, 1629–1632 (1995).

382. Froelich, C. A. & First, E. A. Dominant Intermediate Charcot-Marie-Tooth Disorder Is Not Due to a Catalytic Defect in Tyrosyl-tRNA Synthetase. *Biochemistry* **50**, 7132–7145 (2011).

383. Gonzaga-Jauregui, C. *et al.* Exome Sequence Analysis Suggests that Genetic Burden Contributes to Phenotypic Variability and Complex Neuropathy. *CellReports* **12**, 1169–1183 (2015).

384. Storkebaum, E. *et al.* Dominant mutations in the tyrosyl-tRNA synthetase gene recapitulate in Drosophila features of human Charcot-Marie-Tooth neuropathy. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 11782–11787 (2009).

385. Madeira, F. *et al.* The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res* **47**, W636–W641 (2019).

386. Chien, C.-I. *et al.* Functional Substitution of a Eukaryotic Glycyl-tRNA Synthetase with an Evolutionarily Unrelated Bacterial Cognate Enzyme. *PLoS One* **9**, (2014).

387. Turner, R. J., Lovato, M. & Schimmel, P. One of Two Genes Encoding Glycyl-tRNA Synthetase inSaccharomyces cerevisiae Provides Mitochondrial and Cytoplasmic Functions*. *Journal of Biological Chemistry* **275**, 27681–27688 (2000).

388. Abbott, J. A. *et al.* The Usher Syndrome Type IIIB Histidyl-tRNA Synthetase Mutation Confers Temperature Sensitivity. *Biochemistry* **56**, 3619–3631 (2017).

389. Wolfson, A. D., Pleiss, J. A. & Uhlenbeck, O. C. A new assay for tRNA aminoacylation kinetics. *RNA* **4**, 1019–1023 (1998).

390. Abuduxikuer, K. *et al.* Novel methionyl-tRNA synthetase gene variants/phenotypes in interstitial lung and liver disease: A case report and review of literature. *World J Gastroenterol* **24**, 4208–4216 (2018).

391. Price, V. H., Odom, R. B., Ward, W. H. & Jones, F. T. Trichothiodystrophy: sulfur-deficient brittle hair as a marker for a neuroectodermal symptom complex. *Arch Dermatol* **116**, 1375–1384 (1980).

392. Kaminska, M., Shalak, V. & Mirande, M. The appended C-domain of human methionyl-tRNA synthetase has a tRNA-sequestering function. *Biochemistry* **40**, 14309–14316 (2001).

393. Fecto, F. *et al.* SQSTM1 mutations in familial and sporadic amyotrophic lateral sclerosis. *Arch Neurol* **68**, 1440–1446 (2011).

394. Le Ber, I. *et al.* SQSTM1 Mutations in French Patients With Frontotemporal Dementia or Frontotemporal Dementia With Amyotrophic Lateral Sclerosis. *JAMA Neurol* **70**, 1403–1410 (2013).

395. Haack, T. B. *et al.* Absence of the Autophagy Adaptor SQSTM1/p62 Causes Childhood-Onset Neurodegeneration with Ataxia, Dystonia, and Gaze Palsy. *The American Journal of Human Genetics* **99**, 735–743 (2016).

396. Zhao, C. *et al.* Charcot-Marie-Tooth Disease Type 2A Caused by Mutation in a Microtubule Motor KIF1Bβ. *Cell* **105**, 587–597 (2001).

397. Sabblah, T. T. *et al.* A novel mouse model carrying a human cytoplasmic dynein mutation shows motor behavior deficits consistent with Charcot-Marie-Tooth type 2O disease. *Sci Rep* **8**, 1739 (2018).

398. Dequen, F. *et al.* Reversal of neuropathy phenotypes in conditional mouse model of Charcot–Marie–Tooth disease type 2E. *Human Molecular Genetics* **19**, 2616–2629 (2010).

399. Filali, M., Dequen, F., Lalonde, R. & Julien, J.-P. Sensorimotor and cognitive function of a NEFL(P22S) mutant model of Charcot-Marie-Tooth disease type 2E. *Behav Brain Res* **219**, 175–180 (2011).

400. Boeke, J. D., LaCroute, F. & Fink, G. R. A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. *Mol Gen Genet* **197**, 345–346 (1984).

401. Jin, Y., Jorgensen, E., Hartwieg, E. & Horvitz, H. R. The Caenorhabditis elegans gene unc-25 encodes glutamic acid decarboxylase and is required for synaptic transmission but not synaptic development. *J Neurosci* **19**, 539–548 (1999).

402. Berkowitz, L. A., Knight, A. L., Caldwell, G. A. & Caldwell, K. A. Generation of Stable Transgenic C. elegans Using Microinjection. *J Vis Exp* 833 (2008) doi:10.3791/833.

403. McIntire, S. L., Reimer, R. J., Schuske, K., Edwards, R. H. & Jorgensen, E. M. Identification and characterization of the vesicular GABA transporter. *Nature* **389**, 870–876 (1997).

404. Prior, H., Jawad, A. K., MacConnachie, L. & Beg, A. A. Highly Efficient, Rapid and Co-CRISPR-Independent Genome Editing in Caenorhabditis elegans. *G3 (Bethesda)* **7**, 3693–3698 (2017).

405. Edgley, M. L. & Riddle, D. L. LG II balancer chromosomes in Caenorhabditis elegans: mT1(II;III) and the mIn1 set of dominantly and recessively marked inversions. *Molecular genetics and genomics : MGG* **266**, 385–395 (2001).

406. Oh, S. S., Hayes, J. M., Sims-Robinson, C., Sullivan, K. A. & Feldman, E. L. The effects of anesthesia on measures of nerve conduction velocity in male C57Bl6/J mice. *Neuroscience Letters* **483**, 127–131 (2010).

407. Mcmanus, J. F. A. Histological Demonstration of Mucin after Periodic Acid. *Nature* **158**, 202–202 (1946).

408. Schuske, K., Beg, A. A. & Jorgensen, E. M. The GABA nervous system in C. elegans. *Trends Neurosci* **27**, 407–414 (2004).

409. Blumenthal, T. *Trans-splicing and operons in C. elegans*. *WormBook: The Online Review of C. elegans Biology [Internet]* (WormBook, 2018).

410. Nance, J. & Frøkjær-Jensen, C. The Caenorhabditis elegans Transgenic Toolbox. *Genetics* **212**, 959–990 (2019).

411. Tavee, J. Chapter 14 - Nerve conduction studies: Basic concepts. in *Handbook of Clinical Neurology* (eds. Levin, K. H. & Chauvel, P.) vol. 160 217–224 (Elsevier, 2019).

412. Wagner, C. E., Wheeler, K. M. & Ribbeck, K. Mucins and Their Role in Shaping the Functions of Mucus Barriers. *Annual Review of Cell and Developmental Biology* **34**, 189–215 (2018).

413. Zhang, H. *et al.* l-Threonine improves intestinal mucin synthesis and immune function of intrauterine growth-retarded weanling piglets. *Nutrition* **59**, 182–187 (2019).

414. Faure, M. *et al.* Dietary threonine restriction specifically reduces intestinal mucin synthesis in rats. *J Nutr* **135**, 486–491 (2005).

415. Jeong, S. J. *et al.* Inhibition of MUC1 biosynthesis via threonyl-tRNA synthetase suppresses pancreatic cancer cell migration. *Exp Mol Med* **50**, e424 (2018).

416. Morello, G. M. *et al.* High laboratory mouse pre-weaning mortality associated with litter overlap, advanced dam age, small and large litters. *PLoS One* **15**, e0236290 (2020).

417. Papatheodorou, I. *et al.* Expression Atlas update: from tissues to single cells. *Nucleic Acids Research* **48**, D77–D83 (2020).

418. Nagy, Z. *et al.* The Gp1ba-Cre transgenic mouse: a new model to delineate platelet and leukocyte functions. *Blood* **133**, 331–343 (2019).

419. Chen, W. *et al.* Inactivation of Plin4 downregulates Plin5 and reduces cardiac lipid accumulation in mice. *Am J Physiol Endocrinol Metab* **304**, E770-779 (2013).

420. Zhang, C. *et al.* CRISPR/Cas9-mediated genome editing reveals the synergistic effects of β-defensin family members on sperm maturation in rat epididymis. *FASEB J* **32**, 1354–1363 (2018).

421. Haridas, D. *et al.* MUC16: molecular analysis and its functional implications in benign and malignant conditions. *The FASEB Journal* **28**, 4183–4199 (2014).

422. Raymond, K. *et al.* Expression of the orphan protein Plet-1 during trichilemmal differentiation of anagen hair follicles. *J Invest Dermatol* **130**, 1500–1513 (2010).

423. Killian, G. J. Evidence for the role of oviduct secretions in sperm function, fertilization and embryo development. *Anim Reprod Sci* **82–83**, 141–153 (2004).

424. Cartoni, R. *et al.* Expression of mitofusin 2R94Q in a transgenic mouse leads to Charcot–Marie–Tooth neuropathy type 2A. *Brain* **133**, 1460–1469 (2010).

425. Detmer, S. A., Vande Velde, C., Cleveland, D. W. & Chan, D. C. Hindlimb gait defects due to motor axon loss and reduced distal muscles in a transgenic mouse model of Charcot-Marie-Tooth type 2A. *Hum Mol Genet* **17**, 367–375 (2008).

426. Bouhy, D. *et al.* Characterization of New Transgenic Mouse Models for Two Charcot-Marie-Tooth-Causing HspB1 Mutations using the Rosa26 Locus. *Journal of Neuromuscular Diseases* **3**, 183–200 (2016).

427. Strickland, A. V. *et al.* Characterization of the mitofusin 2 R94W mutation in a knock-in mouse model. *Journal of the Peripheral Nervous System* **19**, 152–164 (2014).

428. Zhou, X.-L., Ruan, Z.-R., Huang, Q., Tan, M. & Wang, E.-D. Translational fidelity maintenance preventing Ser mis-incorporation at Thr codon in protein from eukaryote. *Nucleic acids research* **41**, 302–314 (2013).

429. Chen, Y. *et al.* A threonyl-tRNA synthetase-like protein has tRNA aminoacylation and editing activities. *Nucleic acids research* **46**, 3643–3656 (2018).

430. Boutary, S. *et al.* Treating PMP22 gene duplication-related Charcot-Marie-Tooth disease: the past, the present and the future. *Transl Res* **227**, 100–111 (2021).

431. McCorquodale, D. S. *et al.* Mutation screening of mitofusin 2 in Charcot-Marie-Tooth disease type 2. *J Neurol* **258**, 1234–1239 (2011).

432. Jordanova, A. *et al.* Disrupted function and axonal distribution of mutant tyrosyl-tRNA synthetase in dominant intermediate Charcot-Marie-Tooth neuropathy. *Nat Genet* **38**, 197–202 (2006).

433. 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* **87**, 560–566 (2010).

434. Lee, D. C. *et al.* A recurrent GARS mutation causes distal hereditary motor neuropathy. *J Peripher Nerv Syst* **24**, 320–323 (2019).

435. Mnaimneh, S. *et al.* Exploration of Essential Gene Functions via Titratable Promoter Alleles. *Cell* **118**, 31–44 (2004).

436. Brachmann, R. K., Vidal, M. & Boeke, J. D. Dominant-negative p53 mutations selected in yeast hit cancer hot spots. *Proc Natl Acad Sci U S A* **93**, 4091–4095 (1996).

437. Marutani, M. *et al.* Dominant-Negative Mutations of the Tumor Suppressor p53 Relating to Early Onset of Glioblastoma Multiforme. *Cancer Res* **59**, 4765–4769 (1999).

438. Quimby, B. B. *et al.* Characterization of two mutations associated with epimerase-deficiency galactosemia, by use of a yeast expression system for human UDP-galactose-4-epimerase. *Am J Hum Genet* **61**, 590–598 (1997).

439. Spaulding, E. L. & Burgess, R. W. Accumulating Evidence for Axonal Translation in Neuronal Homeostasis. *Front. Neurosci.* **11**, (2017).

440. Rodnina, M. V. The ribosome in action: Tuning of translational efficiency and protein folding. *Protein Science* **25**, 1390–1406 (2016).

441. Zhang, G. & Ignatova, Z. Folding at the birth of the nascent chain: coordinating translation with co-translational folding. *Curr Opin Struct Biol* **21**, 25–31 (2011).

442. Siller, E., DeZwaan, D. C., Anderson, J. F., Freeman, B. C. & Barral, J. M. Slowing bacterial translation speed enhances eukaryotic protein folding efficiency. *J Mol Biol* **396**, 1310–1318 (2010).

443. Rauscher, R. & Ignatova, Z. Timing during translation matters: synonymous mutations in human pathologies influence protein folding and function. *Biochem Soc Trans* **46**, 937–944 (2018).

444. Behrens, A., Rodschinka, G. & Nedialkova, D. D. High-resolution quantitative profiling of tRNA abundance and modification status in eukaryotes by mim-tRNAseq. *Molecular Cell* **81**, 1802-1815.e7 (2021).

445. Cary, G. A. *et al.* Identification and characterization of a drug sensitive strain enables puromycin-based translational assays in Saccharomyces cerevisiae. *Yeast* **31**, 167–178 (2014).

446. Guydosh, N. R. & Green, R. Dom34 rescues ribosomes in 3' untranslated regions. *Cell* **156**, 950–962 (2014).

447. Okada, H. & Ohya, Y. Fluorescent Labeling of Yeast Cell Wall Components. *Cold Spring Harb Protoc* **2016**, pdb.prot085241 (2016).

448. Ellsworth, R. E. *et al.* The CMT2D locus: refined genetic position and construction of a bacterial clone-based physical map. *Genome Res* **9**, 568–574 (1999).

449. Dubourg, O. *et al.* The G526R glycyl-tRNA synthetase gene mutation in distal hereditary motor neuropathy type V. *Neurology* **66**, 1721–1726 (2006).

450. Matreyek, K. A. *et al.* Multiplex assessment of protein variant abundance by massively parallel sequencing. *Nat Genet* **50**, 874–882 (2018).

451. Davis, J. D. & Wypych, T. P. Cellular and functional heterogeneity of the airway epithelium. *Mucosal Immunol* 1–13 (2021) doi:10.1038/s41385-020-00370-7.

452. Tahmasebi, S., Amiri, M. & Sonenberg, N. Translational Control in Stem Cells. *Front Genet* **9**, 709 (2019).

453. Sampath, P. *et al.* A hierarchical network controls protein translation during murine embryonic stem cell self-renewal and differentiation. *Cell Stem Cell* **2**, 448–460 (2008).

454. Sanchez, C. G. *et al.* Regulation of Ribosome Biogenesis and Protein Synthesis Controls Germline Stem Cell Differentiation. *Cell Stem Cell* **18**, 276–290 (2016).

455. Signer, R. A. J., Magee, J. A., Salic, A. & Morrison, S. J. Haematopoietic stem cells require a highly regulated protein synthesis rate. *Nature* **509**, 49–54 (2014).

456. Zismanov, V. *et al.* Phosphorylation of eIF2α Is a Translational Control Mechanism Regulating Muscle Stem Cell Quiescence and Self-Renewal. *Cell Stem Cell* **18**, 79–90 (2016).

457. Blanco, S. *et al.* Stem cell function and stress response are controlled by protein synthesis. *Nature* **534**, 335–340 (2016).