Expansion of the Allelic, Locus, and Clinical Heterogeneity of Aminoacyl-tRNA Synthetase-Associated Diseases Implicates Impaired Enzyme Function as the Pathogenic Mechanism

by

Laurie Beth Griffin

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Cellular and Molecular Biology)
in the University of Michigan
2016

Doctoral Committee:

Associate Professor Anthony Antonellis, Chair
Associate Professor Catherine E. H. Keegan
Professor Andrew P. Lieberman
Assistant Professor Jordan A. Shavit
Professor Lois S. Weisman

“Science advances one funeral at a time.”

~Max Planck

2016

To my husband Matt,

For being an incredible support every day.

ACKNOWLEDGEMENTS

This work is the result of the love, support, and guidance I received from of an amazing group of people.

My utmost gratitude goes to my advisor, Tony Antonellis. Prior to arriving at Michigan, a friend insisted I rotate with Tony. I ignored his advice. Instead, I drafted a list of qualities I wanted in a mentor: a young faculty member whose lab was small, but productive and full of fun personalities, who was highly motivated to train graduate students and publish good science, whose mentorship style was hands off, yet involved and available, and whose research program involved genetics and human disease. I thought that criteria would lead to a long list of potential mentors; however, every person who I spoke with about the matter instantly came up with the same person: Tony. I am incredibly thankful that they did. Over the past few years, I have been given the privilege of designing and executing my own research projects, with the comfort of knowing Tony would reign me in or push me to expand my skills when necessary. Tony's paradigm for training graduate students is to involve them in all aspects of science. As a result, I have presented at multiple conferences and have written grants and manuscripts that have been positively received by the greater scientific community. Presentation and scientific writing skills are critical to success in science, and I am grateful to have had a mentor who helped me to develop these skills. Most importantly, Tony treats his trainees with an incredible amount of respect and understanding. He has welcomed us to his home for holidays, stayed calm when I

was panicked, and provided me with incredible support and guidance both professionally and personally. I truly cannot thank him enough for everything. I can only hope to emulate his mentorship qualities, scientific creativity, and welcoming demeanor when I am given the opportunity in the future.

I have had the honor of working with an amazing group of people over the past few years. To the former and current 'Team ARS members', Aimee Vester, Rachel Wallen, and Stephanie Oprescu, thank you for allowing me to share your incredible camaraderie, collegiality, scientific curiosity, and reagents. To the two greatest bay mates, Megan Brewer and Liz Fogarty, thank you for endless fabulous conversations and putting up with my spreading tendency. I would be hard pressed to find a more fabulous lunch buddy and friend than in Chetna Gopinath and would be at a loss if it were not for evening chats with Bill Law. I will certainly miss going for a "walk" with you all.

I am incredibly thankful for the guidance I have received from the Michigan biomedical community. I want to thank my committee members, Katy Keegan, Lois Weisman, Andy Lieberman, and Jordan Shavit for their fantastic suggestions that kept me to focused and excited about my projects. Thank you to the Meisler and Kitzman labs for their helpful discussion at lab meeting. I sincerely thank Miriam Meisler for her critical evaluation of my work, willingness to be my co-sponsor, and general mentorship over the past few years. Thank you to Jacob Kitzman for his thoughtful comments and patient guidance on my experimental endeavors. I will truly miss impromptu chats with John Moran. I am glad he was willing to learn my name and truly appreciate his scientific insights and suggestions and keen ability to pass on new tidbits about science, sports, and, most recently, primer numbers on a daily basis. Thank you to Sindhu

Ramchandren for being both my clinical co-sponsor and an amazing role model for a how to be compassionate, caring physician.

I am eternally grateful to the numerous patients and physicians who had the confidence in our lab to develop collaborations and to our scientific collaborators, particularly Ya-Ming Hou at Thomas Jefferson University, for their invaluable contributions to this work.

I would like to thank the Ron Koenig, Ellen Elkin, Hilkka Ketola, and Laurie Koivupalo for all their help and support over the past six years. These four people showed me nothing but care and support during my time at Michigan and are the reason UM's MSTP is such a wonderful program to be a part. And of course, a special thank you to Ron for suggesting Tony as a mentor.

I want to thank everyone who encouraged and helped me keep perspective and maintain a healthy work-life balance. Thank you to my MSTP, med school, and grad school friends for endless advice, entertainment, and willingness to commiserate when necessary. To Jen, Marina, and Wendy, thank you for helping me enjoy my time spent outside. To Rachel, Lisa, Rachel, Caitlin, Cam, Rob, Lilly, and Lex, thank you for always providing me with a listening ear and welcomed distractions. To my family, particularly my mom and dad, thank you for instilling in me a desire and drive to do what makes me happy and for always being a supportive, loving, and encouraging force behind me everyday. Finally, I my greatest thank you goes to my husband Matt. If it were not for his love, grounding perspective, drive to be active, and amazing cooking skills, I would not have been nearly as happy and successful in graduate school. I am forever grateful for his willingness to come to Michigan and his unwavering support my constantly evolving career aspirations.

TABLE OF CONTENTS

DEDICATION ..... **ii**

ACKNOWLEDGEMENTS ..... **iii**

LIST OF FIGURES ..... **xii**

LIST OF TABLES ..... **xiv**

ABSTRACT ..... **xvi**

CHAPTER

1. Introduction to the role of aminoacyl-tRNA synthetases in human disease **1**

Protein translation and human disease 1

Aminoacyl-tRNA synthetases serve a critical housekeeping role in cells. 3

Some ARS enzymes have secondary, non-canonical functions. 8

Aminoacyl-tRNA synthetase (ARS) genes are mutated in human disease. 8

Mitochondrial tRNA synthetase mutations in human disease 8

Cytoplasmic ARS mutations in recessive disease 13

Cytoplasmic ARS mutations in dominantly inherited peripheral neuropathy 14

Genetic tools for the identification of novel loci in dominantly inherited CMT disease 17

Functional characterization of ARS mutations 25

Methods to test for loss of ARS function 28

Biochemical studies: pyrophosphate release and aminoacylation assays 28

Yeast complementation assays 29

Protein localization studies 33
Drosophila neuron complementation assay 34
Dimerization assays 35
Methods to evaluate for dominant toxicity 37
Evaluation of dominant neuronal toxicity in C. elegans 37
Vertebrate models of ARS mutations 39
Potential Mechanisms of ARS-mediate disease 39
Significance of this thesis work 44
2. Evaluating ARS mutations for a pathogenic role in recessive disease 46
INTRODUCTION 46
MATERIALS AND METHODS 48
Patient sample collection, identification of ARS mutations, and linkage studies 48
Aminoacylation assays 49
Generation of yeast ARS expression constructs and haploid yeast strains 49
Yeast complementation assays 52
Yeast growth curves 53
Global protein assessment 53
RESULTS 54
AARS mutations identified in patients with early-onset epileptic encephalopathy. 54
K81T and R751G AARS demonstrate decreased aminoacylation activity in vitro. 56
K81T ALA1 results in reduced growth in yeast complementation assays. 56
K81T / R751G AARS fibroblasts do not show a dramatic reduction in total protein levels. 57

GARS mutations identified in patients with microcephaly. 62
R256Q GRS1 complements loss of endogenous GRS1. 62
DARS mutations identified in patients with hypomyelination. 65
DARS mutations support yeast viability. 68
FARSB mutations identified in a patient with interstitial lung and liver disease. 68
FARSB mutations support yeast viability. 71
Recessive YARS mutations identified a in child with microcephaly. 71
Yeast complementation studies suggest P167T YARS is a hypomorphic allele. 76
Discussion 80
3. Expanding the allelic heterogeneity of ARS mutations in dominant peripheral neuropathy 85
INTRODUCTION 85
MATERIALS AND METHODS 89
Patient sample collection, identification of ARS mutations, and linkage studies 89
Generation of yeast ARS expression constructs and haploid yeast strains 89
Yeast complementation assays 91
Yeast growth curves 93
RESULTS 93
Identification and functional characterization of ARS variants 93
YARS mutation identified in a family with DI-CMT 93
E196Q YARS demonstrates decreased function in yeast. 95
AARS mutation identified in a family with myeloneuropathy. 98
G102R AARS does not complement in yeast. 98

HARS mutations were identified in seven families with peripheral neuropathy..... 102
The majority of disease-associated HARS mutations demonstrate loss of function in yeast. 105
DISCUSSION 108
4. Impaired function is a common characteristic of previously identified *GARS* mutations 116
INTRODUCTION 116
MATERIALS AND METHODS 117
Computational assessment of GARS mutations 117
GARS and GRS1 expression constructs 118
Aminoacylation assays 119
Cell culture and protein localization studies 120
Population screening and segregation analysis 121
RESULTS 122
The majority of GARS mutations reduce aminoacylation activity. 124
The majority of testable GARS mutations dramatically reduce yeast cell viability. 124
The majority of GARS mutations reduce punctate localization in cultured neurons. 129
S581L GARS does not segregate with disease in newly discovered pedigrees. 131
DISCUSSION 131
5. Evaluating the pathogenic mechanism of ARS-mediated CMT disease 139
INTRODUCTION 139
MATERIALS AND METHODS 141
GRS1 expression constructs 141

Yeast complementation assays 142
Yeast trans-complementation assays 143
RESULTS 143
A subset of GRS1 alleles complement in trans to support yeast viability. 143
DISSCUSSION 150
6. Investigating the role of GARS puncta in axonal health and disease 156
INTRODUCTION 156
MATERIALS AND METHODS 159
Aminoacylation assays 160
Yeast complementation assays 161
Cell culture and protein localization studies 162
MN-1 transfections for immunoprecipitation 162
Immunoprecipitation 163
Silver staining for analysis of IP products 164
Trichloroacetic acid (TCA) precipitation 165
Mass Spectrometry and bioinformatics 165
RESULTS 166
245-248ΔETAQ GARS is a loss-of-function allele 166
Immunoprecipitation and mass spectrometry analysis of GARS proteins 168
DISCUSSION 183
7. Summary and Future Directions 187
SUMMARY OF DISSERTATION FINDINGS 187
OUTSTANDING QUESTIONS AND FUTURE DIRECTIONS 196

Defining the complete spectrum of ARS-mediated disease 196

Defining the pathomechanism of ARS-mediated dominant disease..... 199

The role of GARS puncta in disease..... 201

Treatment of ARS-mediated disease 202

CONCLUSIONS..... 203

APPENDIX..... 204

REFERENCES..... 211

LIST OF FIGURES

Figure

1.1. Cartoon of the aminoacylation reaction 4

1.2. Cartoon of the yeast complementation assay 31

1.3. Cartoon of the yeast complementation assay 42

2.1. Localization and conservation of human *AARS* mutations 55

2.2. Yeast complementation assays for recessive *AARS* alleles 60

2.3. Evaluation of protein levels in *AARS* fibroblasts 61

2.4. Localization and conservation of *GARS* mutations 63

2.5. Yeast Complementation assay for R256Q *GRS1* 66

2.6. Localization and conservation of *DARS* mutations 67

2.7. Yeast expressing disease associated *DPS1* alleles 70

2.8. Localization and conservation of *FARSB* mutations 72

2.9. Yeast complementation assay for T256M *FRS1*. 74

2.10. Localization and conservation of *YARS* mutations 75

2.11. Yeast complementation assays of P167T *TYS1* 78

2.12. Growth curve analysis of P167T *TYS1* 79

3.1. Segregation and conservation of E196K *YARS* 94

3.2. Yeast complementation assay for E196Q *TYS1* 97

3.3. Growth curves of yeast expressing E196Q *TYS1*. 99

3.4. Segregation and conservation of G102R *AARS* 100

3.5. Yeast complementation assays for G102R 103

3.6. Segregation of *HARS* mutations in families with peripheral neuropathy. 104

3.7. Conservation of *HARS* mutations 106

3.8. Yeast complementation assay of *HTS1* variants 109

3.9. Yeast complementation assay of *HTS1* variants. 110

4.1. Localization and conservation of *GARS* variants 123

4.2. Characterization of yeast expressing wild-type and mutant *GRS1* 128

4.3. Expression of wild-type and mutant GARS in mouse motor neuron-derived cells 130

4.4. S581L *GARS* does not segregate with CMT disease. 132

5.1. Confirmation of mutant *GRS1* phenotypes in yeast 146

5.2. Yeast *trans*-complementation assay of *GRS1* alleles 147

5.3. Yeast complementation assays using *GRS1 cis* alleles. 151

5.4. Yeast *trans*-complementation assays with *GRS1 cis* alleles 152

6.1. Localization and conservation of *GARS* 245-248ΔETAQ 167

6.2. Aminoacylation assays of 245-248ΔETAQ *GARS* 169

6.3. Yeast complementation studies of 245-248ΔETAQ *GRS1*. 171

6.4. Localization studies of 245-248ΔETAQ *GARS*. 172

6.5. Western blot of protein input for immunoprecipitations 173

6.6. Western blot of immunoprecipitation products 175

6.7. Silver stained gel of immunoprecipitation products 176

6.8. Images of silver stained gel samples used for mass spectrometry 177

LIST OF TABLES

Table

1.1. Human aminoacyl-tRNA synthetases 5

1.2. Human diseases associated with ARS mutations 9

1.3. Characteristics of ARS variants identified in dominantly inherited CMT disease 21

2.1. Human and Yeast orthologous ARS locus 50

2.2. Aminoacylation Kinetics of AARS Variants 58

2.3. Human *AARS* Variants Modeled in the Yeast Ortholog *ALA1* 59

2.4. Human *GARS* Variants Modeled in the Yeast Ortholog *GRS1* 64

2.5. Human *DARS* Variants Modeled in the Yeast Ortholog *DPS1* 69

2.6 Human *FARSB* Variant Modeled in the Yeast Ortholog *FRS1* 73

2.7 Human *YARS* Variant Modeled in the Yeast Ortholog *TYS1* 77

3.1. Human and yeast orthologous ARS locus 92

3.2. Human *YARS* Variants Modeled in the Yeast Ortholog *TYS1* 96

3.3. Human *AARS* Variants Modeled in the Yeast Ortholog *ALA1* 101

3.4. Human *HARS* Variants Modeled in the Yeast Ortholog *HTS1* 107

4.1. Aminoacylation kinetics of GARS protein variants 125

4.2. Human *GARS* Variants Modeled in the Yeast Ortholog *GRS1* 127

4.3. Summary of functional studies on *GARS* variants 134

5.1. Human *GARS* variants modeled in the yeast ortholog *GRS1* 145

6.1. Human 245-248ΔETAQ *GARS* Modeled in the Yeast Ortholog *GRS1* 170

6.2. Spectral counts for GARS in immunoprecipitation samples 179

6.3. Protein demonstrating increased binding to wild-type GARS..... 181

6.4. Protein demonstrating increased binding to 245-248ΔETAQ GARS 182

7.1. Human diseases associated with ARS mutations..... 188

7.2. ARS variants identified in dominantly inherited CMT disease 192

ABSTRACT

Aminoacyl-tRNA synthetases (ARSs) are ubiquitously expressed, essential enzymes that charge tRNA with cognate amino acids, a critical step in protein translation. Mutations in multiple ARS family members have been implicated in dominant and recessive diseases, which span a broad spectrum of phenotypes, including multi-organ disease and tissue specific peripheral neuropathy. The allelic heterogeneity at these loci and the mechanism by which these mutations lead to disease are currently unknown. In this work we sought to: (1) expand the locus and allelic heterogeneity for ARS-mediated recessive and dominant disease; (2) functionally characterize the effects of disease-associated mutations on ARS activity; and (3) investigate the mechanism of pathogenesis in ARS-mediated peripheral neuropathy. Our studies have identified and/or characterized 30 mutations at six ARS loci. Importantly, we identified mutations in four ARS loci not previously associated with recessive phenotypes; provided the necessary genetic evidence to implicate histidyl-tRNA synthetase in inherited peripheral neuropathy; expanded the allelic heterogeneity of glycyl-, alanyl-, and tyrosyl-tRNA synthetase mutations in peripheral neuropathy; and excluded one previously disease-associated mutation (S581L glycyl-tRNA synthetase) in disease pathogenesis. Functional studies of ARS mutations demonstrate that impaired function is a common characteristic of disease-associated ARS mutations and that there is a direct correlation between the level of impaired ARS function and the occurrence of dominant or recessive phenotypes. These data indicate that reduced function is a component of pathogenesis in both recessive and dominant ARS-associated disease. To identify the precise

loss-of-function mechanism involved in dominant ARS-mediated peripheral neuropathy, we performed studies to assess *GARS* variants for a dominant-negative effect and for altered protein interactions that may explain protein mislocalization. Initial results indicate differential interactions between wild-type and mutant GARS that have the potential to address both loss- and gain-of-function hypotheses in the field. Together, this dissertation expands the genetic and allelic heterogeneity in dominant and recessive ARS-associated disease, provides insight into the pathomechanism of *ARS*-mediated peripheral neuropathy, and has the potential to inform future attempts at therapeutic development.

CHAPTER 1

Introduction to the role of aminoacyl-tRNA synthetases in human disease

Protein translation and human disease

Cells are constantly inundated with internal and external signals. In response to these cues, cells must make alterations to their proteomes (Mathews et al. 2007). This process requires protein translation, a coordinated effort that includes mRNA processing and transport, translation initiation, and peptide synthesis. DNA is transcribed into RNA, which is processed into mRNA that can be used as a template for protein synthesis in the ribosome. Ribosomes are protein and RNA complexes that facilitate peptide bond formation in elongating peptide chains. This process occurs in three stages: initiation, elongation, and termination. In eukaryotes, initiation begins when eukaryotic initiation factors (eIFs) bind the 40S ribosomal subunit and methionine ligated $tRNA^{Met}$. This complex, termed the 43S preinitiation complex (PIC), binds the 5'cap of mRNA. Once bound, the complex scans the mRNA until the AUG start codon is recognized by $tRNA^{Met}$. eIFs subsequently hydrolyze GTP and bind the 60S ribosomal subunit to form the 80S ribosome, where elongation occurs. The ribosome binds aminoacyl-tRNAs (tRNAs ligated to cognate amino acids; AA-tRNA) and assembles the growing peptide chain based on the codon sequence in the mRNA. tRNA contains an anticodon that is complementary to the mRNA codon for its cognate amino acid. AA-tRNA binds in the acceptor site (A- site) in the ribosome. As the ligated amino acid is added to the growing peptide chain, the tRNA is transferred to the peptidyl or P-

site. After the amino acid has been incorporated into the peptide chain, the tRNA moves to the exit or E-site just prior to being released from the ribosome. Elongation factors are proteins that facilitate the movement of tRNA between sites and aid in the movement of the ribosome along the mRNA. The elongation process continues until the ribosome reaches the "stop" codon, a codon that lacks a corresponding tRNA. Peptide elongation then halts and the complex disassembles releasing the protein in the termination phase.

Dysfunction of any single component of the protein translation process can lead to human disease. Defects in components of this process have been shown to lead to a spectrum of disease that ranges from systemic to tissue specific phenotypes. For example, mutations in *EIF2AK3* result in Wolcott–Rallison syndrome, a disease of infancy that affects multiple tissues and results in diabetes mellitus, epiphyseal dysplasia, kidney and liver dysfunction, mental retardation, central hypothyroidism and dysfunction of the exocrine pancreas (Mihci et al. 2012). In contrast, mutations in *EIF2B1–5* result in a neurologic condition that includes childhood ataxia with central nervous system hypomyelination (Wilson et al. 2005). Interestingly, neurodegenerative diseases are increasingly associated with impairments in many stages of protein translation. For example, mutations in the RNA binding protein FUS leads to ALS (Vance et al. 2009), while mutations in CLP1, a kinase involved in tRNA biogenesis, results in brain malformations and peripheral nerve dysfunction (Hanada et al. 2013; Karaca et al. 2014; Schaffer et al. 2014). One family of proteins that is critical for protein translation and has been implicated in a variety of human disease, including many neurologic conditions, is the aminoacyl-tRNA synthetase (ARS) family. This work was initiated with the goal of expanding our understanding of the spectrum of phenotypes that occur in ARS-mediated disease and how mutations in ubiquitously expressed proteins lead to tissue-specific disease.

Aminoacyl-tRNA synthetases serve a critical housekeeping role in cells.

The incorporation of appropriate amino acids during the elongation of peptide chains is critical in the protein translation process. This step requires that tRNAs are properly ligated to cognate amino acids for use by the ribosome (Lodish 2008). Aminoacyl-tRNA synthetases (ARSs) are the ubiquitously expressed, essential enzymes that charge tRNA with cognate amino acids in a two-step reaction (equation shown below) (Figure 1.1) (Delarue 1995). In the first step, the ARS enzyme binds its respective amino acid and ATP and generates an amino-adenylate intermediate (AA-AMP), releasing a pyrophosphate (PPi) byproduct. In the second step, the ARS enzyme binds the appropriate tRNA, covalently ligates the amino acid to tRNA, and releases the aminoacyl-tRNA (AA-tRNA) for use during protein translation.

1. $ARS + AA + ATP \rightarrow ARS(AA\text{-}AMP) + PP_i$
2. $ARS(AA\text{-}AMP) + tRNA \rightarrow AA\text{-}tRNA + AMP + ARS$

Protein synthesis, and therefore tRNA charging, must occur in both the cytoplasm and the mitochondria (Bonnefond et al. 2005; Lodish 2008). To meet these demands, human nuclear DNA encodes 37 ARS genes that code for 36 functional enzymes: 17 genes encode 16 cytoplasmic ARS enzymes, 17 genes code for 17 mitochondrial ARS enzymes, and 3 genes encode products that function in both the cytoplasm and the mitochondria, termed 'bifunctional' ARS enzymes (Table 1.1). The nomenclature of ARS enzymes is the single-letter amino acid code followed by 'ARS'. To distinguish between cytoplasmic and mitochondrial ARS enzymes, a "2" is added to denote mitochondrial enzymes. For example, cytoplasmic histidyl-tRNA synthetase is represented by 'HARS', while mitochondrial histidyl-tRNA synthetase is signified as 'HARS2". In general, one gene encodes one ARS enzyme, which charges all the tRNAs for the cognate amino acid. There are two exceptions to this rule. First, a single gene codes for

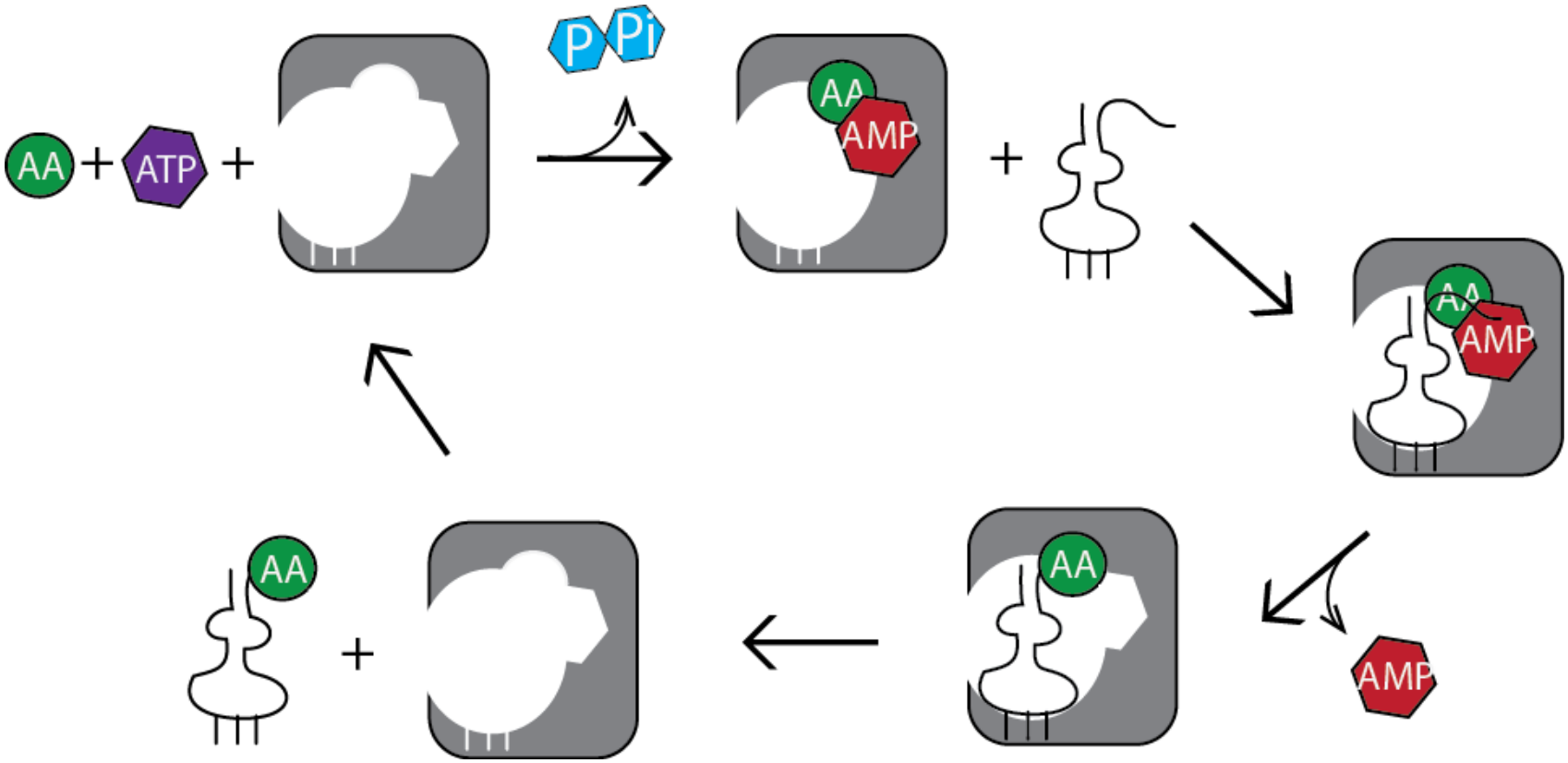


Figure 1.1. Cartoon of the aminoacylation reaction. Aminoacyl-tRNA synthetases (ARS) charge tRNA in a two-step reaction. In the first step, ARS enzymes (grey) bind the amino acid (AA; green) and adenosine triphosphate (ATP; purple) and form an aminoacyl-adenylate intermediate (green and red). During this step, a pyrophosphate is released (PPi; blue). In the second step, ARS binds the tRNA molecule (black), transfers the amino acid onto the tRNA, and releases adenosine monophosphate (AMP; red). The aminoacyl-tRNA is then released for use during protein translation and the ARS enzyme is available for subsequent rounds of aminoacylation.

Table 1.1. Human aminoacyl-tRNA synthetases

Gene Symbol	Gene Name	Class	Oligomerization Status	Cell Compartment	Multisynthetase complex
AARS	alanyl-tRNA synthetase	II	Dimer	Cytoplasm	No
CARS	cysteinyl-tRNA synthetase	I	Monomer	Cytoplasm	No
DARS	aspartyl-tRNA synthetase	II	Dimer	Cytoplasm	Yes
EPRS	glutamyl-prolyl-tRNA synthetase	I	Monomer/Dimer	Cytoplasm	Yes
FARSA	phenylalanyl-tRNA synthetase alpha	II	Heterotetramer	Cytoplasm	No
FARSB	phenylalanyl-tRNA synthetase beta	II	Heterotetramer	Cytoplasm	No
HARS	histidyl-tRNA synthetase	II	Dimer	Cytoplasm	No
IARS	isoleucyl-tRNA synthetase	I	Monomer	Cytoplasm	Yes
LARS	leucyl-tRNA synthetase	I	Monomer	Cytoplasm	Yes
MARS	methionyl-tRNA synthetase	I	Dimer	Cytoplasm	Yes
NARS	asparaginyl-tRNA synthetase	II	Dimer	Cytoplasm	No
RARS	arginyl-tRNA synthetase	I	Monomer	Cytoplasm	Yes
SARS	seryl-tRNA synthetase	II	Dimer	Cytoplasm	No
TARS	threonyl-tRNA synthetase	II	Dimer	Cytoplasm	No
VARS	valyl-tRNA synthetase	I	Monomer	Cytoplasm	No
WARS	tryptophanyl-tRNA synthetase	II	Dimer	Cytoplasm	No
YARS	tyrosyl-tRNA synthetase	I	Dimer	Cytoplasm	No
AARS2	alanyl-tRNA synthetase 2	II	Dimer	Mitochondria	No
CARS2	cysteinyl-tRNA synthetase 2	I	Monomer	Mitochondria	No
DARS2	aspartyl-tRNA synthetase 2	II	Dimer	Mitochondria	No
EARS2	glutamyl-tRNA synthetase 2	I	Monomer	Mitochondria	No
FARS2	phenylalanyl-tRNA synthetase 2	II	Heterotetramer	Mitochondria	No
HARS2	histidyl-tRNA synthetase 2	II	Heterotetramer	Mitochondria	No
IARS2	isoleucyl-tRNA synthetase 2	II	Dimer	Mitochondria	No
LARS2	leucyl-tRNA synthetase 2	I	Monomer	Mitochondria	No
MARS2	methionyl-tRNA synthetase 2	I	Monomer	Mitochondria	No
NARS2	asparaginyl-tRNA synthetase 2	II	Monomer	Mitochondria	No
PARS2	prolyl-tRNA synthetase	II	Dimer	Mitochondria	No
RARS2	arginyl-tRNA synthetase	I	Monomer	Mitochondria	No

Gene Symbol	Gene Name	Class	Oligomerization Status	Cell Compartment	Multisynthetase complex
SARS2	seryl-tRNA synthetase	II	Dimer	Mitochondria	No
TARS2	threonyl-tRNA synthetase	II	Dimer	Mitochondria	No
VARS2	valyl-tRNA synthetase	I	Monomer	Mitochondria	No
WARS2	tryptophanyl-tRNA synthetase	II	Dimer	Mitochondria	No
YARS2	tyrosyl-tRNA synthetase	I	Dimer	Mitochondria	No
GARS	glycyl-tRNA synthetase	II	Dimer	Bifunctional	No
KARS	lysyl-tRNA synthetase	II	Dimer	Bifunctional	Yes
QARS	glutaminyl-tRNA synthetase	I	Monomer	Bifunctional	Yes

glutamyl-prolyl tRNA synthetase (EPRS), a fusion ARS enzyme that charges both $tRNA^{Glu}$ and $tRNA^{Pro}$ in the cytoplasm (Berthonneau and Mirande 2000). The second exception is phenylalanyl-tRNA synthetase (FARS), the enzyme responsible for charging $tRNA^{Phe}$. FARS is a heterotetramer comprised of two subunits independently encoded by the *FARSA* and *FARSB* genes (Rodova et al. 1999).

ARS enzymes can be classified into two structural classes (Arnez and Moras 1997). Class I ARS enzymes have three distinguishing characteristics, including (i) HIGH and KMSKS sequence motifs that form the ATP binding site, (ii) a Rossman fold (parallel β-sheet nucleotide binding motif) in the catalytic domain, and (iii) ligation of the amino acid to the 2'-OH of the terminal adenosine of tRNA. With the exception of the dimeric tyrosyl-tRNA synthetase (YARS) and tryptophyl-tRNA synthetase (WARS), Class I ARS enzymes function as monomers (Arnez and Moras 1997). Alternatively, Class II ARS enzymes contain three homologous motifs (motif 1, 2, and 3). Motif 1 consists of a long α-helix linked to a β-strand and comprises part of the dimer interface, while motif 2 (two antiparallel β -strands connected by a long loop) and motif 3 (β-strand immediately followed by an α-helix) comprise the active site. Class II enzymes function as dimers or tetramers. With the exception of FARS, which utilizes the 2'-OH of tRNA for amino acid ligation, class II enzymes attach amino acids to the 3'-OH of the tRNA (Arnez and Moras 1997). In higher eukaryotes, some ARS enzymes complex in the Multisynthetase Complex (MSC) (Park et al. 2008). The MSC is known to be comprised of eight ARS proteins (EPRS, IARS, LARS, RARS, QARS, MARS, KARS, DARS) and three aminoacyl-tRNA synthetase complex-interacting multifunctional proteins (AIMP1, 2, and 3) that stabilize the complex (Bandyopadhyay and Deutscher 1971; Kerjan et al. 1994; Quevillon et al. 1999; Robinson et al. 2000). Although the remaining ARS proteins have not been purified as part of the

MSC, their association with the MSC cannot be ruled out. It is currently unknown why some ARS enzymes function in this complex, while others appear to charge tRNA independently.

Some ARS enzymes have secondary, non-canonical functions.

In this work, studies will focus on understanding the effects of disease-associated ARS mutations on the primary tRNA charging activity of ARS enzymes; however, it is important to note that some ARS enzymes have non-canonical or secondary functions (reviewed in (Smirnova et al. 2012). For example, KARS regulates transcription involved in HIV virion packaging in human cells (Cen et al. 2002), while EPRS is released from the MSC in response to γ-IFN and regulates translation of mRNA in the inflammatory response (Sampath et al. 2004). Although to date alterations in ARS secondary functions have not been implicated in disease, they may play an unknown role and therefore warrant continued investigation.

Aminoacyl-tRNA synthetase (ARS) genes are mutated in human disease.

To date, mutations in 22 ARS genes have been implicated in a spectrum of inherited human diseases (Table 1.2). More specifically, mitochondrial ARS genes have been implicated in recessive syndromes, while cytoplasmic ARS genes have been implicated in both dominant and recessive phenotypes.

Mitochondrial tRNA synthetase mutations in human disease

Mutations in nuclear-encoded mitochondrial tRNA synthetases have been identified in a wide range of recessive syndromes (Table 1.2). As is typical with mitochondrial disorders, mitochondrial ARS mutations often affect tissues with high metabolic demands such as the brain, muscle, and liver. Recessive mutations in mitochondrial aspartyl- (*DARS2*), glutamyl- (*EARS2*),

Table 1.2. Human diseases associated with ARS mutations

Gene	Locus	Location of Function	Mode of Inheritance	Disease	OMIM no.	Clinical Description	Reference
GARS	7p15	Cytoplasm & Mitochondria	Autosomal dominant	CMT2D or dSMA-V	601472 600794	Upper limb predominant, axonal motor neuropathy with variable sensory involvement	(Antonellis et al. 2006)
YARS	1p35.1	Cytoplasm	Autosomal dominant	DI-CMT	608323	Motor and sensory neuropathy with intermediate nerve conduction velocities	(Jordanova et al. 2006)
AARS	16q22	Cytoplasm	Autosomal dominant	CMT2N	613287	Motor and sensory neuropathy with variable sensorineural deafness	(Latour et al. 2010)
HARS	5q31.3	Cytoplasm	Unknown	Not applicable	142810	Single individual with sensory > motor axonal neuropathy	(Vester et al. 2012)
MARS	12q13.3	Cytoplasm	Autosomal dominant	CMT2U	616280	Motor and sensory neuropathy	(Gonzalez et al. 2013)
MARS	12q13.3	Cytoplasm	Autosomal recessive	ILLD	615486	Interstitial lung and liver disease	(Hadchouel et al. 2015)
KARS	16q23.1	Cytoplasm & Mitochondria	Autosomal recessive	RI-CMTB	613641	Single individual with intermediate CMT, developmental delay, self-abusive behavior, dysmorphic features and vestibular Schwannoma	(McLaughlin et al. 2010)
DARS	2q21.3	Cytoplasm	Autosomal recessive	HBSL	615281	Hypomyelination with brainstem and spinal cord involvement and leg spasticity	(Taft et al. 2013)
QARS	3p21.31	Cytoplasm & Mitochondria	Autosomal recessive	MSCCA	615760	Microcephaly, progressive, seizures, and cerebral and cerebellar atrophy	(Zhang et al. 2014)
RARS	5q34	Cytoplasmic	Autosomal recessive	Leuko-dystrophy	616140	Leukodystrophy, hypomyelination	(Wolf et al. 2014)

Gene	Locus	Location of Function	Mode of Inheritance	Disease	OMIM no.	Clinical Description	Reference
AARS2	6p21.1	Mitochondria	Autosomal recessive	Mitochondrial Infantile CMP	614096	Mitochondrial infantile cardiomyopathy	(Götz et al. 2011)
DARS2	1q25.1	Mitochondria	Autosomal recessive	LBSL	611105	Leukoencephalopathy with brainstem and spinal cord involvement and lactate elevation	(Scheper et al. 2007a)
EARS2	16p12.2	Mitochondria	Autosomal recessive	LTBL	614924	Infantile leukoencephalopathy involving thalamus and brainstem, high lactate, myopathy, hypotonia, seizures, retinitis pigmentosa, respiratory failure	(Steenweg et al. 2012)
FARS2	6p25.1	Mitochondria	Autosomal recessive	Alpers encephalo-pathy	614946	Neonatal encephalopathy, developmental delay, refractory seizures, lactic acidosis, liver disease	(Elo et al. 2012)
HARS2	5q31.3	Mitochondria	Autosomal recessive	Perrault Syndrome 2	614926	Sensorineural deafness (both sexes), ovarian dysgenesis (females)	(Pierce et al. 2011)
IARS2	1q41	Mitochondria	Autosomal recessive	CAGSSS; Leigh Syndrome	616007	Cataracts, growth hormone deficiency with short stature, partial sensorineural deafness, peripheral neuropathy	(Schwartzentruber et al. 2014)
LARS2	3p21.31	Mitochondria	Autosomal recessive	Perrault Syndrome 4	615300	Premature ovarian failure (POF) in females and by progressive hearing loss in both females and males	(Pierce et al. 2013)
MARS2	2q33.1	Mitochondria	Autosomal recessive	Spastic Ataxia 3	611390	Spastic ataxia, leukoencephalopathy	(Bayat et al. 2012)
RARS2	6q16.1	Mitochondria	Autosomal recessive	Spastic Ataxia 3	611390	Spastic ataxia, leukoencephalopathy	(Edvardson et al. 2007)

Gene	Locus	Location of Function	Mode of Inheritance	Disease	OMIM no.	Clinical Description	Reference
SARS2	19q13.2	Mitochondria	Autosomal recessive	HUPRA Syndrome	613845	Hyperuricemia, pulmonary hypertension, infantile renal failure, and alkalosis syndrome	(Belostotsky et al. 2011)
YARS2	12p11.21	Mitochondria	Autosomal recessive	MLASA2	613561	Myopathy, lactic acidosis, and sideroblastic anemia, muscle weakness, dysphagia, exercise intolerance	(Riley et al. 2010)
TARS2	1q21.2	Mitochondria	Autosomal recessive	COXPD21	615918	Axial hypotonia, limb hypertonia, psychomotor delay, and increased serum lactate, thin corpus callosum, cerebral spongiosis, hepatic stenosis	(Diodato et al. 2014)
VARS2	6p21.33	Mitochondria	Autosomal recessive	COXPD20	615917	Ophthalmoplegia, ptosis, ataxia, seizure	(Diodato et al. 2014; Taylor et al. 2014)

and methionyl-tRNA synthetase (*MARS2*) result in various forms of leukoencephalopathy (Scheper et al. 2007a; Bayat et al. 2012; Steenweg et al. 2012), while phenylalanyl- (*FARS2*) and arginyl-tRNA synthetase (*RARS2*) mutations result in encephalopathy and pontocerebellar hypoplasia, respectively (Pierce et al. 2011; Elo et al. 2012; Cassandrini et al. 2013). Mitochondrial valyl- (*VARS2*) and threonyl-tRNA synthetase (*TARS2*) mutations have been identified in patients with combined oxidative phosphorylation deficiency leading to hypotonia and central nervous system symptoms that include ataxia and a thin corpus callosum (Diodato et al. 2014). Mutations in mitochondrial ARS genes have also been implicated in syndromes that include peripheral nerve dysfunction. Mitochondrial histidyl- (*HARS2*) and leucyl-tRNA synthetase (*LARS2*) both cause Perrault syndrome, a condition characterized by sensorineural hearing loss in both males and females and ovarian failure in females (Pierce et al. 2011; Pierce et al. 2013). Mitochondrial isoleucyl-tRNA synthetase (*IARS2*) mutations have been identified in multiple patients with a syndromic phenotype that includes partial sensorineural deafness and peripheral neuropathy, in addition to cataracts and growth hormone deficiency (Schwartzentruber et al. 2014). In contrast to the mitochondrial mutations that cause neuronal phenotypes, mitochondrial tyrosyl-tRNA synthetase (*YARS2*) mutations result in recessive myopathy characteristic of mitochondrial respiratory chain disorders (Riley et al. 2010). Recessive mutations in mitochondrial alanyl- (*AARS2*) and seryl-tRNA synthetase (*SARS2*) lead to cardiomyopathy and pulmonary and renal failure, respectively (Belostotsky et al. 2011; Götz et al. 2011). Functional studies (see below) performed on a subset of these mutations indicate that although the mutations result in dramatically reduced aminoacylation activity, the alleles retain some level of activity sufficient to support patient viability (Edvardson et al. 2007; Scheper et al. 2007a; Belostotsky et al. 2011; Pierce et al. 2011; Bayat et al. 2012; Elo et al.

2012; Cassandrini et al. 2013). To date, effects on non-canonical ARS functions have not been reported for any of the mitochondrial ARS disease alleles.

Cytoplasmic ARS mutations in recessive disease

Similar to mitochondrial ARS mutations, cytoplasmic ARS mutations have been identified in a spectrum of recessive syndromes, including multiple with neurologic manifestations. Loss-of-function cytoplasmic methionyl-tRNA synthetase (*MARS*) mutations were identified in patients with interstitial lung and liver disease (Hadchouel et al. 2015). Lysyl-tRNA synthetase (*KARS*) mutations were identified in an individual with motor and sensory neuropathy, developmental delay, self-abusive behavior, dysmorphic features and vestibular Schwannoma, and shown to adversely affect enzyme function (McLaughlin et al. 2010). Glutamyl-tRNA synthetase (*QARS*) mutations that dramatically reduce aminoacylation activity were identified in children in two unrelated families who presented with seizures and progressive microcephaly due to cerebral and cerebellar atrophy (Zhang et al. 2014). Lastly, mutations in cytoplasmic aspartyl-tRNA synthetase (*DARS*) result in hypomyelination of the brain stem and spinal cord and leg spasticity (Taft et al. 2013), while those in arginyl-tRNA synthetase (*RARS*) result in hypomyelination in all areas of the brain (Wolf et al. 2014). Although predicted to be loss-of-function, *DARS* or *RARS* mutations have not been tested for effects on aminoacylation. As with the mitochondrial ARS alleles, only effects on the primary tRNA charging activity have been observed for disease-associated mutations; effects on non-canonical functions have not been described. Taken together, the mechanism for cytoplasmic and mitochondrial ARS-associated recessive diseases is believed to be the result of loss of tRNA charging in cells.

Cytoplasmic ARS mutations in dominantly inherited peripheral neuropathy

Mutations in five cytoplasmic ARS genes have been identified in patients with inherited peripheral neuropathy (Antonellis et al. 2003; Jordanova et al. 2006; Latour et al. 2010; McLaughlin et al. 2011; Vester et al. 2012; Gonzalez et al. 2013). Peripheral neuropathies include toxic, syndromic, or inherited neurodegenerative diseases that affect peripheral nerves (Dyck and Thomas 2005). The peripheral nervous system (PNS) includes motor and sensory neurons and myelinating Schwann cells that reside outside the brain and spinal cord, which comprise the central nervous system (CNS) (Levitan and Kaczmarek 2002). Peripheral nerves are responsible for transmitting information bi-directionally between the tissues of the body and the CNS. Dysfunction of the PNS leads to impaired movement and sensation that have a negative impact on the health and quality of life of affected individuals. It is estimated that up to 10% of the population worldwide suffer from peripheral neuropathy (Martyn and Hughes 1997).

Inherited peripheral neuropathies (IPNs) are a group of heterogeneous neurologic disorders characterized by impaired motor and sensory nerve function in the distal extremities that are caused by genetic lesions passed down through generations (Murakami et al. 1996). These genetic mutations result in IPN by disrupting the health or maintenance of one of the cell types that comprise a peripheral nerve: neurons or myelinating Schwann cells. Charcot-Marie-Tooth (CMT) disease is the most common class of inherited peripheral neuropathies, affecting up to one in 2,500 individuals worldwide (Skre 1974). CMT disease is diagnosed based on the presence of both motor and sensory nerve dysfunction and is named for the three physicians, Drs. Jean-Martin Charcot, Pierre Marie, and Howard Henry Tooth, who first described the condition in the 1880s. As CMT disease progresses, loss of nerve function ultimately leads to severe muscle wasting, reduced or absent deep-tendon reflexes, pes cavus (high foot arches), and

steppage gait (high stepping gait due to foot drop) (Dyck and Lambert 1968). Patients with CMT disease primarily suffer from loss of motor and sensory nerve function in the distal extremities, leading to impaired movement and sensation in the hands and lower legs that moves proximally as the disease progresses (Patzkó and Shy 2010). Although individuals with CMT disease typically live normal life spans, quality of life can be dramatically affected due to loss of mobility (*e.g.* reliance on walking aids) and other secondary consequences of loss of sensation (*e.g.* limb amputations) (Timmerman et al. 2013). CMT disease is extremely heterogeneous, with disease severity (*i.e.* mild to severe neuropathy) and age of onset (*i.e.* childhood through late adulthood) varying dramatically not only between subtypes of CMT disease, but also among individuals within the same family (Reilly et al. 2011).

Clinical evaluation of affected individuals is necessary to determine if CMT disease is a result of Schwann cell dysfunction or defects in the neurons themselves. Patients undergo an electrophysiologic test to determine motor nerve conduction velocity (MNCVs) of the affected nerves (Dyck and Thomas 2005). MNCVs are determined by introducing an electrical impulse to a nerve and measuring the time required for a muscle at a fixed distance away to respond to the stimulation (*i.e.* latency of response). The distance (millimeters) is then divided by the latency of response (milliseconds) to calculate the MNCV. Decreased MNCVs (<38m/s) indicate the myelin sheath is impaired due to a defect in the Schwann cell and the patient receives a diagnosis of CMT type 1 (CMT1) or demyelinating CMT disease. If performed, sural nerve biopsies from patients with CMT1 demonstrate evidence of reoccurring demyelination and attempts at remyelination by the Schwann cells that lead to loose concentric rings of myelin termed “onion bulb” formations (Dyck and Thomas 2005). Alternatively, patients can have normal MNCV (> 38m/s), but decreased nerve amplitudes. These patients are diagnosed with CMT type 2 (CMT2)

or axonal CMT disease. In these cases, the genetic lesion is presumed to affect the neuron itself and nerve biopsies can demonstrate loss of myelinated fibers, axonal degeneration, and regenerating clusters. A third class, termed intermediate CMT disease, designates cases where the MNCVs are in the intermediate range (25-45 m/s) or when there is evidence of both axonal and demyelinating disease in a single individual (Nicholson and Myers 2006).

For each of the three classes of CMT described above, autosomal dominant, recessive and X-linked disease have been identified clinically. To date over 80 genetic loci and 1000 genetic mutations have been implicated in various forms of CMT disease, with an estimated 80-100 genetic loci still undiscovered (Timmerman et al. 2014). Mutations in CMT1 have been identified in Schwann cell-specific genes involved in myelination (Lupski et al. 1991; Raeymaekers et al. 1991; Hayasaka et al. 1993), while genes associated with CMT2 disease encode proteins critical for neuronal function (Shy and Patzkó 2011), including axonal transport proteins (Millecamps and Julien 2013). As a result of molecular diagnostics, CMT disease is further sub-classified based on the genetic lesion responsible for disease. For example, patients with demyelinating CMT disease caused by a duplication of peripheral myelin protein 22 (*PMP22*) are diagnosed with CMT1A (Lupski et al. 1991; Raeymaekers et al. 1991), while individuals with decreased MNCVs and a mutation in myelin protein zero (*MPZ*) are diagnosed with CMT1B (Hayasaka et al. 1993). However, it is important to note that different genetic lesions at the same locus can result in different phenotypes. For example, *PMP22* duplications lead to CMT1A; however *PMP22* deletions lead to Hereditary Neuropathy with liability to Pressure Palsies (HNPP) (Chance et al. 1993) and individuals with *PMP22* missense mutations can have CMT1A, HNPP or CMT1E (van Paassen et al. 2014), classified based on differences in clinical presentation. Lastly, patients with the identical genetic lesion can also have different

phenotypes. For example, clinical, electrophysiologic, and histologic analysis of 61 patients with a duplication of PMP22 showed variable involvement of the corticospinal tract (indicating CNS involvement), MNCVs that ranged from 5-34m/s, and a variety of histologic changes, indicating that genetic modifiers or environmental exposures may also play a role in disease (Thomas et al. 1997).

In 2003, mutations in glycyl-tRNA synthetase (*GARS*) were implicated in IPN, marking the first report of ARS enzymes in human disease (Antonellis et al. 2003). Charcot-Marie-Tooth Disease type 2D (CMT2D) and distal spinal muscular atrophy type V (dSMA-V) are two autosomal dominant axonal IPNs that are characterized by an upper limb predominant presentation (Christodoulou et al. 1995; Ionasescu et al. 1996; Pericak-Vance et al. 1997; Sambuughin et al. 1998; Ellsworth et al. 1999). This phenotype is in contrast to the more typical lower limb predominant affects seen for IPNs (Reilly et al. 2011). Patients with CMT2D display motor and sensory defects, while patients with dSMA-V have motor deficits, but demonstrate normal sensory nerve activity. Prior to the implication of *GARS* in disease, five families were identified with CMT2D or dSMA-V (Christodoulou et al. 1995; Ionasescu et al. 1996; Pericak-Vance et al. 1997; Sambuughin et al. 1998; Ellsworth et al. 1999). By pooling the genetic information from the 5 families into a single cohort for genetic studies, Antonellis and colleagues utilized a systematic genetic approach to implicate *GARS* in upper limb predominant peripheral neuropathy (Antonellis et al. 2003; Burton et al. 2005).

Genetic tools for the identification of novel loci in dominantly inherited CMT disease

To identify a novel genetic locus in disease pathogenesis, investigators can utilize linkage studies to narrow the genomic region harboring the responsible mutation. However, investigators must first ensure the phenotype of interest is both monogenetic (*i.e.* caused by a single gene) and

measurable (*e.g.* upper limb predominant peripheral neuropathy) (Burton et al. 2005). To determine that the disease trait is monogenic, a careful family history of disease and familial structure can be ascertained to determine if the phenotype is transmitted according to Mendelian laws of dominant, recessive, or X-linked inheritance. Phenotypic information must be carefully considered for confounding factors such as disease age of onset, incomplete penetrance, variable expressivity, and phenocopy. Phenocopy refers to the presence of the trait of interest (*e.g.* peripheral neuropathy) that is due to environmental conditions (*e.g.* diabetic or toxin-induced neuropathy) rather than a genetic variant. Furthermore, to be confident that a monogenic disease and not familial clustering (*i.e.* distant family members that happen to have that same trait by chance) is being observed, the trait must occur within the family at a rate greater than expected by chance in the general population (Burton et al. 2005).

Once the mode of inheritance is determined, linkage analysis can be used to narrow the region of interest containing the pathogenic genetic lesion (Dawn Teare and Barrett 2005). Genetic linkage analysis evaluates the co-segregation of genetic loci and phenotypic traits and utilizes the likelihood that a stretch of DNA is inherited intact based on the distance between the two ends and the rate of recombination between those ends (Burton et al. 2005; Dawn Teare and Barrett 2005). To perform genetic linkage, DNA from affected and unaffected individuals is evaluated for the presence of genetic markers that exist at known locations throughout the genome. The genetic markers used for linkage studies are either microsatellite repeats (sequences that vary in length) or single nucleotide polymorphisms that have no known functional consequences. Individuals from a family are genotyped at these loci and the smallest haplotype (set of DNA variations inherited together) that is common among all affected family members is denoted as a region of interest (ROI). Linkage analysis may result in one or multiple ROIs depending on the

number of informative segregation events. Although both affected and unaffected individuals are used for linkage studies, once a ROI is established, only recombination events in affected individuals should be used to narrow the ROI to minimize confounding factors, such as phenotyping concerns, late onset of disease, and incomplete penetrance, which may improperly exclude genomic segments (Dawn Teare and Barrett 2005). The length of the ROI may vary considerably depending on the number of informative recombination events that occurred between generations. Thus, ROIs may be small and contain only one or a few genes or large and contain many genes.

A logarithm of the odds of linkage (LOD) score is often used to report linkage (MORTON 1955). LOD scores are calculated using the equation

$$\log_{10}[(1-\Theta)^{NR}*\Theta^{R} / (0.5)^{NR+R}]$$

where Θ is the recombination fraction (the probability of recombination between two genetic loci during meiosis; Θ = R / [NR + R]), 'NR" is the number of non-recombinant offspring, and 'R' is the number of recombinant offspring. A positive LOD score is suggestive of linkage, while a negative LOD score indicates an absence of linkage. A LOD score of +3 is the threshold for accepting linkage with a 5% error rate (Kruglyak and Lander 1995). However, when performing a genome wide analysis, a LOD of 3.3 must be met in order to correct for multiple testing (Kruglyak and Lander 1995). A LOD score <-2 excludes linkage, a score from -2 to +3 is inconclusive, and a score < 5 should be considered provisional. Limitations for genetic linkage include errors both in genotyping and phenotyping of individuals (misdiagnosis, phenocopy, late onset, incomplete penetrance, etc.) and the limited resolution of genotyping platforms (*i.e* the number of genetic markers assessed).

In the event that multiple or a single large ROI results from a linkage study of a single family, linkage analysis from multiple families with the same phenotype can be combined to further narrow the region. In this case, LOD scores are additive (Dawn Teare and Barrett 2005). For example, the linkage of *GARS* mutations in IPN was performed through the genetic evaluation of five unrelated families with upper limb predominant IPN, whose disease was all linked to the short arm of chromosome 7 (Christodoulou et al. 1995; Ionasescu et al. 1996; Pericak-Vance et al. 1997; Sambuughin et al. 1998; Ellsworth et al. 1999; Antonellis et al. 2003). Combining the linkage data from the five families allowed researchers to narrow the region of interest to ~980kb of chromosome 7p that contained only 11 candidate genes (Antonellis et al. 2006).

Once an ROI that contains a tractable number of candidate genes is established, mapping of the specific pathogenic variant(s) can begin. Investigators can either prioritize genes based on evidence in the literature (i.e. a candidate gene approach) or sequence all the genes in the ROI to look for variants in a single gene that may explain the common phenotype seen in affected families. In the case of *GARS*, the exons of the 11 candidate genes were sequenced in affected individuals from each family (Antonellis et al. 2006). The only gene to contain mutations in all five families was *GARS,* including one mutation that was present in two families (Table 1.3). Segregation studies were performed to determine if *GARS* mutations were present in all affected individuals. Indeed, all affected family members carried a *GARS* variant. Further analysis of ethnically matched control individuals with no evidence of peripheral neuropathy did not reveal the presence of *GARS* mutations. Taken together, the genetic evidence strongly implicated *GARS* mutations in CMT disease (Antonellis et al. 2003). Subsequent to this initial study, nine additional *GARS* mutations have been identified in patients with CMT disease (Table 1.3).

Table 1.3. Characteristics of ARS variants identified in dominantly inherited CMT disease

Gene	Variant	Conservation	Enzyme Activity	Yeast Viability	Fly Comp.	Puncta Formation	Dimerization	References
GARS	A57V	NP	NP	NP	NP	NP	NP	(Rohkamm et al. 2007)
GARS	E71G	Yeast	Normal	Viable	Reduced	Yes	NP	(Antonellis et al. 2006; Nangle et al. 2007)
GARS	L129P	Bacteria	Reduced	Reduced	No	No	Reduced	(Antonellis et al. 2006; Nangle et al. 2007)
GARS	D146N	NP	NP	NP	NP	NP	NP	(Lee et al. 2012a)
GARS	S211F	NP	NP	NP	NP	NP	NP	(Lee et al. 2012a)
GARS	G240R	Bacteria	Reduced	Viable	NP	No	Reduced	(Antonellis et al. 2006; Nangle et al. 2007)
GARS	P244L	NP	NP	NP	NP	NP	NP	(Abe and Hayasaka 2009)
GARS	I280F	NP	NP	NP	NP	NP	NP	(James et al. 2006)
GARS	H418R	NP	NP	Lethal	NP	No	Reduced	(Antonellis et al. 2006; Nangle et al. 2007)
GARS	D500N	NP	Normal	NP	NP	NP	Normal	(Del Bo et al. 2006; Nangle et al. 2007; Stum et al. 2011)
GARS	G526R	Bacteria	Reduced	Lethal	NP	Yes	Normal	(Antonellis et al. 2006; Nangle et al. 2007; Xie et al. 2007)

Gene	Variant	Conservation	Enzyme Activity	Yeast Viability	Fly Comp.	Puncta Formation	Dimerization	References
GARS	S581L	NP	Normal	NP	NP	NP	Normal	(James et al. 2006; Cader et al. 2007; Nangle et al. 2007)
GARS	G598A	NP	Reduced	Viable	NP	No	NP	(James et al. 2006; Stum et al. 2011)
YARS	G41R	Bacteria	Reduced	Lethal	NP	No	NP	(Jordanova et al. 2006)
YARS	E196K	Yeast	Reduced	Reduced	NP	No	NP	(Jordanova et al. 2006)
YARS	153-156 ΔVKQV	Yeast	NP	NP	NP	NP	NP	(Jordanova et al. 2006)
AARS	N71Y	Yeast	Reduced	Lethal	NP	NP	NP	(Lin et al. 2011; McLaughlin et al. 2011)
AARS	R329H	Zebrafish	Reduced	Lethal	NP	NP	NP	(Latour et al. 2010; McLaughlin et al. 2011)
AARS	D893N	Mouse	NP	NP	NP	NP	NP	(Zhao et al. 2012)
HARS	R137Q	Yeast	NP	Lethal	NP	NP	NP	(Vester et al. 2012)
MARS	R618C	Bacteria	NP	Lethal	NP	NP	NP	(Gonzalez et al. 2013)

Red text indicates a result consistent with a loss-of-function effect.
We considered a >90% reduction in activity as dramatically 'Reduced'.
NA – not applicable; mutation could not be modeled in yeast ortholog.
NP- not performed; analysis not included in the literature.

(Sivakumar et al. 2005; Del Bo et al. 2006; James et al. 2006; Rohkamm et al. 2007; Abe and Hayasaka 2009; Eskuri et al. 2012). For a subset of these mutations, the genetic and functional evidence is insufficient to implicate these variants in disease. Therefore, these mutations must be viewed cautiously as variants of unknown significance until appropriate studies have been performed (see chapter 4).

Since the identification of *GARS* mutations in CMT2D and dSMA-V, mutations in two additional ARS enzymes were identified in linkage studies of families with CMT disease (Jordanova et al. 2006; Latour et al. 2010; Vester et al. 2012; Gonzalez et al. 2013). The identification of tyrosyl-tRNA synthetase (*YARS*) mutations followed a similar pipeline as described above for *GARS*. Linkage analysis in two unrelated families who were affected by dominant intermediate CMT disease (DI-CMT), enabled Jordanova and colleagues to identify pathogenic mutations in *YARS* (Table 1.3) (Jordanova et al. 2006). The first family had 15 affected and 11 unaffected individuals spanning four generations that were included in the linkage analysis, while the second family had a family history of disease that spanned seven generations and including 39 affected individuals, 18 of whom were included in the study along with nine unaffected family members (Jordanova et al. 2003). The ROI calculated using linkage analysis in both families contains 37 genes and 19 predicted genes. Genes were prioritized for sequencing and ultimately a mutation in *YARS* (either G41R or E196K *YARS*) was found to segregate with disease in each family. Candidate gene screening of *YARS* in unrelated individuals with CMT revealed a third *YARS* mutation (153-156ΔVKQV). Subsequent to the genetic implication of *YARS* in CMT disease, Latour *et al*. used linkage analysis in two families with individuals in four and five generations who were affected by dominant axonal CMT disease to narrow the ROI to 130 genes on chromosome 16 (Latour et al. 2010). Using a candidate gene

approach, the investigators selected five genes for sequencing based on evidence in the literature, including alanyl-tRNA synthetase (*AARS*). The R329H *AARS* mutation was found to segregate in both families. Subsequently, this mutation was determined to have independently arisen in an unrelated Australian family, classifying R329H *AARS* as a recurrent mutation (McLaughlin et al. 2011). Two additional *AARS* mutations (N71Y and D893N) have since been identified in patients with CMT disease (Lin et al. 2011; Zhao et al. 2012). Based on the evidence thus far, three ARS genes have strong genetic evidence supporting their role in the pathogenesis of CMT disease.

Since three members of the ARS family had been implicated in CMT disease using genetic approaches, researchers hypothesized that mutations in other members of the ARS enzyme family could also result in peripheral neuropathy. To test this hypothesis, 355 patients with CMT disease and no known disease causing mutations in a CMT associated gene were screened for mutations in the 37 ARS enzymes (McLaughlin et al. 2010; Vester et al. 2012). Biallelic lysyl-tRNA synthetase (*KARS*) mutations were identified in an individual with a recessive syndrome that included axonal peripheral neuropathy (Table 1.2) (McLaughlin et al. 2010). Additionally, a mutation in histidyl-tRNA synthetase (*HARS*), R137Q *HARS,* was identified in a single individual with CMT disease (Table 1.3). Unfortunately, no additional family members were available for genetic studies to determine the inheritance pattern and segregation of the *HARS* variant; therefore, the genetic evidence was insufficient to implicate *HARS* in CMT disease (Vester et al. 2012).

As the cost of high throughput sequencing rapidly decreases, the utilization of whole exome sequencing (WES) (Ng et al. 2009; Turner et al. 2009) and whole genome sequencing (WGS) (Ng and Kirkness 2010) in the clinical setting is increasingly utilized to identify genetic variants

(Boyd 2013). WES allows for the identification of variants in coding regions of the genome (Ng et al. 2010); however, this technique is limited in that it only identifies variants in exons or at splice sites and thus does not cover non-coding or copy number variations that may be important to disease. In contrast, WGS provides coverage of both coding and non-coding regions; however, the significance of variation in non-coding regions is more difficult to interpret as the role of many of these regions is still underappreciated (Dewey et al. 2014). Currently, WES is more commonly used for identifying pathogenic variants since the majority of disease-associated variants identified to date affect coding regions. WES is a powerful tool in cases of idiopathic disease where linkage studies cannot be performed due to small pedigree size or limited access to DNA samples. Using WES, R618C methionyl-tRNA synthetase (*MARS*) was identified in a single small pedigree with late onset, incompletely penetrant axonal CMT disease (Table 1.3) (Gonzalez et al. 2013); however, until additional variants are identified, the role of *MARS* in CMT disease is still unknown.

Functional characterization of ARS mutations

Gene mutations can result in disease by a variety of mechanisms, which can be broadly classified into loss-of-function or gain-of function mechanisms (Rosenberg 2008). Loss-of-function mechanisms are caused by reduced or absent expression and/or activity of a mutant gene product, and are often the pathogenic mechanisms responsible for recessive disease. However, dominant loss-of-function mechanisms occur when gene dosage is critical to normal cell function. Loss-of-function alleles can be classified as null (protein product is non-functional), hypomorphic (protein product has a reduced, but not absent, level of activity), or antimorphic (product antagonizes the activity of the wild-type gene product). Haploinsufficiency occurs when a 50% reduction in protein function (*i.e.* one null allele in a heterozygous individual) is sufficient

to cause a disease phenotype. This mechanism occurs in diseases such as Marfan syndrome that results from heterozygosity of a loss-of-function mutation in the fibrillin gene (Dietz et al. 1991). Alternatively, a dominant-negative mechanism of disease occurs when a mutation leads to decreased protein function that is able to bind and antagonize the activity of the wild-type gene product (*i.e.* antimorphic). Antimorphs occur in cases where the protein subunits must form a homodimer to function, thus enabling the mutant subunit to interfere with the wild-type subunit's activity. For example, loss-of-function mutations in the transcriptional repressor Forkhead L2 (FOXL2) lead to premature ovarian failure by dimerizing with wild-type FOXL2 and impairing the FOXL2 heterodimer's ability to maintain transcriptional repression of genes involved in ovarian follicle differentiation (Kuo et al. 2011). Loss-of-function mutations typically affect evolutionarily conserved residues. When a high degree of allelic heterogeneity (*i.e.* multiple alleles in the same gene leading to the same phenotype) is observed, a loss-of-function mechanism should be strongly considered.

A gain-of-function mechanism should be suspected when multiple unrelated cases of a dominant disease are caused by a specific recurrent mutation. In contrast to loss-of-function mechanisms, gain-of-function mechanisms occur when a mutant protein either has increased activity (hypermorphic) or gains a novel function (*i.e.* protein aggregation or a novel binding partner; neomorphic) that leads to cellular toxicity. An example of a hypermorphic gain-of-function mutation is the activating mutation in the *c-kit* gene that results in constitutive activation of the KIT receptor tyrosine kinase, and the formation of gastrointestinal stromal tumors (GISTs) (Hirota et al. 1998). Neomorphic mutations have been reported in dihydrolipoamide dehydrogenase (DLD), a dimeric protein component of the pyruvate, α-ketoglutarate, and branched chain amino acid dehydrogenase complexes. Interestingly, DLD mutations located at

the homodimer interface impair dimer formation and allow DLD to acquire a neomorphic protease activity that leads to cardiomyopathy in patients (Brautigam et al. 2005; Babady et al. 2007). Importantly, all DLD mutations cause a loss of enzyme function that leads to a metabolic syndrome in infancy. This example illustrates an important point: loss- and gain-of-function mechanisms may not be mutually exclusive in disease pathogenesis. Loss of canonical protein function may be a prerequisite for disease pathogenesis by allowing the mutant protein product to adopt neomorphic activity resulting in a toxic gain of function. This dual mechanism has been hypothesized in some dominantly inherited neurodegenerative diseases (Winklhofer et al. 2008; Paine 2015). In Huntington disease, a toxic gain-of-function mutation causes the huntingtin protein to aggregate in neurons, leading to neurotoxicity; however, new evidence has been presented that wild-type huntingtin protects neurons, which may indicate a loss-of-function component in disease (Paine 2015). Deletion of huntingtin is not sufficient for disease, indicating that loss-of-function alone is not the pathomechanism. Thus the loss of normal huntingtin protein function and the gain of the toxic aggregates may both be required for Huntington's disease to manifest (Paine 2015).

While there is evidence that ARS mutations associated with both dominant and recessive disease demonstrate loss-of-function (Table 1.3), not all mutations have been functionally characterized (Taft et al. 2013; Wallen and Antonellis 2013). Thus, much is currently unknown about the mechanism of pathogenesis for ARS-mediated disease. Functional studies evaluating the consequences of ARS mutations on enzyme activity is a critical step toward implicating ARS variants in both dominant and recessive disease and understanding the mechanism by which ARS mutations lead to inherited disease.

Methods to test for loss of ARS function

Biochemical studies: pyrophosphate release and aminoacylation assays

Since the primary function of ARS enzymes is to charge tRNA, it is critical to understand how mutations impact aminoacylation activity. Pyrophosphate and aminoacylation assays are commonly used to determine the ability of mutant ARS enzymes to complete steps necessary for successful aminoacylation *in vitro* (Francklyn et al. 2008). Pyrophosphate release assays are colorimetric assays that detect pyrophosphate release upon the completion of first step of the aminoacylation reaction (Kiga et al. 2002). In this assay, varying concentrations of purified recombinant ARS enzyme is incubated with tRNA, ATP, cognate amino acid, and pyrophosphatase. A colorimetric indicator that reacts to the presence of free phosphates that are produced by the cleavage of pyrophosphates (PPi) released during amino acid activation by pyrophosphatase is added at set intervals. Absorbance is then used to calculate the phosphate release as an indicator of the efficiency of formation of the amino-adenylate intermediate. This assay has been used to study *YARS* and *HARS2* mutations (Jordanova et al. 2006; Pierce et al. 2011), but is limited in that it only assess the amino acid activation step of the aminoacylation reaction.

To assess an ARS enzyme's ability to complete both steps of the tRNA charging reaction, *in vitro* aminoacylation assays are employed (Francklyn et al. 2008). Aminoacylation assays calculate the steady state kinetics of the aminoacylation reaction by monitoring the formation of amino acid-ligated tRNA. Recombinant purified ARS enzyme is incubated with tRNA, ATP, and varying concentrations of radiolabeled ^{3}H or ^{14}C amino acid under optimal temperature and buffer conditions for the ARS enzyme of interest. Aliquots of the reaction mixture are taken at various time points and spotted on Whatmann filter paper that binds the tRNA. The filter paper

contains trichloroacetic acid (TCA) to quench the reaction and precipitate the aminoacylated tRNA. After washing the filter paper to remove any unincorporated amino acids, the filter paper is dried and a liquid scintillation counter is used to assess for radioactivity levels. The amount of radioactivity retained on filter pads is corrected for quenching effects to determine the amount of synthesized amino acid-ligated tRNA. Steady-state kinetics is then determined by fitting the initial rate of aminoacylation as a function of tRNA concentration to the Michaelis–Menten equation (Schreier and Schimmel 1972). Aminoacylation assays are more commonly used than pyrophosphate release assays to test the impact of ARS mutations on enzyme function since it assesses for completion of the entire reaction. The consequences of mutations in *GARS*, *AARS*, *KARS*, *QARS*, and *FARS2* have been evaluated using the aminoacylation assay (Antonellis et al. 2006; Nangle et al. 2007; McLaughlin et al. 2010; McLaughlin et al. 2011; Elo et al. 2012; Zhang et al. 2014). The predominant limitation of these assays is that they are performed outside of the context of a cell. Thus any post-translational modification and/or stabilizing cellular factors that may have an impact on the function of the ARS enzyme *in vivo* are absent. Nonetheless, aminoacylation assays provide important insights into the effects of mutations on the primary tRNA charging function of ARS enzymes.

Yeast complementation assays

To study ARS mutations *in vivo*, investigators have adopted *Saccharomyces cerevisiae*, baker's yeast, as a model organism. Yeast are small, single cell, eukaryotic organisms that are genetically tractable (Xiao 2006). Importantly, knockout allele strains exist for the majority of yeast genes and haploid yeast can be easily generated, making it an excellent model to assess the effect of a single allele *in vivo* (Winzeler et al. 1999). To evaluate the effects of ARS mutations, mutant ARS alleles are introduced into haploid yeast strains lacking the ARS locus of interest to see if

the alleles retain enough activity to support yeast cell viability. To generate haploid yeast strains, diploid yeast that have one orthologous ARS allele disrupted by a selective marker, either *HIS3* or *KanMX* (Winzeler et al. 1999; Turner et al. 2000), are induced to undergo sporulation. Since ARS enzymes are essential, a *URA3*-bearing plasmid containing a wild-type copy of the yeast ARS locus (including the promoter and 3' downstream sequence) is introduced into the diploid yeast strain prior to sporulation. During sporulation, a yeast cell produces four meiotic products that each contains one allele per gene. These meiotic spores are dissected and selected on appropriate media to isolate individual haploid yeast that have the endogenous ARS locus deleted and viability maintained by the ARS locus on the *URA3* maintenance vector. Human mutations can then be modeled in the yeast ortholog on a *LEU2*-bearing experimental vector and introduced into the haploid yeast strain (Figure 1.2). Yeast carrying both the *URA3* and *LEU2* vectors are selected for on media lacking uracil and leucine and subsequently plated on 5-fluoroorotic acid (5-FOA) media and incubated for 2-3 days. 5-FOA is toxic to yeast expressing *URA3* and thus selects for yeast that have spontaneously lost the *URA3*-bearing maintenance vector (Boeke et al. 1987). Therefore, only yeast with a functional copy of the ARS gene on the *LEU2* vector will survive on 5-FOA media. Visual inspection of growth on solid 5-FOA medium or OD_{600} absorbance readings in liquid 5-FOA medium are used to determine if the experimental (wild-type or mutant) ARS allele can complement loss of the endogenous allele. Wild-type and other functional ARS alleles will support robust yeast cell growth while functionally null alleles are unable to sustain yeast viability, similar to an empty vector control. Some mutations will support an intermediate level of growth compared to the wild-type ARS and the empty vector; these alleles would be considered hypomorphic. A slight modification to this approach was used by Jordanova *et al*. to test *YARS* mutations (Jordanova et al. 2006). In their study, the *LEU2*

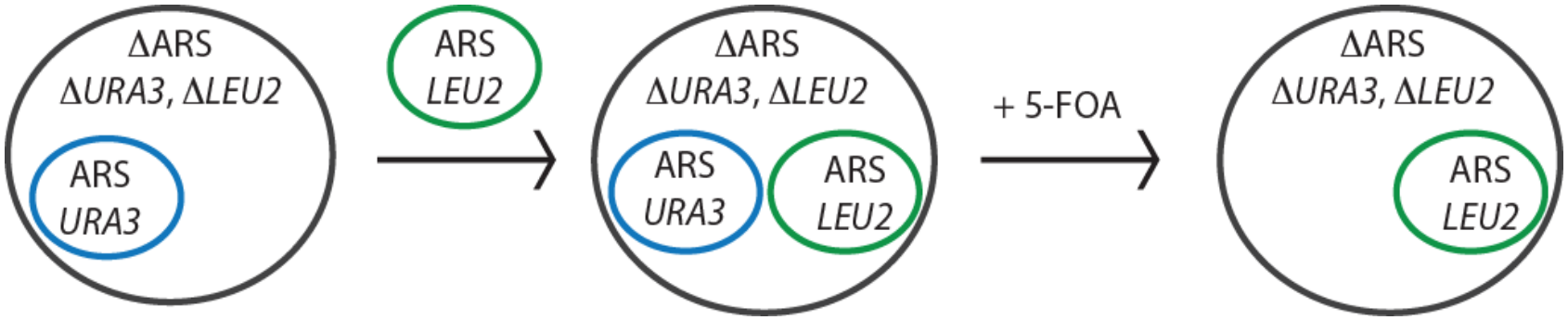


Figure 1.2. Cartoon of the yeast complementation assay. A haploid yeast strain [that has the endogenous ARS locus deleted and bearing a *URA3* maintenance vector containing the wild-type ARS locus (blue)] is transformed with a *LEU2*-bearing pRS315 vector containing wild-type or mutant ARS or no insert (green). Liquid cultures for each strain are grown to saturation in media lacking uracil and leucine to select for the presence of the two vectors. Yeast are spotted on selective solid growth medium containing 5-FOA to select for the spontaneous loss of the *URA3* maintenance vector (Boeke et al. 1987). Visual inspection of growth is used determine whether the mutant ARS allele complements loss of endogenous ARS.

vector containing wild-type or mutant yeast *TYS1*, the yeast ortholog of *YARS*, was introduced into the diploid strain that lacked a *URA3* maintenance vector. Tetrad dissections were performed to determine if the experimental allele could sustain yeast viability in the haploid yeast lacking the endogenous ARS allele.

It is important to note that the yeast complementation assay does not directly test for enzyme activity; instead it tests for the ability of the enzyme to complement all aspects of ARS function *in vivo*. However, there are limitations to this assay: primarily the use of the orthologous yeast gene to model human mutations. Mutations that affect residues not conserved between human and yeast or in domains that are absent in the yeast ortholog (*i.e.* WHEP domain, a domain that is a found in five human ARS enzymes, but is absent in all *S. cerevisiae* ARS proteins (Ray and Fox 2014)) cannot be modeled. Additionally, some mutations may affect human protein function, but may not have a severe enough effect on the yeast enzyme to cause a reduction in cell growth. To date, all mutations that demonstrate decreased yeast cell growth also show impaired function in aminoacylation assays, though the converse is not true (Antonellis et al. 2006; Nangle et al. 2007; McLaughlin et al. 2010). This may be due to post-translational modifications or stabilization of ARS by cellular components in yeast that are not present in the *in vitro* aminoacylation assay. It has been shown that some human ARS genes, such as *YARS* (Jordanova et al. 2006), can complement loss of the yeast ortholog. Therefore, modifications to the standard yeast complementation assays that allow use of the human protein will help address some of these limitations. Lastly, yeast are small, unicellular organisms that are unable to recapitulate the demands of long, terminally differentiated cells (i.e. neurons). Since dominant ARS-associated disease only affects neurons, impairment of ARS function that may be significant in neurons may be too subtle to detect an effect in yeast. Nonetheless, yeast

complementation assays are currently the best *in vivo* screen for pathogenicity of ARS mutations (Motley et al. 2015).

Protein localization studies

Many proteins must localize to specific sites in the cell for proper functioning. Disruption of this localization pattern can result in impaired protein function. For example, loss-of-function mutations in the *CFTR* gene, a chloride channel, lead to cystic fibrosis (Kerem et al. 1989). A subset of these mutations have no effect on the proper conduction of the channel in functional assays, but instead cause disease by impairing the localization of CFTR to the membrane where its function is required (Smit et al. 1995; Fanen et al. 2014). Thus, it is important to consider whether ARS enzymes have site-specific localization requirements necessary for proper ARS activity. As previously discussed, a subset of ARS enzymes localize to the MSC while other ARS enzymes appear to function independent of the complex (Park et al. 2008). GARS and YARS proteins are not known to associate with the MCS and have both been shown to form punctate structures in cells (Antonellis et al. 2006; Jordanova et al. 2006). Specifically, endogenous wild-type GARS localize to 'puncta' within the soma and axons in the human peripheral nerve and in a variety of non-neuronal and neuronal cultured cells (Antonellis et al. 2006). YARS was shown to localize in a "tear-drop" pattern that was specific to growth cones and axonal termini in cells that form neurite projections (Jordanova et al. 2006). In contrast to the ubiquitous presence of GARS puncta in all cells, YARS tear-drops were not observed in non-neuronal cells. Specifically, COS-7 cells, which are derived from monkey kidney tissue, demonstrated a homogenous distribution of YARS (Jordanova et al. 2006). To test the effect of mutations on GARS and YARS localization, the open reading frame of each human enzyme was tagged with enhanced green fluorescent protein (EGFP; C-terminal tag for GARS; N-terminal

tag for YARS) and wild-type and mutant constructs were overexpressed in neuronal cell lines (Antonellis et al. 2006; Jordanova et al. 2006). When modeled *in vitro*, wild-type GARS-EGFP forms puncta consistent with endogenous GARS; however, three of the five *GARS* mutations impair the punctate localization in cultured neurons (Table 1.3) (Antonellis et al. 2006), suggesting altered localization in axons may be a component of CMT2D and dSMA-V disease pathogenesis. The remaining two mutant GARS proteins formed puncta in cell bodies and neurite projections, consistent with previous observations of wild-type GARS localization to axons. Similar to GARS, overexpressed wild-type YARS-EGFP forms axonal tear-drop structures consistent with endogenous YARS, and *YARS* mutations tested result in loss of these structures (Table 1.3) (Jordanova et al. 2006). It is important to note that the composition, function, and relevance to disease are currently unknown for GARS and YARS puncta (see chapter 6). However, these studies imply that protein localization may be an important loss-of-function component for some *GARS* and *YARS* CMT-associated mutations.

Not all disease-associated ARS enzymes form punctate structures in cells. Localization studies of AARS demonstrated a diffuse pattern of localization for the endogenous protein (McLaughlin et al. 2011); furthermore overexpression of *AARS* revealed a similar localization pattern for both wild-type and mutant AARS (McLaughlin et al. 2011). Still, to determine if abnormal localization may be playing a role in disease, it will be important to assess newly identified *GARS* and *YARS* mutations and any newly implicated ARS enzymes for localization effects in neuronal cells.

Drosophila neuron complementation assay

Although aminoacylation, yeast complementation, and localization assays provide valuable information regarding ARS function *in vitro* and *in vivo*; however, many ARS mutations lead to

tissue-specific disease, which is not be reflected in these studies. For example, E71G *GARS* leads to peripheral neuropathy, but has not demonstrated abnormal function in aminoacylation, yeast complementation, or localization studies (Antonellis et al. 2003; Antonellis et al. 2006). Therefore, it is critical to evaluate mutations, such as E71G, in the context of a neuron to understand its functional consequences. During a forward genetic screen for mutations that result in dendritic and axonal morphology defects of olfactory projection neurons, a mutation was identified in Drosophila *gars*, P98L *gars* (Chihara et al. 2007). These neurons each innervate a single glomerulus in the Drosophila brain in a stereotypical arborization pattern. $gars^{P98L/P98L}$ flies showed normal growth and guidance of dendritic and axonal stalks but failed to properly arborize the distal axons and dendrites. To elucidate if the axonal arborization defect was a result of *gars* loss-of-function, investigators generated a null *gars* allele ($gars^{E34}$). Flies homozygous for $gars^{E34}$ in olfactory neurons demonstrated the same arborization defects. To test whether human wild-type, E71G or L129P *GARS* demonstrated loss-of-function characteristics in their Drosophila model, investigators performed rescue experiments with each allele (Table 1.3). Wild-type *GARS* rescued the defect, but L129P *GARS* did not complement. E71G *GARS* partially rescued the arborization defect, but not to the same extent as wild-type *GARS*. This study provided the first functional data indicating that E71G is a loss-of-function *GARS* allele and was the first to evaluate ARS mutations in the context of axons *in vivo*. Thus, evaluating ARS mutations in a fly complementation model may provide the resolution necessary to detect subtle loss of function characteristics of ARS alleles that are only detrimental in the context of neurons.

Dimerization assays

A subset of ARS enzymes function as dimers. In fact, with the exception of *MARS,* whose role in CMT disease is still unclear (Gonzalez et al. 2013), all ARS enzymes identified in dominantly-

inherited CMT disease function as dimers (Antonellis et al. 2003; Jordanova et al. 2006; Latour et al. 2010; Vester et al. 2012). Loss of dimerization may result in loss of enzyme function. Alternatively, ARS mutants that have impaired function in other assays may maintain the ability to dimerize with wild-type ARS, potentially leading to an antagonization of wild-type activity (*i.e.* dominant negative pathogenic mechanism). Therefore, it is important to determine if disease-associated mutations affect dimerization as this may be a critical component of ARS-mediated disease. To test for dimerization, three assays have been used. First, mutant ARS enzymes can be tagged and overexpressed in cells and immunoprecipitation (IP) of the tagged mutant protein can be performed. Western blots of IP products can then be used to detect tagged and endogenous ARS to determine if wild-type and mutant proteins dimerize (Nangle et al. 2007; Xie et al. 2007). Alternatively, analytical centrifugation can be used to determine if a mutant protein exists primarily as a monomer or dimer in solution. For analytical ultracentrifugation experiments, high velocity centrifugation is used to sediment proteins (Howlett et al. 2006). Dimers and monomers will sediment at different rates due to differences in mass, thus allowing investigators to quantify the difference between monomeric and dimeric interactions. These studies have been used to evaluate the dimerization of mutant GARS proteins (Table 1.3) and have shown that *GARS* mutations variably affect dimerization; some *GARS* mutations increase dimerization (ex. G598A *GARS*) while others nearly ablate dimeric interactions (ex. G240R *GARS*) (Table 1.3) (Nangle et al. 2007; Xie et al. 2007). As the mechanism of ARS-mediated CMT is further investigated, increasing or inhibiting wild-type /mutant dimerization may prove to be an important consideration in therapeutic development.

Methods to evaluate for dominant toxicity

Evaluation of dominant neuronal toxicity in C. elegans

Genetic studies have clearly defined a role for ARS mutations in dominantly inherited CMT disease (Antonellis et al. 2006; Jordanova et al. 2006; Latour et al. 2010). Thus, when a variant of unknown significance is identified in a case where segregation studies cannot be performed (*e.g.* sporadic CMT disease), assessing the mutation for dominant neurotoxicity may be an important step in implicating the variant in disease. *Caenorhabditis elegans* (*C. elegans*) is a transparent nematode (roundworm) that provides an excellent model system for testing for ARS-mediated toxicity. In contrast to biochemical and yeast studies, *C. elegans* can be used to understand a variant's effect in the context of an axon, a particularly important consideration in understanding the axonal pathology of ARS-mediated peripheral neuropathy. Furthermore, *C. elegans* have characteristics that make them particularly useful for studying neuron biology (Yook 2005). First, *C. elegans* have a short lifespan and can be synchronized to specific developmental stages that allows for efficient assessment of variants associated with late-onset disease in large cohorts of animals. Second, *C. elegans* have a well-characterized non-myelinated nervous system that makes axonal defects easy to visualize. Specifically, the GABAergic nervous system consists of 26 GABAergic neurons, including 19 inhibitory (D-type) GABA motor neurons that project commissural axons from the nerve cell bodies in the ventral nerve cord to the dorsal nerve cord (Schuske et al. 2004). These neurons are responsible for innervating the body wall muscles of the worm and are critical for locomotion. Dysfunction of these neurons can be assessed using a thrash assay, which quantifies the bending locomotion of individual animals and can be used as an indicator of mutant protein toxicity on neuromuscular function in transgenic animals (Vester et al. 2012). Specifically, animals with abnormal D-type GABA neuron function will fail to bend and the number of bends over time can be quantified and

compared to control worms. Finally, *C. elegans* are genetically tractable and transgenes can be expressed in single or multiple copy allowing for overexpression and gene dosage studies.

A *C. elegans* model was developed to study the *HARS* mutation R137Q (Vester et al. 2012). In this model, wild-type or mutant cDNA sequences are cloned downstream of the glutamic acid decarboxylase gene (a GABA biosynthetic enzyme; *unc-25*) promoter and are injected into the *C. elegans oxIs-12* strain that expresses GFP under the GABA vesicular transporter gene (*unc-47*) promoter to illuminate the GABA neurons for microscopy (Hobert 2002). Offspring from stably transmitting lines are synchronized to the fourth larval stage and aged to 1 or 4 day old adults to assess for developmental versus degenerative effects, respectively, of mutant ARS expression. Confocal imaging can be used to visualize malformations of the ventral or dorsal nerve cords or abnormal branching of axon projections. Overexpression of R137Q *HARS*, but not wild-type *HARS*, resulted in abnormal axonal morphology, including abnormal branching and axonal blebbing, failure of axonal commissures to extend to the dorsal nerve cord, and large dorsal nerve cord gaps (Vester et al. 2012). These abnormalities of the GABA nervous system progressed as animals aged. When assessed for neuromuscular function using thrash assays, impaired motor function was evident in older (4 day old) adult animals, but not in larval or 1-day old adults, indicating a worsening of neurotoxicity with age. Decreased locomotion directly correlated with the increase in abnormal axon morphology, suggesting a progressive loss of motor neuron function and muscular innervation, consistent with the late onset, progressive peripheral neuropathy phenotype associated with human ARS mutations. This *C. elegans* model demonstrated that R137Q *HARS* was dominantly neurotoxic. A similar approach could be used to assess additional ARS variants of unknown significance for dominant neurotoxicity, providing evidence for their role in dominantly inherited CMT disease. In addition, this system can be used

for mechanistic studies. For example, this assay could be employed to compare variants with variable activity of levels or dimerization ability to determine which characteristics are critical components of ARS-mediated disease.

Vertebrate models of ARS mutations

While invertebrate models have been valuable in studying aspects of ARS mutations, vertebrate models offer a unique resource to understanding the pathogenesis of ARS-mediated disease. To date there are four mouse models of ARS mutations. The "*sticky*" mouse is homozygous for A734E *Aars* and suffers from progressive ataxia (Sarna and Hawkes 2011). A743E *Aars* affects the editing activity of AARS, resulting in the mischarging of $tRNA^{Ala}$, misincorporation of amino acids into proteins, and the accumulation of misfolded proteins in neurons. This accumulation ultimately leads to neurotoxicity, Purkinje cell loss, and ataxia. Three mouse models contain mutations in *Gars*. Two models carry missense alleles, C201R and P234KY, that lead to mild and severe peripheral neuropathy, respectively, in heterozygous animals (Seburn et al. 2006; Achilli et al. 2009). Homozygosity for either allele is lethal by two weeks of age (Motley et al. 2011). The third *Gars* mutation is a gene-trap allele (Seburn et al. 2006). Unlike mice heterozygous for the missense mutations, mice heterozygous for the gene-trap allele have a wild-type phenotype. Researchers have used these models to test hypotheses pertaining to both loss- and gain-of-function mechanisms in *GARS*-mediated CMT disease (Motley et al. 2011; He et al. 2015).

Potential Mechanisms of ARS-mediated disease

As discussed above, there is a growing body of evidence suggesting that loss of ARS function plays an important role in both recessive and dominant disease. In this section I will discuss the hypotheses that have been proposed to explain the mechanism of ARS-mediated disease and the

current evidence in support of each hypothesis.

For recessive disorders, the loss of tRNA charging activity is believed to be the pathogenic mechanism. In fact, all mutations functionally characterized to date result in a dramatic reduction of enzyme activity (Scheper et al. 2007a; McLaughlin et al. 2010; Riley et al. 2010; Belostotsky et al. 2011; Götz et al. 2011; Pierce et al. 2011; Bayat et al. 2012; Elo et al. 2012; Steenweg et al. 2012; Cassandrini et al. 2013; Pierce et al. 2013; Diodato et al. 2014; Wolf et al. 2014; Zhang et al. 2014; Hadchouel et al. 2015). Mutant ARS enzymes do retain a basal level of activity to sustain life in these individuals, but it appears inadequate for normal growth and development. It remains unclear why recessive mutations in ARS enzymes lead to a wide spectrum of disease (*i.e.* microcephaly to interstitial lung and liver disease) phenotypes and why some tissues (*e.g.* brain and muscle) are more often affected than others. One could hypothesize that the demands on protein synthesis are greater in some tissues (such as muscle) or that some tissues are less resistant to defects in protein synthesis (such as neuronal tissue). Further investigation is necessary to discern the mechanisms underlying each disease and to identify other ARS loci that may be involved in recessive phenotypes. These studies will be paramount to understanding ARS biology and to informing therapeutic development. Although the details of the loss-of-function mechanisms are still being elucidated, efforts to improve tRNA charging, such as increasing amino acid concentrations or identification of pharmacologic agents that increase aminoacylation, would be logical initial avenues for therapeutic exploration in ARS-mediated recessive disease.

The implication of multiple ARS family members in a similar phenotype strongly suggests that a common pathogenic mechanism is involved in ARS-mediated CMT disease; however, the precise mechanism by which ARS mutations lead to dominant disease remains an open question

in the field. To date, there are five major hypotheses regarding the mechanism (Wallen and Antonellis 2013): (1) haploinsufficiency due to impaired tRNA charging in the cytoplasm or axon projections leads to decreased protein synthesis; (2) protein mislocalization leads to impaired local protein translation; (3) a dominant-negative effect results in a severe reduction in tRNA charging (>50%) sufficient for disease; (4) increased levels of uncharged tRNA results in a reduction in global protein synthesis; and (5) a toxic gain-of-function (*i.e.* abnormal protein or RNA binding) leads to neurotoxicity (Figure 1.3).

Multiple observations support a loss-of-function mechanism in ARS-mediated CMT disease. First, ARS mutations associated with CMT disease affect conserved residues (Table 1.3) and do not cluster in a specific functional domain. Secondly, 12 CMT-associated ARS alleles have been functionally characterized to date (Table 1.3) revealing that: (a) nine of 11 result in impaired tRNA charging in biochemical assays (Antonellis et al. 2006; Jordanova et al. 2006; Nangle et al. 2007; Xie et al. 2007; McLaughlin et al. 2011); (b) seven of 10 impair cellular growth when modeled in yeast (Antonellis et al. 2006; Jordanova et al. 2006; McLaughlin et al. 2011; Stum et al. 2011; Vester et al. 2012); (c) five of seven impair puncta formation within cultured neurons (Antonellis et al. 2006; Jordanova et al. 2006); and (d) two of two mutations tested are unable to complement deletion of the fly ortholog (Chihara et al. 2007). Together, 10 of 12 characterized ARS mutations have at least one loss-of-function characteristic, indicating that impaired tRNA

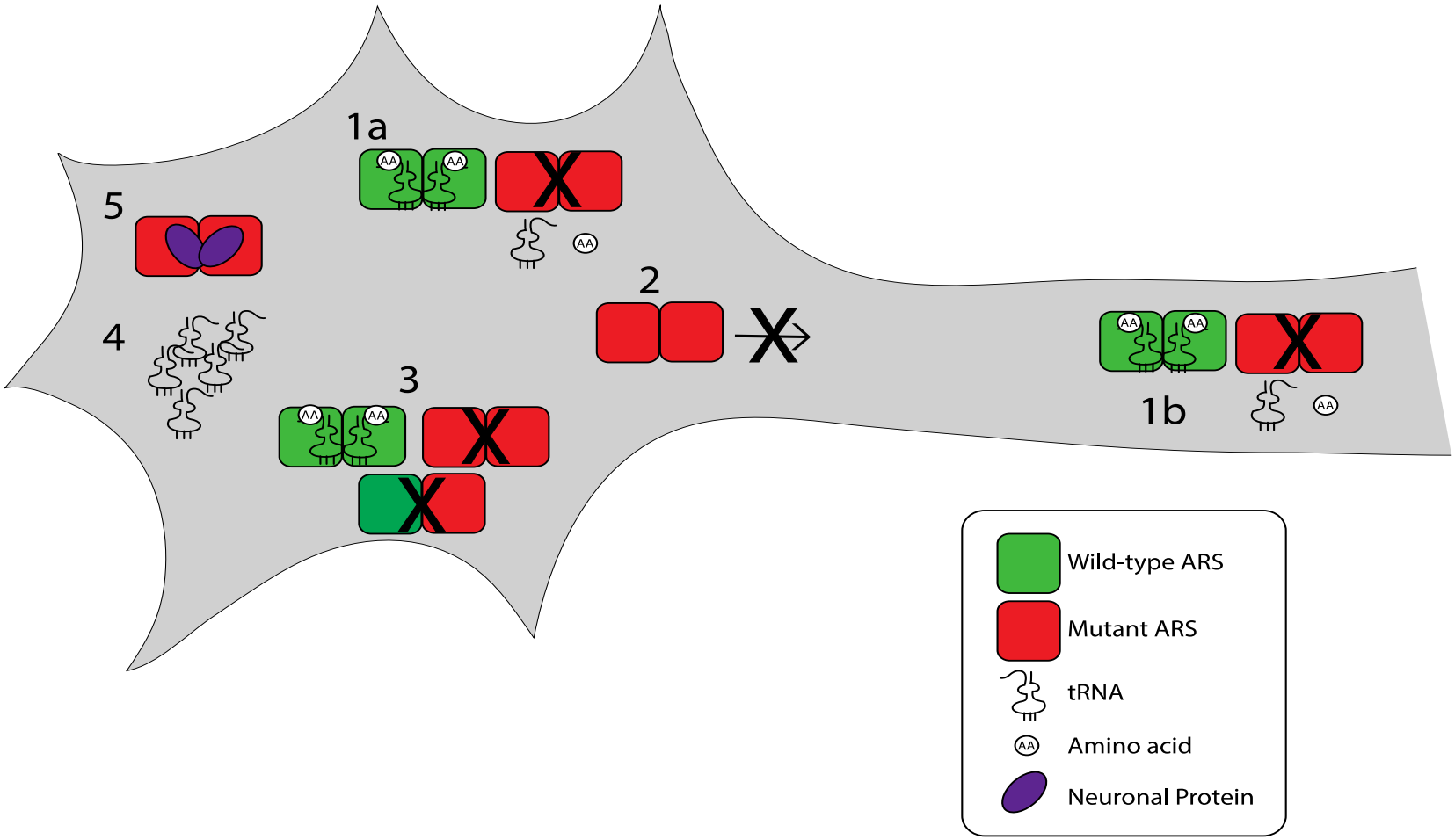


Figure 1.3. Potential pathogenic mechanisms of ARS-mediated CMT disease. There are five proposed mechanisms by which ARS mutations lead to CMT disease. Loss-of-function ARS enzymes could result in a 50% reduction in tRNA charging in the neuron cell body (1a) or in the axon (1b). Mutant ARS protein may mislocalize in cells (2). Mutant ARS protein may bind wild-type ARS and antagonize aminoacylation activity, resulting in a severe reduction in tRNA charging (dominant-negative effect; 3). Uncharged tRNA may lead to a global decreased in protein synthesis (4). Lastly, mutant ARS may acquire a novel binding partner (proteins or RNA) that results in neurotoxicity (5) (adapted from (Wallen and Antonellis 2013).

charging is a component of ARS-mediated CMT disease. However, the precise loss-of-function mechanism by which impaired ARS function leads to CMT disease has not been elucidated. All dominantly inherited ARS alleles are missense mutations (Wallen and Antonellis 2013); frameshift and nonsense alleles have not been identified. As previously discussed, mice heterozygous for *Gars* loss-of-function missense mutations demonstrate axonal degeneration (Seburn et al. 2006; Achilli et al. 2009), while mice heterozygous for the null gene-trap *Gars* allele are phenotypically normal (Seburn et al. 2006), suggesting that haploinsufficiency is not the pathological mechanism. The apparent discrepancy between impaired ARS function and dominant CMT disease may be explained by a dominant-negative effect. A dominant-negative mechanism is supported by data showing that: (1) ARS enzymes implicated in CMT disease function as oligomers (Antonellis et al. 2006; Jordanova et al. 2006; Latour et al. 2010; Vester et al. 2012); (2) ARS missense mutations tested to date result in stable proteins; (3) ARS mutations demonstrate loss-of-function characteristics (Wallen and Antonellis 2013); (4) *GARS* and *YARS* mutations allow wild-type:mutant dimer formation (Jordanova et al. 2006; Nangle et al. 2007; Qin et al. 2014); and (5) *YARS* mutations co-expressed with wild-type YARS in yeast lead to a negative effect on yeast growth (Jordanova et al. 2006).

Additional studies in the literature have addressed the remaining hypotheses. A reduction in ARS enzyme function would be predicted to lead to a decrease in charged tRNA. It is known that at times of nutrient deprivation uncharged $tRNA^{His}$ can lead to activation of general control non-repressible 2 kinase (GCN2) (Dong et al. 2000). GCN2 phosphorylates eukaryotic initiation factor 2a (eIF2a) to inhibit global protein translation. Although increased eIF2a phosphorylation has not been reported, a recent study of a *Gars* fly model showed a decrease in global protein translation in neurons expressing mutant *Gars* (Niehues et al. 2015). However, overexpression of

wild-type *Gars* in these flies did not rescue the phenotype, suggesting to the authors that a toxic gain of function, not a decrease in tRNA charging, was responsible for the global protein translation defect. Studies on mouse *Gars* mutations that result in a peripheral neuropathy phenotype were also not rescued by overexpression of wild-type mouse *GARS* (Motley et al. 2011). Finally, evidence in the literature suggests that *GARS* mutations may alter the structure of GARS proteins and that these changes could result in novel protein interactions for mutant GARS (He et al. 2015). Researchers exploring this possibility in a mouse model of *Gars* identified an interaction between the neuronal receptor neuropillin1 (NRP1) and mouse P234KY GARS, but not with wild-type GARS. However, this interaction was not observed for human GARS mutations *in vivo* and its significance to other disease-associated ARS mutations is unknown.

Two additional hypotheses for ARS-mediated CMT have largely been excluded, namely protein aggregation and editing defects. To date there is no evidence that protein aggregation is involved in ARS-mediated CMT disease. Since some ARS enzymes have the ability to inappropriately charge amino acids onto tRNA, editing domains in these enzymes evolved to minimize this occurrence. As discussed above, a mutation in the mouse *Aars* gene that affects the editing activity of mouse AARS leads to ataxia (Sarna and Hawkes 2011). Since the phenotype in ARS-mediated disease does not include ataxia, tRNA mischarging is not thought to be part of the mechanism of ARS-mediated CMT disease.

Significance of this thesis work

Peripheral neuropathies pose a significant health-care burden. Currently, there is no treatment for any of the 80+ forms of CMT disease (Timmerman et al. 2014). Instead, general physical and occupational therapy, orthopedic surgery, ankle-foot orthosis and mobility assisting devices (*i.e.*

walking aids and wheel chairs) are the only options for patients. Since CMT disease has a high level of locus and allelic heterogeneity, therapeutic interventions will likely need to be specific to the genetic lesion (Reilly et al. 2011). Thus, it is critical to identify new CMT-associated genes and alleles to help elucidate the specific mechanism by which mutated ARS genes lead to CMT disease to inform the development of targeted therapies. In order to do so, functional studies of specific gene lesions will play a critical role in elucidating the pathogenic mechanisms of genetic disease (Bouhy and Timmerman 2013).

In this thesis, I will describe genetic and functional studies implicating novel ARS alleles in both recessive and dominantly inherited human disease (Chapter 2, 3 and 4). The findings of these functional studies will then serve as the rationale for studies into the mechanism of ARS-mediated CMT disease (Chapter 5 and 6). Together, this body of work aspires to make contributions to understanding the allelic spectrum of ARS mutation and the pathomechanism of ARS-mediated diseases with the ultimate goal of informing therapeutic studies in the future.

CHAPTER 2

Evaluating ARS mutations for a pathogenic role in recessive disease

INTRODUCTION

With advancements in next generation sequencing, new genes and variants are continually being identified in cases of idiopathic disease. High throughput sequencing can be extremely useful in understanding the genetic underpinnings of *de novo* and recessive diseases where linkage studies cannot be used; however, careful evaluation of these variants is needed before classifying mutations as pathogenic (MacArthur et al. 2014). Recently, the National Institutes of Health (NIH) established a program to expedite the diagnosis in medical cases where standard clinical evaluations have proved inadequate. This program, termed the Undiagnosed Disease Program (UDP) (http://www.genome.gov/27544402), chooses individual clinical cases submitted by physicians for further study. Selected patients undergo an extensive clinical evaluation to carefully characterize their phenotype. Whole exome sequencing (WES) is performed on patient DNA to identify putative disease-causing mutations. Variants identified are then provided to researchers for follow-up studies investigating the pathogenicity of these mutations and the pathomechanism by which these variants cause disease. Interestingly, of the cases considered by the UDP, multiple patients have been found to be compound heterozygous for mutations in ARS enzymes. These ARS variants are particularly interesting for follow-up studies as 18 ARS genes have already been reported in the literature in patients with recessive syndromes, including 14 implicated in phenotypes that include neurologic deficits (Scheper et al. 2007a; McLaughlin et

al. 2010; Riley et al. 2010; Belostotsky et al. 2011; Götz et al. 2011; Pierce et al. 2011; Bayat et al. 2012; Elo et al. 2012; Steenweg et al. 2012; Cassandrini et al. 2013; Pierce et al. 2013; Taft et al. 2013; Diodato et al. 2014; Wolf et al. 2014; Zhang et al. 2014; Hadchouel et al. 2015). In addition to mutations identified as part of the UDP, ARS mutations have recently been identified in patients with recessive diseases by clinicians at the University of Michigan (Ann Arbor, MI) and the Clinic for Special Children (Lancaster, PA).

While ARS mutations identified in recessive phenotypes are presumed to cause disease by a loss-of-function mechanism, not all ARS variants reported in the literature have been functionally characterized. Cataloging the functional consequences of alleles identified in patients with recessive diseases will expand our understanding of ARS variants and may lead to potential insights into the relationship between a particular ARS gene and a given phenotype. Importantly, understanding the relationship between the function of an ARS variant and the phenotype in heterozygotic carriers may also lead to insights into dominantly inherited ARS-mediated CMT disease. Although literature on ARS-associated CMT has revealed loss-of-function characteristics for some implicated variants, the role of loss of ARS function in disease pathogenesis is still widely debated in the field (see Chapters 3-5).

In this chapter, we will evaluate the functional consequences of recessive ARS mutations that were identified as part of the UDP (*GARS*, *AARS*, and *DARS*) and by physicians at the University of Michigan (*FARSB*) and the Clinic for Special Children (*YARS*). We will consider the role of these variants in both dominant and recessive phenotypes in these families. Of the genes discussed in this chapter, only *DARS* mutations have been previously implicated in recessive disease (Taft et al. 2013); thus this work has the potential to implicate novel ARS genes in recessive phenotypes. We hypothesize that recessive mutations will result in a dramatic

reduction in ARS activity; however, at least one of the two alleles carried by each patient is predicted to retain some level of activity necessary to sustain viability, as complete loss of ARS function would not be predicted to be compatible with life (Delarue 1995).

A subset of data in this chapter was previously published in the *American Journal of Human Genetics* (Simons et al. 2015). Permission was requested for the reproduction of figures. The author performed all the studies in this chapter with the following exceptions: (1) Clinical collaborators collected the phenotypic data and performed mutation sequencing; (2) Aminoacylation assays were performed by Katrina Lu and Ya-ming Hou at Thomas Jefferson University; (3) The University of Michigan Sequencing core performed the DNA sequencing reactions described in this work; and (4) Haploid Δ*ALA1* and Δ*GRS1* yeast strains were previously generated by Heather McLaughlin (Antonellis Laboratory, University of Michigan) (McLaughlin et al. 2011) and by the Schimmel Laboratory at Scripps Research Institute (Turner et al. 2000).

MATERIALS AND METHODS

Patient sample collection, identification of ARS mutations, and linkage studies

Collaborating physicians performed the clinical evaluation and subsequent classification of patient phenotypes based on exam findings. DNA samples were collected with patient consent as per Institutional Review Board protocols at the participating institutions. *GARS, AARS,* and *DARS* mutations were identified as part of the NIH UDP. Additional *DARS* mutations previously reported by Taft *et al.* were also evaluated (Taft et al. 2013). *YARS* mutations were identified at the Regeneron Genetics Center in collaboration with clinicians at the Clinic for Special Children (Strasburg, PA). *FARSB* mutations were identified at the Michigan Molecular Genetic Laboratory at University of Michigan. Variants were identified via whole exome

sequencing and segregation studies were performed in additional family members when possible. Each identified ARS mutation was assessed for its presence in the general population via database searches of the Exome Variant Server (snp.gs.washington.edu/EVS), the 1000 Genomes Project database (www.1000genomes.org), and dbSNP (www.ncbi.nlm.nih.gov/snp). Conservation of the affected amino acid residue was determined by comparison of evolutionarily divergent species using ClustalW2 (Larkin et al. 2007).

Aminoacylation assays

Wild-type and mutant human *AARS* open reading frames were cloned in frame with a 6X His C-terminal tag in the pET-DEST42 gateway vector (Life Technologies) and expressed in *E. coli*. Protein was purified with the Co^{2+}-affinity resin per manufacturer's protocol (Novagen). Human $tRNA^{Ala}$ was prepared, purified, and annealed as previously described (Hou et al. 1993). Aminoacylation assays were performed to determine the steady-state kinetics of mutant AARS enzymes. Briefly, 20 mM of recombinant human wild-type or mutant AARS was incubated at 37°C with $tRNA^{Ala}$ in varying concentrations (0.6–16 mM) and ^{3}H-alanine (20 mM) in a buffer containing 20 mM KCl, 50 mM HEPES (pH 7.5), 4.0 mM DTT, 0.2 mg/ml BSA, 10 mM $MgCl_2$, and 2 mM ATP. Incorporation of ^{3}H-Ala (specific activity of 16,500 dpm/pmole) onto $tRNA^{Ala}$ to generate ^{3}H-Ala-$tRNA^{Ala}$ was determined by calculating acid precipitable counts on filter pads. Counts were corrected for the filter quench factor and were calculated as pmoles of ^{3}H-Ala-$tRNA^{Ala}$. Data were fit to the Michaelis-Menten equation. Four independent experiments were performed, and error bars represent standard deviations.

Generation of yeast ARS expression constructs and haploid yeast strains

PCR primers were designed to amplify the orthologous *S. cerevisiae* ARS locus (Table 2.1)

Table 2.1. Human and Yeast orthologous ARS locus

Human ARS	Yeast ARS
AARS	*ALA1*
DARS	*DPS1*
GARS	*GRS1*
FARSB	*FRS1*
YARS	*TYS1*

including the endogenous promoter from yeast genomic DNA. Gateway (Invitrogen) cloning sequences (attB1 (forward) and attB2 (reverse); primer sequences in Appendix A) were included to facilitate recombination into the pDONR221 vector. Subsequent to PCR, BP Clonase (Invitrogen) reactions were employed to clone the wild-type ARS locus into pDONR221 and resulting colonies were fully sequenced to rule out PCR-induced errors. Next, mutation-containing oligonucleotides (primer sequences in Appendix A) were employed to model the human mutations at the corresponding residue in the orthologous yeast gene (Table 2.1). The QuickChange II XL Site-Directed Mutagenesis Kit (Stratagene) was used to mutate each locus in the pDONR221 entry construct. Resulting clones were purified and sequenced to confirm successful mutagenesis and rule out PCR-induced errors. The wild-type and mutant ARS/pDONR221 entry clones were subsequently recombined into a Gateway-compatible *URA3*-bearing pRS316 or *LEU2*-bearing pRS315 destination vector using LR Clonase (Invitrogen). Resulting clones were purified and digested with *Bsr*GI (New England Biosystems) to confirm successful recombination.

Δ*ALA1* and Δ*GRS1* haploid strains containing a wild-type copy of the *ALA1* or *GRS1* locus, respectively, on a *URA3*-bearing vector were previously generated (Turner et al. 2000; McLaughlin et al. 2011). To test the remaining ARS variants, commercially available diploid heterozygous yeast strains were purchased for each ARS ortholog of interest (Open Biosystems) to generate additional haploid yeast strains. The strains utilized were: (1) *YARS*: diploid heterozygous Δ*TYS1* yeast strain *(MATa/α, his3Δ1/ his3Δ1, leu2Δ0/ leu2Δ0, met15Δ0/ MET15, ura3Δ0/ ura3Δ0)*; (2) *DARS*: diploid heterozygous Δ*DPS1* yeast strain (*MATa/α, his3Δ1/ his3Δ1, leu2Δ0/ leu2Δ0, met15Δ0/ MET15, ura3Δ0/ ura3Δ0*); and (3) *FARSB*: diploid heterozygous Δ*FRS1* yeast strain (*MATa/α, his3Δ1/ his3Δ1, leu2Δ0/ leu2Δ0, met15Δ0/ MET15, ura3Δ0/*

ura3Δ0). Each diploid strain was generated by replacing one endogenous locus of interest with a KAN^R cassette (Winzeler et al. 1999). The heterozygous diploid strain was transformed with an *URA3*-bearing pRS316 vector containing the wild-type ARS locus (Table 2.1) (see above for construct generation). Lithium acetate yeast transformations were performed at 30°C using 200 ng of plasmid DNA. Transformed yeast were grown on yeast media lacking uracil. Resulting diploid yeast containing the wild-type pRS316 maintenance vector were treated to induce sporulation. Briefly, the diploid strain was patched twice onto GNA presporulation plates (5% D-glucose [Fisher Scientific], 3% nutrient broth [Becton, Dickinson and Company], 1% yeast extract [Acros Organics], and 2% agar [Teknova]). One microliter of cells was then transferred into 2 ml of supplemented liquid sporulation medium (1% potassium acetate [Fisher Scientific], 0.005% zinc acetate [Fisher Scientific], 1X ura supplement, 1X his supplement, and 1X leu supplement) and incubated for 5 days at 25°C followed by 3 days at 30°C. Sporulated yeast were dissected using a MSM 400 dissection microscope (Singer Instruments) and plated on yeast extract, peptone, and dextrose (YPD) plates (Becton, Dickinson and Company). Resulting spores were individually patched onto solid growth medium containing Geneticin (G418) or 0.1% 5-fluoroorotic acid (5-FOA), or media lacking uracil (Teknova, Hollister, CA). Two resulting haploid yeast clones that grew on G418 medium and medium lacking uracil, but failed to grow on 5-FOA medium, were selected for use in the yeast complementation assays.

Yeast complementation assays

Two haploid ΔARS strains [harboring a maintenance vector to express wild-type ARS and *URA3*] per locus were then transformed with an empty vector (‘Empty pRS315’ in figures) or the appropriate wild-type or mutant ARS in a *LEU2*-bearing pRS315 vector and selected on medium lacking uracil and leucine (Teknova). For each transformation, at least two independent pRS315

plasmid preparations were used and at least two colonies from each plasmid were selected for additional analysis. Selected colonies were grown to saturation for 2 days at 37°C in liquid media lacking uracil and leucine. A 10 µl aliquot of each culture was spotted undiluted or diluted 1:10 or 1:100 in H_2O on plates containing 0.1% 5-FOA (Teknova) or medium lacking uracil and leucine and incubated at 30°C for 48-72 hours. Survival was determined by visual inspection of growth.

Yeast growth curves

Growth curves were generated using pseudo-heterozygous yeast, which contain both wild-type *TYS1* on the *URA3*-bearing maintenance vector and wild-type, G41R, or P167T *TYS1* or no insert on the *LEU2*-bearing vector. Three yeast colonies from each transformation were selected for growth curves. Yeast were grown to saturation for 2 days at 30°C in liquid media lacking uracil and leucine. Saturated cultures were diluted to an OD_{600} = 0.200 +/- 0.015 based on OD_{600} readings recorded on the Absorbance module E7061 on the Glomax®-Multi Microplate Multimode Reader Detection System (Promega, Madison, WI). 20 µl of OD_{600} = 0.200 +/- 0.015 cultures was added to 4 mL of media lacking uracil and leucine or media contain 0.1% 5FOA and incubated at 30°C shaking. Growth curves were generated using the average OD_{600} measurements from three biological replicate cultures for each pseudo-heterozygous strain at the indicated time points. Error bars represent standard deviation. Growth curves were performed twice using independent yeast transformation samples.

Global protein assessment

To compare total protein levels in wild-type and K81T/R751G *AARS* fibroblasts, total protein was extracted from 500,000 cells per culture using 200 µl of a cocktail containing RIPA buffer and 100X protease inhibitor (Thermo Scientific, Rockford, IL). Cells were lysed by rocking for

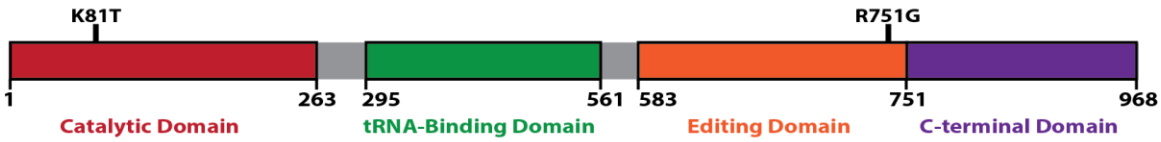
30 min at 4°C. Lysates were centrifuged at 4°C for 30 min at 15,000 rpm to remove cellular debris. 20 μl of lysates were mixed with 20 μl of 2X SDS-Glycine running buffer (Life Technologies, Grand Island, NY) and 2 μl of β-mercaptoethanol and run on a 4-12% Tris-glycine for 2 hours at 100 volts (Life Technologies). The protein gel was then washed in distilled water, stained with SimplyBlue SafeStain (Invitrogen, Carlsbad, CA) for 1 hour, and washed again in distilled water overnight. Relative protein levels were determined by visual inspection of staining.

RESULTS

AARS mutations identified in patients with early-onset epileptic encephalopathy.

Three children, including two siblings and one unrelated child, were diagnosed with microcephaly and refractory myoclonic epilepsy in infancy. All three children demonstrated spasticity, blepharospasm (involuntary tight closure of eyelids), orobuccal dyskinesia (involuntary movement of the mouth), dystonia of limbs, chorea, and absent peripheral deep tendon reflexes. MRI studies revealed progressive atrophy of both grey and white matter and hypomyelination in all patients. Whole exome sequencing determined that all three patients carried biallelic *AARS* mutations. The two siblings were compound heterozygous for K81T (c.242A>C; p.Lys81Thr) and R751G (c.2251A>G; p.Arg751Gly) *AARS*. The K81T *AARS* allele was maternally inherited and was neither identified in general population databases nor in any of the control exomes examined. The *AARS* R751G mutation was paternally inherited and was present at an allele frequency of 0.00005% (6/120,000) in the dbSNP database, but absent from other general population databases and control exomes. Of note, all six occurrences of the R751G *AARS* allele were in heterozygous individuals. The third child tested was homozygous for the R751G *AARS* allele. K81T *AARS* is located in the aminoacylation domain of AARS, while

A



B

	AARS K81T	AARS R751G
Human	CIRAGG**K**QNDLDD	IAKGIR**R**IVAVTG
Mouse	CIRAGG**K**HNDLDD	IAKGIR**R**IVAVTG
Zebrafish	CIRAGG**K**HNDLDD	IAKGIR**R**IVAVTG
Worm	CIRAGG**K**HNDLDD	IAKGIR**R**IVALTG
Yeast	CIRAGG**K**HNDLED	IAKGIR**R**IVAVTG
Bacteria	CVRAGG**K**HNDLEN	TAAGVR**R**IEAVTG

Figure 2.1. Localization and conservation of human *AARS* mutations. (A) A cartoon of the known functional domains of the AARS protein including the catalytic domain (red), the tRNA binding domain (green), the editing domain (orange), and the C-terminal domain (purple). The residues of each domain are indicated (bottom). K81T and R751G *AARS* are mapped on the protein structure (top). (B) A multiple-species sequence alignment demonstrates the conservation of the affected residues for K81T and R751G *AARS* mutations and the flanking protein sequences in multiple and evolutionarily diverse species. The affected residue is in red text.

R751G is located in the editing domain (Figure 2.1A). Both residues are conserved from human to *E. coli* (Figure 2.1B).

K81T and R751G AARS *demonstrate decreased aminoacylation activity* in vitro.

ARS mutations previously implicated in recessive disease have demonstrated decreased aminoacylation, including recessive *KARS* and *QARS* mutations (McLaughlin et al. 2010; Zhang et al. 2014). *In vitro* aminoacylation assays were used to test the ability of human K81T and R751G *AARS* to charge $tRNA^{Ala}$ with tritium labeled alanine. Both the K_m (a measure of the substrate concentration necessary for effective catalysis) and the k_{cat} (a direct measure of catalysis generating the final product under ideal conditions including enzyme saturation) were determined. The K_m/k_{cat} was used to calculate aminoacylation for wild-type and mutant AARS. The ratio of mutant AARS activity compared to wild-type activity was used to determine the overall effects of each mutation on aminoacylation. K81T AARS had a 2-fold increase Km, resulting in a 2-fold decrease in aminoacylation (K_m/k_{cat}) compared to wild-type AARS (Table 2.2). In contrast, R751G AARS showed both an increase in K_m by 2-fold and a decrease in k_{cat} by 5-fold, resulting in a 90% reduction in tRNA charging efficiency compared to wild-type AARS (Table 2.2). These data indicate that both K81T and R751G *AARS* are loss of function *in vitro*.

K81T ALA1 *results in reduced growth in yeast complementation assays.*

Disease-associated aminoacyl-tRNA synthetases mutations have been tested for loss-of-function characteristics *in vivo* using yeast complementation assays, including mutations in *AARS* that have been implicated in peripheral neuropathy (Antonellis et al. 2006; Jordanova et al. 2006; McLaughlin et al. 2010; McLaughlin et al. 2011; Pierce et al. 2011; Vester et al. 2012; Cassandrini et al. 2013; Gonzalez et al. 2013; Diodato et al. 2014; Griffin et al. 2014; Gonzaga-

Jauregui et al. 2015; Hadchouel et al. 2015; Safka Brozkova et al. 2015). To assess the functional consequences of K81T and R715G *AARS in vivo,* a previously validated haploid yeast strain that has the endogenous yeast ortholog of *AARS, ALA1,* deleted and viability maintained by expression of wild-type *ALA1* on a *URA3* selective vector (McLaughlin et al. 2011) was transformed with either an empty vector or one containing the wild-type, K81T, or R751G *ALA1* alleles. The mutations were modeled by mutating the orthologous amino acid residue in the yeast protein (Table 2.3), and mutations are referred to by the human mutation for clarity. Yeast were then selected on media containing 5-FOA to select for yeast cells that have spontaneously lost the *URA3*- maintenance vector (Boeke et al. 1987). Only yeast cells that express a functional *ALA1* allele on the *LEU2*-bearing pRS315 vector are able to survive. Wild-type *ALA1* sustained yeast viability, while the empty vector ('Empty pRS315') was unable to complement loss of the endogenous *ALA1* allele (Figure 2.2), consistent with the *ALA1* being an essential gene in yeast (McLaughlin et al. 2011). Yeast expressing R751G *ALA1* showed survival comparable to expression of wild-type *ALA1*, while yeast expressing K81T *ALA1* demonstrated reduced growth on 5-FOA media (Figure 2.2). These results strongly suggest that K81T *AARS* represents a hypomorphic *AARS* allele *in vivo*.

K81T / R751G AARS *fibroblasts do not show a dramatic reduction in total protein levels.* tRNA charging is an essential first step in protein translation. Thus, mutations that result in impaired charging could impair global protein translation, resulting in decreased protein levels in cells (Wallen and Antonellis 2013). To determine if the K81T and R751G *AARS* mutations resulted in decreased protein translation, protein was extracted from patient fibroblasts generated from an individual who was compound heterozygous for K81T and R751G *AARS* and from wild-type control fibroblasts. Protein was extracted from equal cell numbers based on cell counts and

Table 2.2. Aminoacylation Kinetics of AARS Variants

Variant	K_m (μM)	k_{cat} (s^{-1})	K_m/ k_{cat} ($\mu M^{-1}s^{-1}$)	Ratio to WT
Wild-type	3.1 ± 0.7	0.4 ± 0.1	0.13 ± 0.04	1
K81T	6.0 ± 1.8	0.4 ± 0.1	0.07 ± 0.03	1/2
R751G	6.3 ± 1.4	0.07 ± 0.03	0.01 ± 0.005	1/10

Table 2.3. Human *AARS* Variants Modeled in the Yeast Ortholog *ALA1*

Human AARS[a]	Yeast ALA1[b]
K81T	K85T
R751G	R747G

[a]Amino acid coordinates correspond to GenBank accession number BAA06808.1.
[b]Amino acid coordinates correspond to GenBank accession number EDV1089.

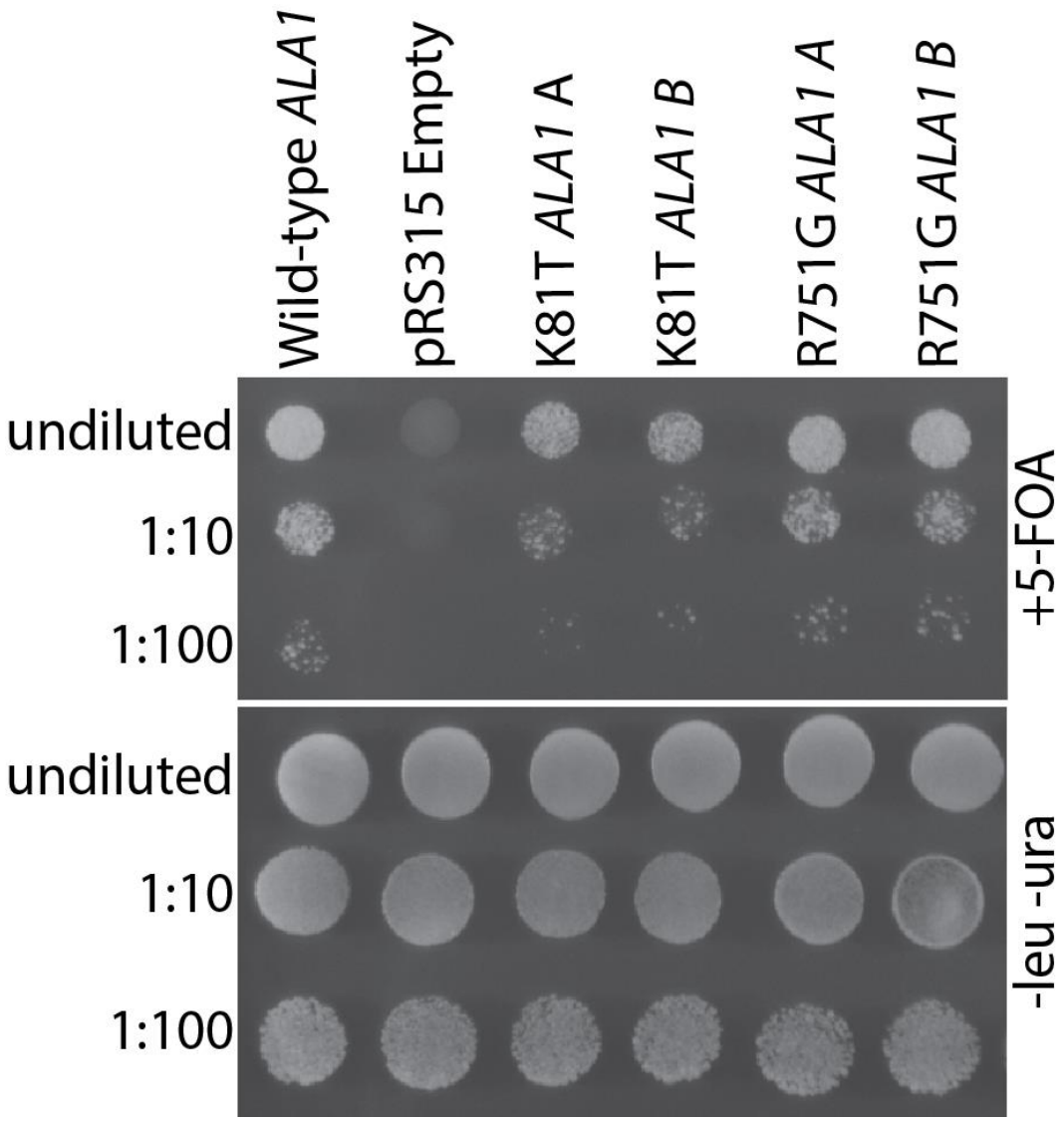


Figure 2.2. Yeast complementation assays for recessive *AARS* alleles. Haploid Δ*ALA1* yeast strains were transformed with a vector containing an insert to express wild-type, K81T, or R751G *ALA1* or no insert ('pRS315 Empty'). Resulting cultures were spotted (undiluted and diluted 1:10 and 1:100) on complete agar growth medium plates containing 5-FOA (top panel) or on SD –leu -ura agar growth medium plates (bottom panel) and incubated at 30ºC for two days. Experiments were performed using independently generated *ALA1* expression constructs for each mutant allele (indicated by 'A' and 'B' along the top). Culture dilutions are indicated on the left and media conditions are indicated on the right.

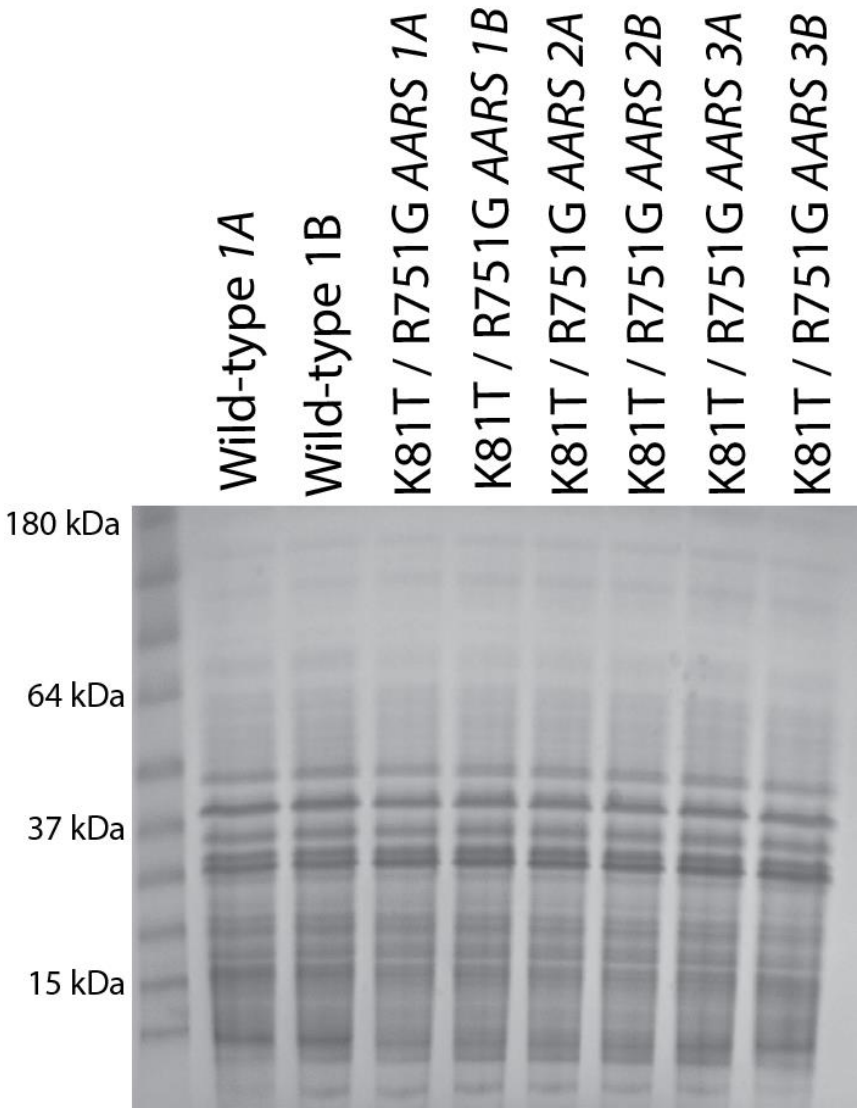


Figure 2.3. Evaluation of protein levels in *AARS* fibroblasts. Protein extracted K81T / R751G *AARS* patient fibroblasts were compared to protein from wild-type fibroblasts. Protein lysates were prepared from equal numbers of cells for each sample. Three independently prepared protein lysates from patient fibroblasts are shown (1, 2, and 3) with two replicates from each isolation included for comparison (A and B).

equal volumes of total protein were evaluated. The total protein levels in patient fibroblast samples were comparable to levels of total protein from control fibroblast extracts (Figure 2.3), indicating that compound heterozygosity of K81T and R751G *AARS* does not lead to a severe reduction in global protein levels.

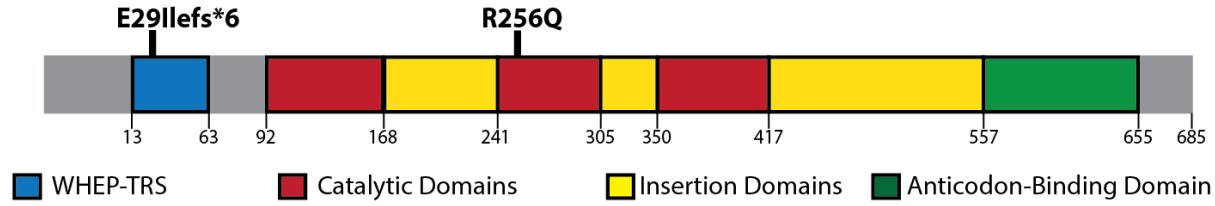
GARS mutations identified in patients with microcephaly.

An individual with idiopathic microcephaly was evaluated at the NIH as part of the UDP. The patient was diagnosed with growth retardation *in utero* and had decreased body weight and short stature at the time of clinical evaluation. The patient also had an atrial septal defect, abnormal facial features, including a smooth philtrum and increased distance between the eyes (hypertelorism), and demonstrated a failure to thrive. Whole exome sequencing revealed the patient was compound heterozygous for E29Ifs*6 [c.84_87del, p.Glu29Ilefs*6] and R256Q [c.767G>A, p.Arg265Gln] *GARS*. mRNA transcribed from the E29Ifs*6 *GARS* allele is predicted to undergo nonsense-mediated decay (NMD) and thus would not be expected to produce protein product (Amrani et al. 2006). In contrast, the R256Q *GARS* allele is predicted to code for an expressed, stable protein that affects a residue in the catalytic domain that is conserved from human to worm (Figure 2.4A and B).

R256Q GRS1 *complements loss of endogenous* GRS1.

To test the functionality of R256Q *GARS*, we utilized the yeast complementation assay that had been previously developed to study CMT2D-associated *GARS* mutations (Antonellis et al. 2006; Turner et al. 2009). Briefly, based on conservation of the surrounding amino acids, we were confident that the orthologous residue to human R256 GARS in yeast is K239 GRS1 (Table 2.4). Both residues are polar basic amino acids, and thus the presence at this position of an arginine or lysine is not predicted to have a dramatic impact on the structure and function of the enzyme.

A



B

GARS
R256Q

Human	IFLNFKRLLEFNQ
Mouse	IFLNFKRLLEFNQ
Zebrafish	IFLNFKRLLEFNQ
Worm	IFVNFKRLLEFNQ
Yeast	QFLNFNKLLEFNN

Figure 2.4. Localization and conservation of *GARS* mutations. (A) The position of E29Ifs*6 and R256Q *GARS* are indicated on a cartoon of GARS protein structure (top). The functional domains of GARS are depicted with the amino acid positions marked for each domain (bottom): WHEP-TRS (blue), catalytic domains (red), insertion domains (yellow), and anticodon-binding domain (green). (B) Alignment of R256 GARS (red text) and surrounding amino acids in evolutionarily diverse species are shown. Note that the predicted orthologous residue in yeast is a lysine (purple text).

Table 2.4. Human *GARS* Variants Modeled in the Yeast Ortholog *GRS1*

Human GARS[a]	Yeast GRS1[b]
R256Q	K239Q

[a]Amino acid coordinates correspond to GenBank accession number AAA57001.1.
[b]Amino acid coordinates correspond to GenBank accession number NP_009679.2.

However, the substitution of an uncharged polar residue, such as glutamine that would occur as a result of the R256Q mutation, may have an appreciable effect on GARS activity. For consistency with other variants described in this thesis, the mutation in yeast is referred to by the human nomenclature (R256Q). Yeast strains lacking *GRS1* with viability maintained by a *URA3*-bearing vector encoding wild-type *GRS1* (Turner et al. 2009) were transformed with wild-type or R256Q *GRS1* or an empty *LEU2*-bearing vector. Resulting yeast were plated on 5-FOA media to select for yeast that spontaneously lost the *URA3* maintenance vector (Boeke et al. 1987). Visual inspection of yeast growth revealed that the R256Q expressing yeast grew to levels comparable to yeast expressing wild-type *GRS1* (Figure 2.5). This indicates that in yeast, the R256Q variant is sufficiently active to complement loss of endogenous *GRS1*.

DARS mutations identified in patients with hypomyelination.

Mutations in *DARS* had been previously identified in ten patients with recessive leukoencephalopathy (Taft et al. 2013); however, functional studies were not performed on any of these variants. Recently, two additional *DARS* mutations, H280L (c.1099G>C, p.Asp367His) and D367H (c.839A>T, p.His280Leu), were identified in a patient with recessive hypomyelination enrolled in the NIH's UDP protocol. Similar to the previously reported cases (Taft et al. 2013), this patient had decreased myelination in both the brain and the spinal cord and spasticity. Both H280L and D367H *DARS* mutations reside in the catalytic domain of *DARS* and are conserved from human to yeast (Figure 2.6A and B). Interestingly, D367H *DARS* affects the same residue as D367Y, a mutation previously implicated in recessive hypomyelination (Taft et al. 2013). Of the previously reported mutations used in this study, A274V, D367Y, and R494G *DARS* reside in the catalytic domain and are conserved from human to yeast (Figure 2.6A and

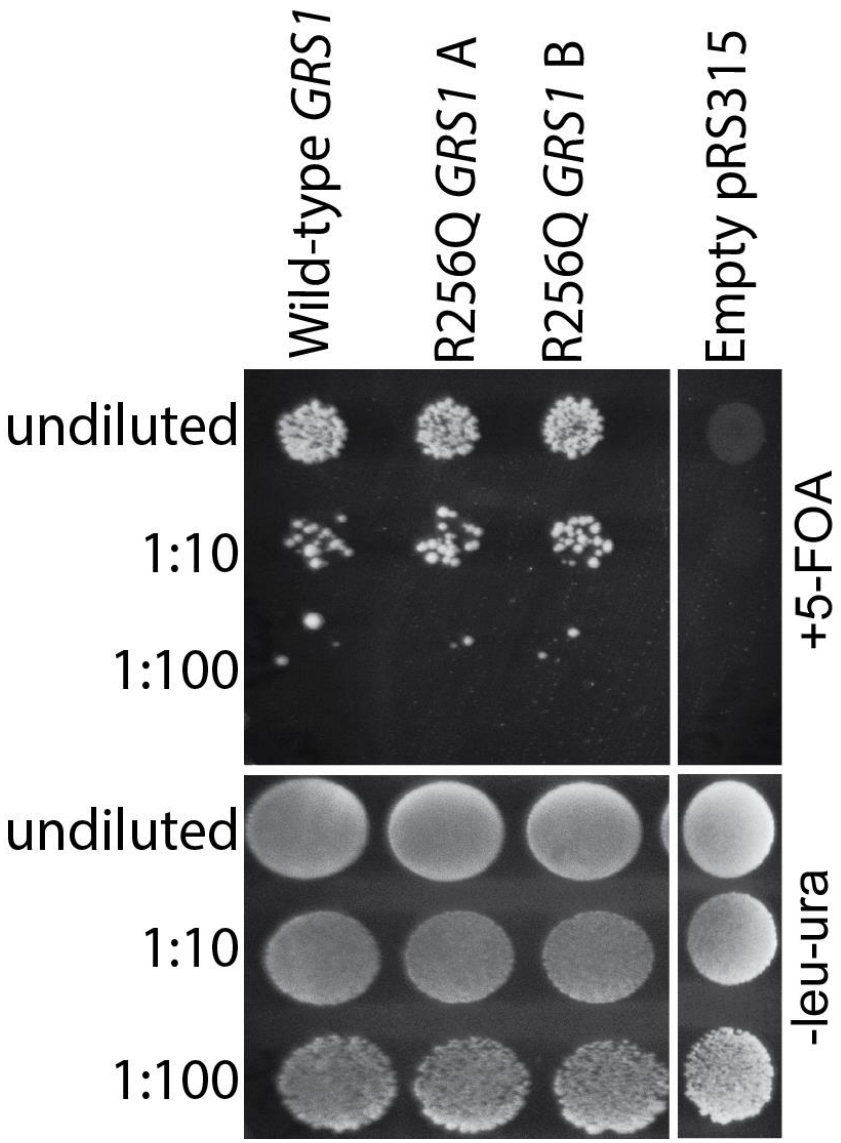
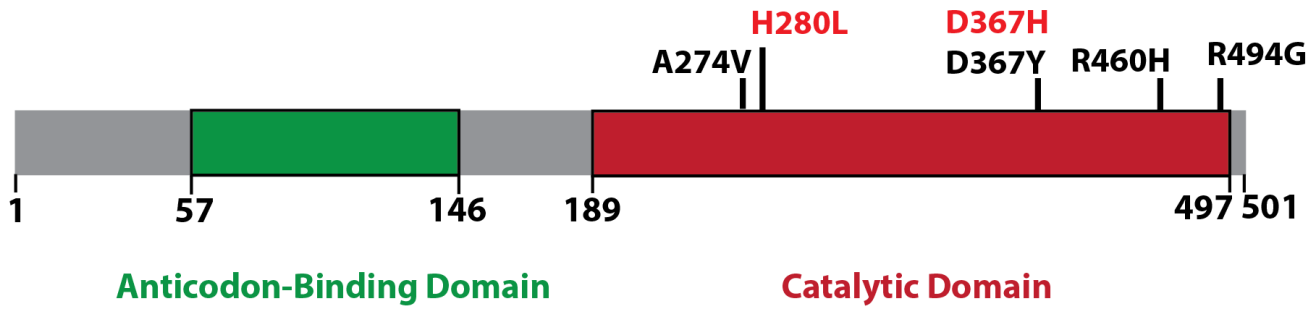


Figure 2.5. Yeast Complementation assay for R256Q *GRS1*. Haploid yeast lacking endogenous *GRS1* were transformed with wild-type or R256Q *GRS1* or an empty vector ('Empty pRS315) (top) and grown on plates containing 5-FOA or lacking leucine and uracil at various concentrations. Yeast expressing wild-type or R256Q *GRS1* grew robustly, while yeast containing only an empty vector were unable to survive on 5-FOA.

A



B

	DARS A274V	DARS H280L	DARS D367H/Y	DARS R460H	DARS R494G
Human	IGPVFR**A**EDSNTH	AEDSNT**H**RHLTEF	EMGDED**D**LSTPNE	AYIDSF**R**FGAPPH	QTSMFP**R**DPKRLT
Mouse	IGPVFR**A**EDSNTH	AEDSNT**H**RHLTEF	EMDDEE**D**LSTPNE	AYIDSF**R**FGAPPH	QTSMFP**R**DPKRLT
Zebrafish	VGPVFR**A**EDSNTH	AEDSNT**H**RHLTEF	EMGDEE**D**LSTPNE	AYIDSF**R**YGAPPH	QTSMFP**R**DPKRLT
Worm	IGPVFR**A**EDSNTH	AEDSNT**H**RHMTEF	EIGDED**D**LSTPVE	SYIDSFKYGCPPH	LASLFP**R**DPKRLT
Yeast	IGPVFR**A**ENSNTH	AENSNT**H**RHMTEF	EIGDFE**D**LSTENE	DYCDGF**S**YGCPPH	RASLFP**R**DPKRLR

Figure 2.6. Localization and conservation of *DARS* mutations. (A) DARS has two major functional domains, the anticodon-binding domain (green) and the catalytic domain (red). All *DARS* mutations used in this study affect residues in the catalytic domain (indicated on top of cartoon). The newly identified mutations, H280L and D367H *DARS*, are in red text. (B) DARS protein alignments of multiple evolutionarily diverse species is depicted. The affected residue is in red, bold text. All affected residues are conserved from human to yeast except R460 DARS, which has a serine (purple, bold text) at the orthologous residue.

B). R460H *DARS* also affects an amino acid in the catalytic domain, but is only conserved in vertebrates.

DARS *mutations support yeast viability.*

Although predicted to be deleterious using multiple damage-prediction algorithms, *DARS* mutations previously implicated in recessive leukoencephalopathy have not been previously tested for effects on enzyme activity (Taft et al. 2013). As such, we developed a yeast complementation assay to test previously implicated mutations (A274V, D367Y, R460H, and R494G *DARS*) and the two newly identified *DARS* mutations (H280L and D367H). Haploid yeast strains lacking *DPS1*, the yeast ortholog of *DARS*, were generated. Viability of the strain was maintained using an *URA3*-bearing plasmid that contains the wild-type *DPS1* locus. A274V, H280L, D367H, D367Y, R460H, and R494G *DARS* were modeled in the *DPSI* gene based on conservation (Table 2.5). Wild-type or mutant *DPS1* or an empty vector control was introduced into the haploid Δ*DPS1* strain. Yeast were subsequently selected for the spontaneous loss of the *URA3*-maintenance vector using 5-FOA media (Boeke et al. 1987). Visual inspection of the plates revealed that all alleles were able to sustain yeast viability; however A274V and D367Y *DPS1* showed decreased growth compared to wild-type *DPS1* expressing yeast that could be appreciated at 1:10 and 1:100 dilutions (Figure 2.7). Thus, when modeled in yeast, none of the tested *DARS* alleles had a dramatic reduction in activity.

FARSB *mutations identified in a patient with interstitial lung and liver disease.*

An infant presented to the University of Michigan Medical Genetics clinic with liver and interstitial lung disease. Exome sequencing revealed the patient was compound heterozygous for T256M (c.767C>T; pThr256Met) and H496Kfs*14 (c.1486delCinsAA; p.His496Lysfs*14)

Table 2.5. Human *DARS* Variants Modeled in the Yeast Ortholog *DPS1*

Human DARS[a]	Yeast DPS1[b]
A274V	A326V
H280L	H332L
D367H	D421H
D367Y	D421Y
R460H	S516H
R494G	R550G

[a]Amino acid coordinates correspond to GenBank accession number AAI07750.1.
[b]Amino acid coordinates correspond to GenBank accession number CAA66172.1.

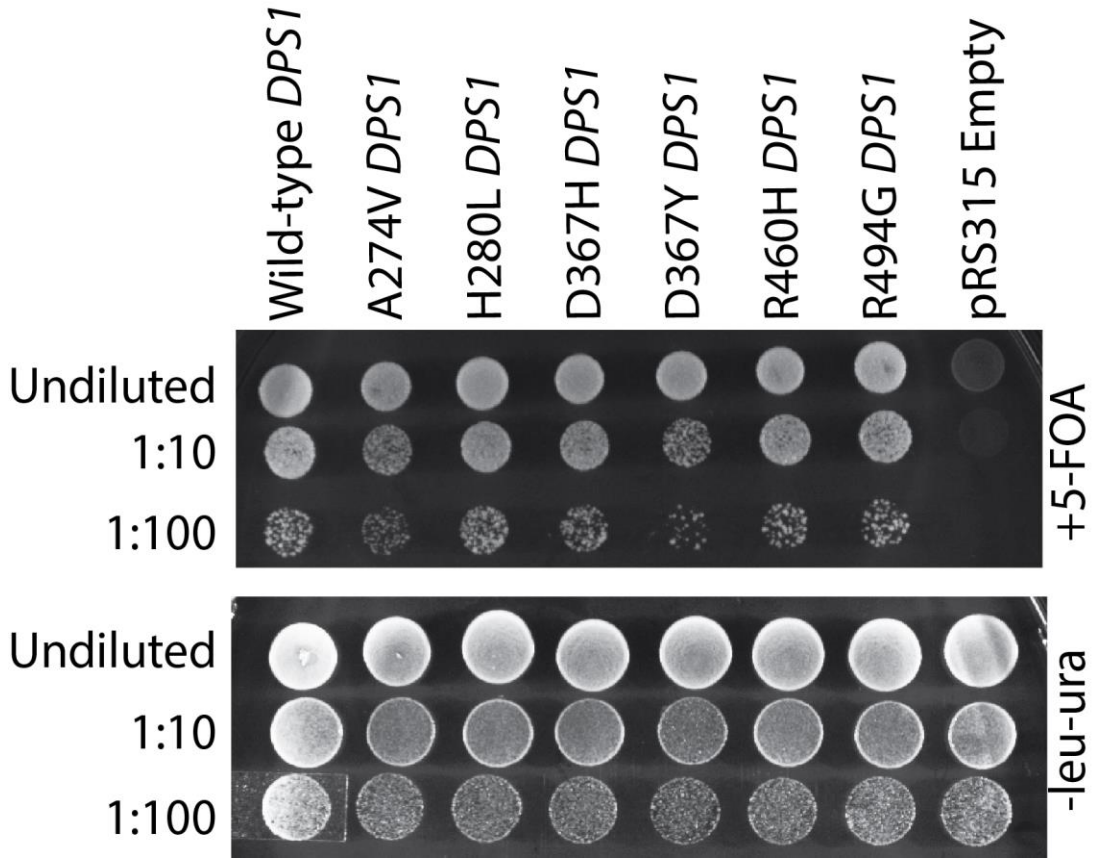


Figure 2.7. Yeast expressing disease associated *DPS1* alleles. Yeast lacking endogenous *DPS1* were transformed with *LEU2*-bearing vectors containing wild-type of mutant *DPS1* or no insert ('pRS315 Empty') (top). All mutations were able to sustain yeast viability; however, A274V and D367Y *DPS1* showed a mild reduction in yeast growth. Note the subtle differences at 1:10 and 1:100 dilutions on 5-FOA.

FARSB. The H496Kfs*14 *FARSB* mRNA is predicted to undergo nonsense mediated decay (Amrani et al. 2006). T256M *FARSB*, in contrast, is expected to be expressed and stable and affects a residue in the tRNA binding domain that is conserved from human to yeast (Figure 2.8A and B).

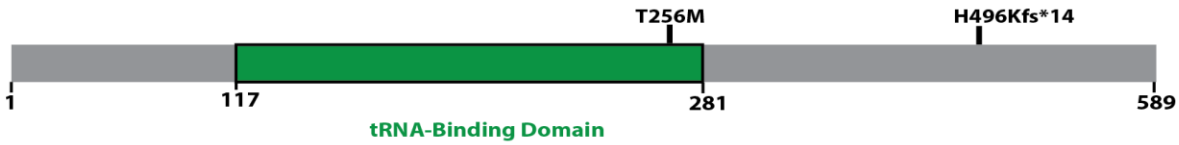
FARSB *mutations support yeast viability.*

In order to functionally test the T256M *FARSB* mutation, we developed a haploid yeast strain, similar to strains used in previously described studies. T256M *FARSB* was modeled in the yeast ortholog *FRS1* (Table 2.6). Δ*FRS1* [containing the *FRS1* locus on a *URA3* maintenance vector] was transformed with a *LEU2*-vector containing wild-type or T256M *FRS1* or an empty vector. After plating on 5-FOA plates to select against cells carrying the maintenance vector (Boeke et al. 1987), it was determined that T256M *FRS1* supports yeast growth to wild-type *FRS1* levels, indicating T256M *FRS1* does not demonstrate loss of function in this assay (Figure 2.9).

Recessive YARS *mutations identified in a child with microcephaly.*

An infant child from a consanguineous family presented to the Clinic for Special Children with microcephaly. This child and a sibling both ultimately died in early infancy. Exome sequencing performed by Regeneron revealed that these patients were homozygous for P167T (c.499C>A; p.Pro167Thr) *YARS*. Because mutations in *YARS* had been previously implicated in dominant intermediate CMT (DI-CMT), a family history of neuropathy was investigated. The parents reported symptoms consistent with a mild peripheral neuropathy. P167T *YARS* affects a residue conserved from human to yeast that is located in the catalytic domain (Figure 2.10).

A



B

FARSB
T256M

Human	NIFIEC**T**GTDFTK
Mouse	NIFIEC**T**GTDFTK
Zebrafish	NVFIEC**T**ATDLTK
Worm	NVFIEA**T**ATDKQK
Yeast	NILIDI**T**ATDKTK

Figure 2.8. Localization and conservation of *FARSB* mutations. (A) The position of T256M and H496Kfs*14 *FARSB* are depicted on the cartoon of the FARSB protein. FARSB contains a tRNA-binding domain (green), but lacks a catalytic domain. (B) T256M *FARS*B affects a residue conserved from human to yeast (bold, red text).

Table 2.6 Human *FARSB* Variant Modeled in the Yeast Ortholog *FRS1*

Human FARSB[a]	Yeast FRS1[b]
T256M	T246M

[a]Amino acid coordinates correspond to GenBank accession number NP_005678.3.
[b]Amino acid coordinates correspond to GenBank accession number P15624.3.

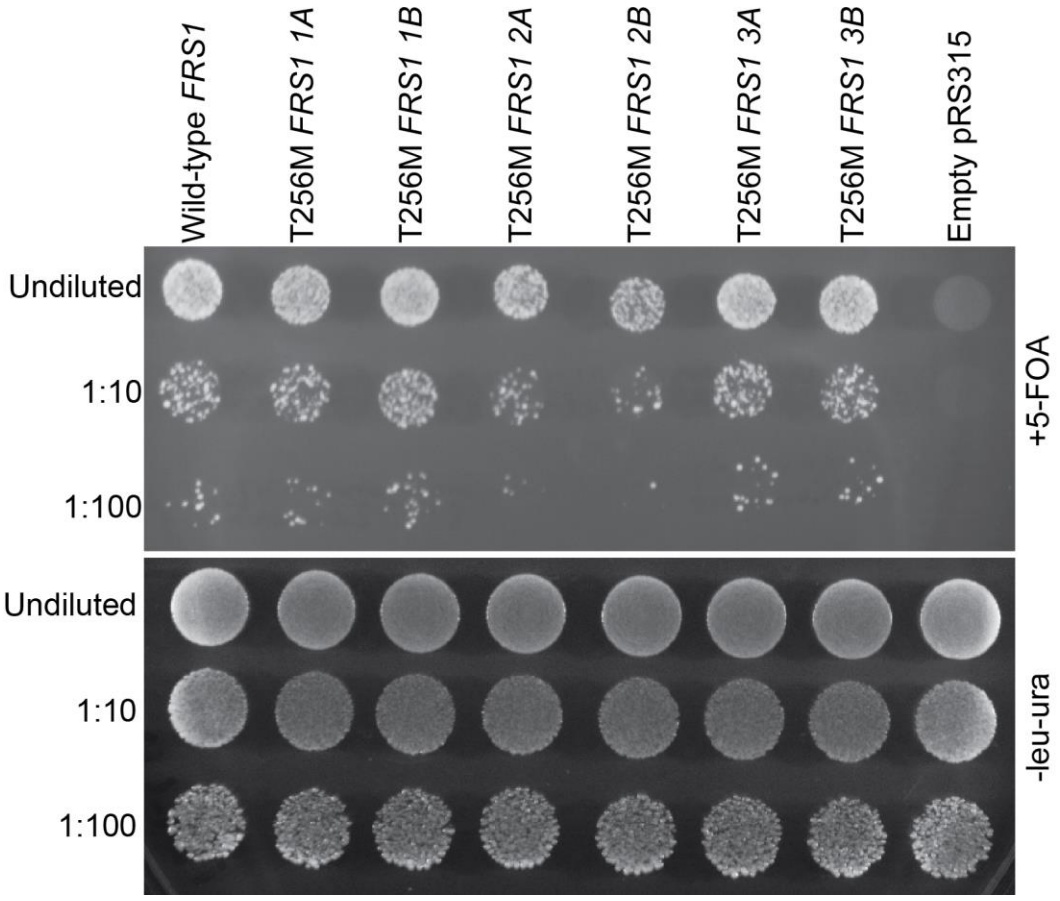


Figure 2.9. Yeast complementation assay for T256M *FRS1*. Yeast lacking endogenous *FRS1* were transformed with a *LEU2*-bearing vector containing wild-type or T256M *FRS1* or an empty vector ('Empty pRS315') (top) and spotted on 5-FOA plates. T256M *FRS1* sustained yeast viability to levels comparable to wild-type *FRS1*. Transformations were performed with three independent plasmid preps (1, 2 and 3) and two colonies (A and B) were selected from each transformation for evaluation.

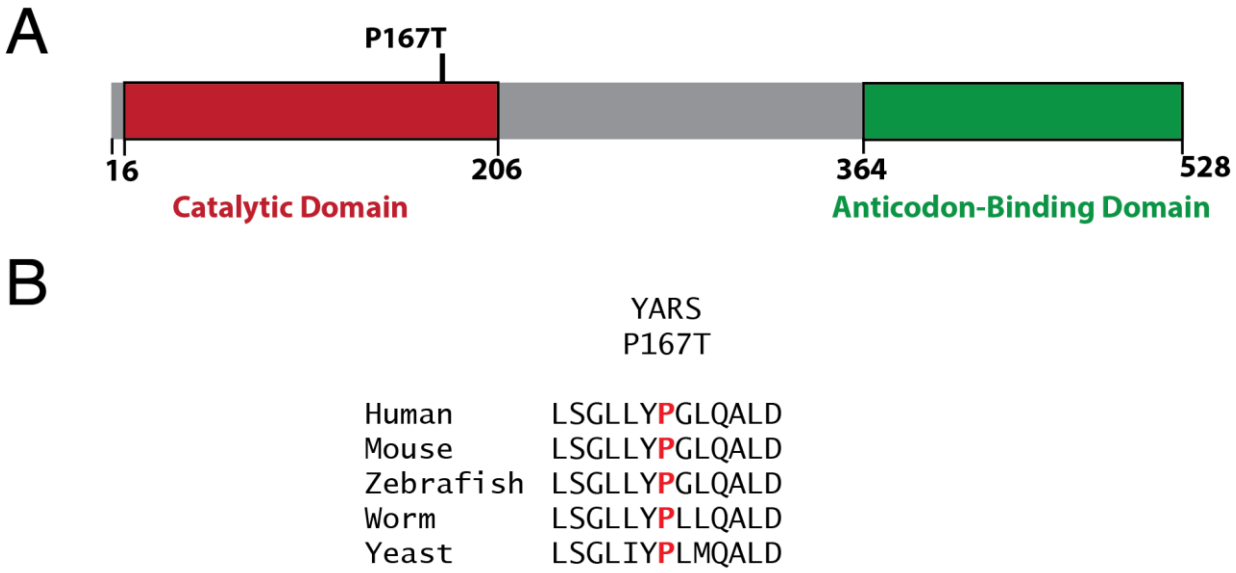


Figure 2.10. Localization and conservation of *YARS* mutations. (A) A cartoon depiction of YARS illustrates the location of the catalytic domain (red) and an anticodon-binding domain (green). The amino acid residues comprising each domain are listed below the cartoon. The position of P167T *YARS* in the catalytic domain is illustrated on top of the cartoon. (B) P167T *YARS* affects a residue conserved from human to yeast.

Yeast complementation studies suggest P167T YARS *is a hypomorphic allele.*

To understand the functional consequences of the P167T *YARS* on enzyme activity, we employed a haploid yeast strain that was deleted for the endogenous *TYS1* locus, the yeast ortholog of *YARS*. As with other ARS yeast models, viability was maintained using a wild-type copy of *TYS1* on an *URA3*-containing vector, allowing for selection against the maintenance vector using 5-FOA media (Boeke et al. 1987). Mutations of interest were modeled based on conservation between human YARS and yeast TYS1 (Table 2.7). Wild-type, G41R (a previously reported loss of function allele implicated in dominantly inherited CMT disease (Jordanova et al. 2006)), or P167T *TYS1* or an empty vector were introduced into the haploid yeast strain. After selection on solid 5-FOA media, P167T *TYS1* showed a dramatic reduction in yeast growth compared to wild-type (Figure 2.11); however the P167T *TYS1* allele supported significantly more yeast growth than either the functional null G41R *TYS1* and the empty vector control. To quantify the difference in growth between wild-type and P167T *TYS1* expressing yeast, we performed growth curves in liquid 5-FOA media using pseudo-heterozygous strains carrying both a wild-type allele on a *URA3*-bearing vector and the experimental allele on a *LEU2*-bearing vector (Figure 2.12). Yeast carrying the G41R allele or an empty vector were unable to grow in liquid 5-FOA media. Wild-type *TYS1* expressing yeast entered mid log phase at ~ 90 hours. P167T *TYS1* expressing yeast had a 15-hour lag in growth before reaching mid-log phase compared to wild-type expressing yeast (Figure 2.12). This is consistent with the slower growth observed for P167T *TYS1* yeast on solid 5-FOA media. These data suggest that P167T *YARS* is a hypomorphic allele.

Table 2.7 Human *YARS* Variant Modeled in the Yeast Ortholog *TYS1*

Human YARS[a]	Yeast TYS1[b]
G41R	G45R
P167T	P171T

[a]Amino acid coordinates correspond to GenBank accession number AAH16689.1.
[b]Amino acid coordinates correspond to GenBank accession number AAB59329.1.

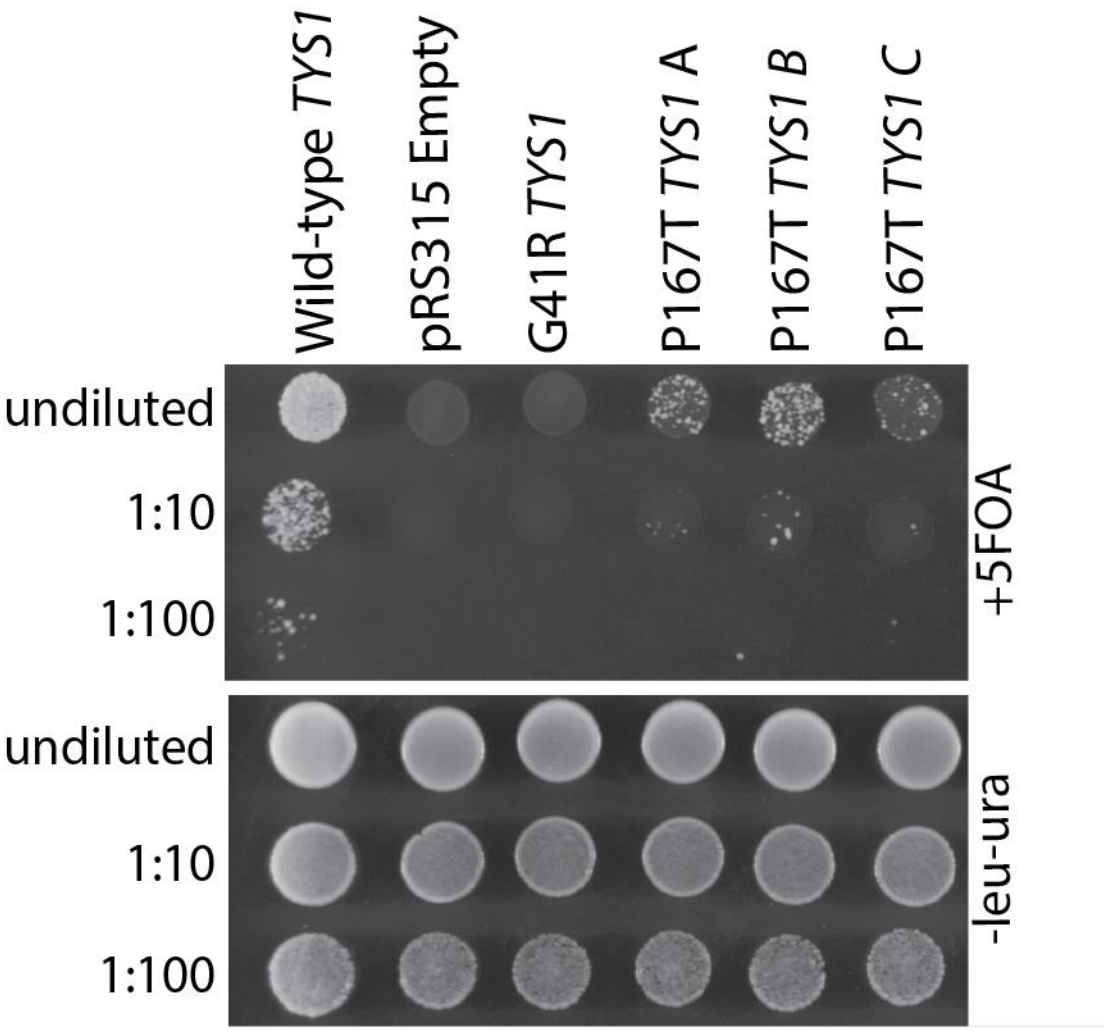


Figure 2.11. Yeast complementation assays of P167T *TYS1*. Yeast lacking endogenous *TYS1* were transformed with vectors containing wild-type, G41R, or P167T *TYS1* or an empty vector (pRS315 empty) and plated undiluted and diluted (1:10 and 1:100) on 5-FOA or –leu -ura media. Yeast expressing P167T *TYS1* showed an intermediate level of growth on 5-FOA media compared to the wild-type and the empty vector control yeast strains. G41R *TYS1*, a known complete loss-of-function allele in yeast (Jordanova et al. 2006), did not support yeast cell growth.

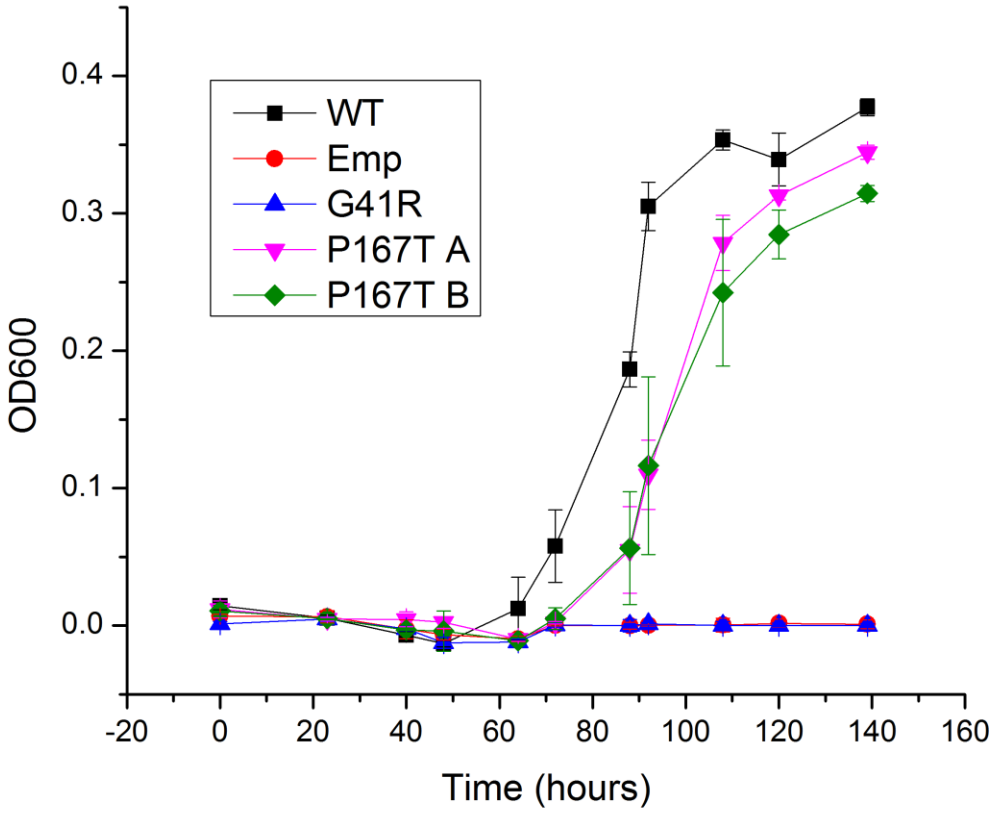


Figure 2.12. Growth curve analysis of P167T *TYS1*. Yeast expressing wild-type, G41R, or P167T *TYS1*, or no insert ('Emp') from a *LEU2*-bearing vector and carrying a *URA3* maintenance vector were grown in liquid 5-FOA media to select for the spontaneous loss of the *URA3* vector (Boeke et al. 1987) for 150 hours. OD_{600} readings were taken at the intervals indicated on the X-axis (hours). Error bars represent standard deviation of the average OD_{600} of three biological replicates for each strain. Note the ~15 hour lag between mid-log phase of wild-type and P167T *TYS1* expressing yeast.

Discussion

Here we present the functional characterization of mutations identified in patients who suffer from autosomal recessive syndromes with no known etiology. Recessive mutations in *AARS, GARS, DARS, FARSB,* and *YARS* were found to have varying levels of activity. K81T and R751G *AARS* and P167T *YARS* demonstrated loss-of-function characteristics in either aminoacylation or yeast complementation assays, while mutations in *GARS, FARSB* and *DARS* did not show reduced function when modeled in yeast. Thus, these data indicate a loss-of-function mechanism for recessive *AARS*- and *YARS*-associated syndromes, but is insufficient to implicate or refute a similar mechanism for *GARS, FARSB,* and *DARS* mutations.

Aminoacylation and yeast complementation assays have been used to evaluate ARS-mutations associated with a variety of human diseases (Antonellis et al. 2006; Jordanova et al. 2006; McLaughlin et al. 2010; McLaughlin et al. 2011; Pierce et al. 2011; Vester et al. 2012; Cassandrini et al. 2013; Gonzalez et al. 2013; Diodato et al. 2014; Griffin et al. 2014; Gonzaga-Jauregui et al. 2015; Hadchouel et al. 2015; Safka Brozkova et al. 2015). Direct comparison of results from these two assays has demonstrated that all mutations that show loss of function in yeast assays result in a dramatic reduction in aminoacylation activity *in vitro*. However, the converse is not true. A subset of mutations that support yeast cell growth show decreased activity in aminoacylation assays. For example, L133H *KARS,* which is implicated in a recessive syndrome that includes peripheral neuropathy and developmental delay, has a >90% decrease in aminoacylation studies, but supports yeast cell growth to wild-type levels (McLaughlin et al. 2010). This may indicate that the mutation has a greater impact on the activity of the human protein than the yeast ortholog or that the tissues affected in disease (*i.e.* neurons) are exquisitely sensitive to ARS dysfunction compared to yeast cells. Thus while decreased ARS activity

appears detrimental to a subset of specialized cells as evidenced by the tissue specific phenotypes observed in dominant and recessive ARS-mediated disease, decreased activity in yeast may not be sufficient to cause impaired growth. This may be due to the size of a yeast cell compared to a neuron (Wallen and Antonellis 2013) or the metabolic demands of yeast compared to specialized human tissues. There are two approaches to address this issue. The first is to perform aminoacylation assays on mutations that do not demonstrate loss of function in yeast complementation assays. Indeed, we have generated constructs to express human *GARS* and *DARS* mutant alleles and aminoacylation studies are underway. Alternatively, recent studies have demonstrated that deletion of some endogenous yeast loci can be rescued by the expression of the human ortholog (Jordanova et al. 2006; Chien et al. 2014; Kachroo et al. 2015). For example, deletion of *TYS1* was rescued by expression of human wild-type *YARS*, but not G41R *YARS* (Jordanova et al. 2006). Recently, it was shown that deletion of *GRS1* could be rescued by a plasmid expressing a truncated form of human *GARS* that lacks the mitochondrial targeting sequence and WHEP domain, two domains not conserved between human and yeast (Chien et al. 2014). Human mutations can be modeled in be this truncated form of GARS and assessed from complementation in a 'humanized' yeast assay. This would allow testing of disease-associated mutations for effects on the human protein *in vivo*. The humanized yeast assay would also expand the spectrum of mutations that can be evaluated, as conservation would no longer be a limiting factor. Indeed, R256Q *GARS* has already been tested in the humanized yeast assay and it is unable to support yeast viability compared to wild-type *GARS* (Stephanie Oprescu, personal communication). However, we would predict since *GARS* is critical for organismal viability and the compound heterozygous genotype E26Ifs6*/ R256Q *GARS* is not embryonic lethal, R256Q *GARS* must retain at least a low level of activity. Indeed, aminoacylation data show that R256Q

GARS activity is 1/920 of wild-type levels (Ya-ming Hou, personal communication). The detectable activity of R256Q *GARS* in the aminoacylation assays compared to the humanized yeast assay may be due to increased stabilization of the full length GARS protein compared to the truncated protein used to complement in yeast. Together these data indicate that R256Q *GARS* is a severe, but not complete, loss-of-function mutation.

Developing "humanized" yeast complementation assays for other ARS genes will be important for expanding our ability to assess variants for an effect on ARS enzyme function. It is currently unknown how many human ARS enzymes can complement in yeast. A recent study found that of 11 full-length ARS enzymes tested, seven (*KARS, AARS, QARS, CARS, SARS , LARS,* and *YARS*) were able to complement, while four (*DARS, FARSB, FARSA,* and *IARS*) were unable to support yeast viability (Kachroo et al. 2015). However, modifications, such as removal of non-conserved domains (*e.g.* the WHEP domain of GARS), may improve rescue by human genes and allow for complementation of most, if not all, human ARS genes in yeast.

Although a loss-of-function mechanism is likely the pathomechanism in recessive ARS-associated syndromes, it is unknown how these mutations lead to cell specific phenotypes. For example, *AARS* mutations that lead to hypomyelination may be due to an abnormality in myelin development or a delay in myelination due to a defect in the neuron. *GARS, AARS,* and *YARS* mutations have been implicated in dominantly inherited CMT disease. For *GARS* and *AARS* mutations, the resulting CMT disease is axonal in nature, indicating that dysfunction of these enzymes primarily affects the neurons, not the myelinating Schwann cells. Patients with *YARS* mutations have DI-CMT and demonstrate evidence of both neuron and Schwann cell dysfunction. However, future research into how these cells are specifically affected by ARS mutations is needed (see chapter 6).

Since mutations in some genes cause both recessive and dominant disorders, it appears that the level of ARS function in a cell is critical to the severity of disease. Mouse models of *Gars* mutations have shown a dose dependent spectrum of phenotypes. Heterozygosity for one of two loss-of-function missense *Gars* alleles leads to a peripheral neuropathy phenotype and homozygosity or compound heterozygosity of these alleles leads to an even more severe phenotype than what is seen in heterozygous state, including decreased life span (Seburn et al. 2006; Achilli et al. 2009; Motley et al. 2015). Thus, one would predict that heterozygous parents of the probands described in this chapter would be at risk for developing dominant CMT disease. In fact, family members who are heterozygous for the P167T *YARS* allele reported mild neuropathy symptoms. This provides the first evidence that both dominantly and recessively inherited disease can be due to the same loss-of-function ARS mutation and that the severity of disease is likely due overall reduction in enzyme function. Heterozygous parents carrying *AARS* and *GARS* mutations do not currently demonstrate symptoms of a peripheral neuropathy; however, because ARS-mediated CMT disease often presents late in life, these individuals should be continually screened for evidence of disease as they age. Based on the mouse studies of null and missense *GARS* alleles, we would hypothesize that individuals that carry the missense *AARS* and *GARS* alleles (K81T and R751G *AARS* and R256Q *GARS*) will develop late-onset axonal neuropathy, while the parent carrying a null allele (E29Ilefs6* *GARS*) would not develop disease. This would further support the hypothesis that an expressed stable, loss-of-function protein is necessary for CMT disease pathogenesis. Although mutations in *DARS* and *FARSB* have not been identified in patients with peripheral neuropathy, continued follow-up of these families would be informative in determining if additional ARS genes are capable of causing both recessive and dominant disease.

Future research on ARS-mediated recessive disorders should focus on improving ARS function in cells. One option would be to screen for pharmacologic compounds that improve mutant ARS function in yeast models. A second therapeutic avenue that should be considered is altering the kinetics of ARS aminoacylation reactions. This may be accomplished by providing more substrate, specifically the cognate amino acid in cases where the Km is increased, to improve activity of the mutant enzyme. Although a very simplistic approach, there are currently no other therapeutic options; therefore every conceivable treatment option should be explored. Altogether, understanding recessive ARS mutations not only provides opportunities to understand these rare diseases, but may also shed light into the mechanism of dominantly inherited ARS-mediated disease.

CHAPTER 3

Expanding the allelic heterogeneity of ARS mutations in dominant peripheral neuropathy

INTRODUCTION

Mutations in five ARS loci have been identified in patients with dominant peripheral neuropathy: *GARS, YARS, AARS, HARS,* and *MARS* (Antonellis et al. 2006; Jordanova et al. 2006; Latour et al. 2010; Vester et al. 2012; Gonzalez et al. 2013). However, only variants in *GARS, YARS,* and *AARS* have strong genetic evidence, specifically segregation studies in large pedigrees and identification of variants in multiple unrelated families, supporting their role in CMT disease (Antonellis et al. 2003; Jordanova et al. 2006; Latour et al. 2010; Lin et al. 2011; McLaughlin et al. 2011; Zhao et al. 2012). To date, 13 *GARS* mutations have been identified in patients with CMT2D or dSMA-V (James et al. 2006; Rohkamm et al. 2007; Abe and Hayasaka 2009; Lee et al. 2012a; Lee et al. 2012b), while only three *YARS* variants and three *AARS* variants have been implicated in DI-CMT and CMT2N, respectively. (Jordanova et al. 2006; Latour et al. 2010; Lin et al. 2011; McLaughlin et al. 2011; Zhao et al. 2012). While the functional consequences for many *GARS* variants remains uncertain (see chapter 4), biochemical and yeast complementation assays have demonstrated that all previously tested *YARS* and *AARS* mutations demonstrate loss of function in at least one of these assays (Jordanova et al. 2006; McLaughlin et al. 2011). Specifically, G41R and E196K *YARS* result in decreased activity in pyrophosphate assays, which

test for the completion of the first step in tRNA charging, and in yeast complementation assays (Jordanova et al. 2006). N71Y and R329H *AARS* both result in a severe reduction in aminoacylation and fail to support yeast cell growth in complementation assays (McLaughlin et al. 2011). Importantly, a non-disease associated mutation, E778A *AARS*, supported yeast viability and had normal aminoacylation kinetics. Taken together, the functional characterization of *YARS* and *AARS* mutations provide evidence that loss of function is a characteristic of disease-associated ARS alleles. However, additional mutations are needed to further support this hypothesis and allow for a greater understanding of the mechanism of ARS- mediated disease.

When this thesis work commenced, the genetic evidence for the pathogenic role of *HARS* and *MARS* variants in CMT disease was insufficient to implicate these loci. Only one *HARS* mutation (R137Q) had been identified by candidate gene sequencing of a single individual with late-onset, sporadic axonal peripheral neuropathy (Vester et al. 2012). When tested in yeast complementation assays, R137Q *HTS1*, the yeast ortholog of *HARS*, was unable to complement loss of the endogenous *HTS1*, consistent with the loss-of-function characteristics previously reported for other pathogenic ARS variants (Antonellis et al. 2006; Jordanova et al. 2006; McLaughlin et al. 2011; Stum et al. 2011). Additionally, R137Q *HARS* was neurotoxic when overexpressed in the GABA neurons of *C. elegans*, supporting its role in dominant neurotoxicity (Vester et al. 2012). However, this variant was also identified in a single individual included in the ClinSeq™ database (Biesecker et al. 2009) who showed no evidence of peripheral neuropathy. Given the late onset of the disease and known pattern of incomplete penetrance for some ARS mutations (Sivakumar et al. 2005), the inclusion of this variant in the ClinSeq™ database is not sufficient to exclude R137Q *HARS* as a pathogenic variant. However, the lack of segregation studies of the R137Q *HARS* prevented the implication of *HARS* as a CMT disease

gene (Vester et al. 2012).

Similar to *HARS*, only one *MARS* variant has been identified in patients with peripheral neuropathy. Specifically, exome sequencing identified R618C *MARS* in two individuals with peripheral neuropathy and an unaffected obligate carrier in a small pedigree with late-onset, incompletely penetrant axonal CMT disease (Gonzalez et al. 2013). Yeast complementation studies revealed that R618C *MARS* could not support yeast viability. While the genetic evidence is insufficient to implicate *MARS* in peripheral neuropathy, these data indicate *MARS* variants identified in individuals with peripheral neuropathy should be carefully evaluated for a role in the pathogenesis CMT disease (Gonzalez et al. 2013). Although the loss-of-function data for *HARS* and *MARS* variants is consistent with other disease-associated ARS variants, the identification of additional variants in patients with CMT disease is required to implicate these ARS loci in CMT disease.

ARS variants are increasingly being screened for and, subsequently, identified in patients with CMT disease. In the absence of strong genetic evidence, it is critical to functionally evaluate these variants before implicating them in disease for multiple reasons (MacArthur et al. 2014): First, identification of a novel ARS variant in an individual with CMT disease is not sufficient to implicate the variant in disease. As discussed above, loss-of-function characteristics have proven to be the best predictor of pathogenesis for ARS variants (Wallen and Antonellis 2013); however, since the pathomechanism of disease is still unknown, the question remains as to whether loss of function is the sole indicator of pathogenicity. Thus, functional characterizations of new variants will not only help establish a pipeline for implicating rare or *de novo* ARS mutations in CMT disease, but also will provide evidence for the mechanism of ARS-mediated pathogenesis. Expanding the allelic heterogeneity of ARS mutations will not only provide

critical information for patients and clinicians to consider, but also provide researchers with both active and loss-of-function alleles for use in future studies into the mechanism of ARS-mediated CMT disease. Second, identification of new variants in patients with CMT disease is necessary to provide the genetic evidence to implicate *HARS*, *MARS*, and any additional ARS enzymes not previously associated CMT disease. Third, characterization of novel ARS mutations in patients with a variety of peripheral neuropathy phenotypes (*e.g.* axonal versus demyelinating disease; peripheral neuropathy versus myeloneuropathy) would expand our understanding of the spectrum of disease manifestations for ARS-mediated disease; Finally, implication of new ARS loci in CMT disease would provide valuable information as to the ARS characteristics necessary for disease pathogenesis. For example, the three strongly implicated ARS enzymes (GARS, YARS, and AARS), as well as HARS, all function as dimers. However, MARS functions as a monomer. Thus further research may support or refute a role for oligomerization in the mechanism of ARS-associated CMT disease.

This chapter focuses on the identification, genetic validation, and functional characterization of newly identified mutations in *YARS*, *AARS*, and *HARS* in patients with dominant peripheral neuropathy. Mutations were identified using various sequencing techniques, evaluated for segregation in pedigrees when available, screened for inclusion in general population databases, assessed for conservation, and functionally characterized based on previous studies evaluating ARS mutations. We hypothesize that pathogenic ARS variants will segregate with disease in pedigrees, be absent or present at very low frequency in the general population, affect highly conserved amino acid residues, and demonstrate loss-of-function characteristics in yeast complementation assays.

A subset of data in this chapter was previously published in *Cell Reports*, *Brain*, and *Neurology* (Gonzaga-Jauregui et al. 2015; Motley et al. 2015; Safka Brozkova et al. 2015). Permission was requested to reproduce figures. The author performed all the studies in this chapter with the following exceptions: (1) Clinical collaborators collected the phenotypic data and performed sequencing and segregation studies for the mutations presented; (2) Haploid Δ*ALA1* and Δ*HTS1* yeast strains were generated by Heather McLaughlin (McLaughlin et al. 2011) and Aimee Vester (Vester et al. 2012), respectively; and (4) The University of Michigan Sequencing core performed the DNA sequencing reactions described in this work.

MATERIALS AND METHODS

Patient sample collection, identification of ARS mutations, and linkage studies

Clinical collaborators performed the clinical evaluation, classified patient phenotypes based on clinical exam findings, and collected DNA samples with patient consent as per the Institutional Review Board protocols from the participating institutions. *YARS* and *HARS* variants were identified via whole exome sequencing. *AARS* variants were identified upon candidate gene sequencing of disease-associated ARS loci. Genetic studies were performed to determine segregation. Each identified ARS mutation was assessed for its presence in the general population via database searches of the Exome Variant Server (snp.gs.washington.edu/EVS), the 1000 Genomes Project database (www.1000genomes.org), and dbSNP (www.ncbi.nlm.nih.gov/snp). Conservation of the affected amino acid residues was determined by comparison of evolutionarily divergent species using ClustalW2 (Larkin et al. 2007).

Generation of yeast ARS expression constructs and haploid yeast strains

PCR primers were designed to amplify the orthologous *S. cerevisiae* ARS locus (Table 3.1) including the endogenous promoter from yeast genomic DNA. Gateway (Invitrogen) cloning

sequences (attB1 (forward) and attB2 (reverse); primer sequences in Appendix A) were included to facilitate recombination into the pDONR221 vector. Subsequent to PCR, BP Clonase (Invitrogen) reactions were employed to clone the wild-type ARS locus into pDONR221 and resulting colonies were fully sequenced to rule out PCR-induced errors. Next, mutation-containing oligonucleotides (primer sequences in Appendix A) were employed to model the human mutations at the corresponding residue in the orthologous yeast gene (Table 3.1). The QuickChange II XL Site-Directed Mutagenesis Kit (Stratagene) was used to mutate each locus in the pDONR221 entry construct. Resulting clones were purified and sequenced to confirm successful mutagenesis and rule out PCR-induced errors. The wild-type and mutant ARS/pDONR221 entry clones were subsequently recombined into a Gateway-compatible *URA3*-bearing pRS316 or *LEU2*-bearing pRS315 destination vector using LR Clonase (Invitrogen). Resulting clones were purified and digested with *Bsr*GI (New England Biosystems) to confirm successful recombination.

Commercially available diploid heterozygous yeast strains were purchased for each ARS ortholog of interest (Open Biosystems). The strains utilized were: (1) *YARS*: diploid heterozygous Δ*TYS1* yeast strain (*MATa/α, his3Δ1/ his3Δ1, leu2Δ0/ leu2Δ0, met15Δ0/ MET15, ura3Δ0/ ura3Δ0*); (2) *AARS*: diploid heterozygous Δ*ALA1* yeast strain (*MATa/α, his3Δ1/ his3Δ1, leu2Δ0/ leu2Δ0, met15Δ0/ MET15, ura3Δ0/ ura3Δ0*) (McLaughlin et al. 2011); and (3) *HARS*: diploid heterozygous Δ*HTS1* yeast strain (*MATa/α, his3Δ1/ his3Δ1, leu2Δ0/ leu2Δ0, met15Δ0/ MET15, ura3Δ0/ ura3Δ0*) (Vester et al. 2012). Each strain was transformed with a *URA3*-bearing pRS316 vector containing the wild-type ARS locus (see above). Lithium acetate yeast transformations were performed at 30°C using 200 ng of plasmid DNA. Transformed yeast were grown on yeast media lacking uracil. Resulting haploid yeast containing the wild-type pRS316

maintenance vector were treated to induce sporulation. Briefly, the diploid strain was patched twice onto GNA presporulation plates (5% D-glucose [Fisher Scientific], 3% nutrient broth [Becton, Dickinson and Company], 1% yeast extract [Acros Organics], and 2% agar [Teknova]). One microliter of cells was then transferred into 2 ml of supplemented liquid sporulation medium (1% potassium acetate [Fisher Scientific], 0.005% zinc acetate [Fisher Scientific], 1X ura supplement, 1X his supplement, and 1X leu supplement) and incubated for 5 days at 25°C followed by 3 days at 30°C. Sporulated yeast were dissected using a MSM 400 dissection microscope (Singer Instruments) and plated on yeast extract, peptone, and dextrose (YPD) plates (Becton, Dickinson and Company). Resulting spores were individually patched onto solid growth medium containing Geneticin (G418) or 0.1% 5-fluoroorotic acid (5-FOA), or media lacking uracil (Teknova, Hollister, CA). Two spores that grew on G418 and yeast medium lacking uracil, but failed to grow on 5-FOA medium, were selected for use in the yeast viability assays.

Yeast complementation assays

Two haploid ΔARS strains [harboring a maintenance vector to express wild-type ARS and *URA3*] per locus were then transformed with an empty vector ('Empty pRS315' in figures) or the appropriate wild-type or mutant ARS in a *LEU2*-bearing pRS315 vector and selected on medium lacking uracil and leucine (Teknova). For each transformation, at least two independent pRS315 plasmid preparations were used and at least two colonies from each plasmid were selected for additional analysis and grown to saturation for 2 days at 37°C in liquid media lacking uracil and leucine. A 10 μl aliquot of each culture was spotted undiluted or diluted 1:10 or 1:100 in H_2O onto plates containing 0.1% 5-FOA (Teknova) or medium lacking uracil and leucine and incubated at 30°C for 48-72 hours. Survival was determined by visual inspection of growth.

Table 3.1. Human and yeast orthologous ARS locus

Human ARS	Yeast ARS
AARS	*ALA1*
HARS	*HTS1*
YARS	*TYS1*

Yeast growth curves

Growth curve assays were performed on pseudo-heterozygous yeast containing both wild-type *TYS1* on the *URA3*-bearing maintenance vector and wild-type, G41R, or E196Q *TYS1* or empty *LEU2*-bearing vector (see above). Three yeast colonies from each transformation were selected for growth curves. Yeast were grown to saturation for 2 days at 30°C in liquid media lacking uracil and leucine. Saturated cultures were diluted to an OD_{600} = 0.200 +/- 0.015 based on OD_{600} readings recorded on the Absorbance module E7061 on the Glomax®-Multi Microplate Multimode Reader Detection System (Promega, Madison, WI). 20 µl of OD_{600} = 0.2 +/- 0.015 cultures was added to 4 mL of media lacking uracil and leucine or media contain 0.1% 5-FOA (Teknova) and incubated at 30°C shaking. The average OD600 reading from the three biological replicate cultures for each strain were used to generate the growth curves based on the average OD_{600} measurements at the indicated time points. Error bars represent standard deviation. Growth curves were performed twice on yeast generated from independent transformations.

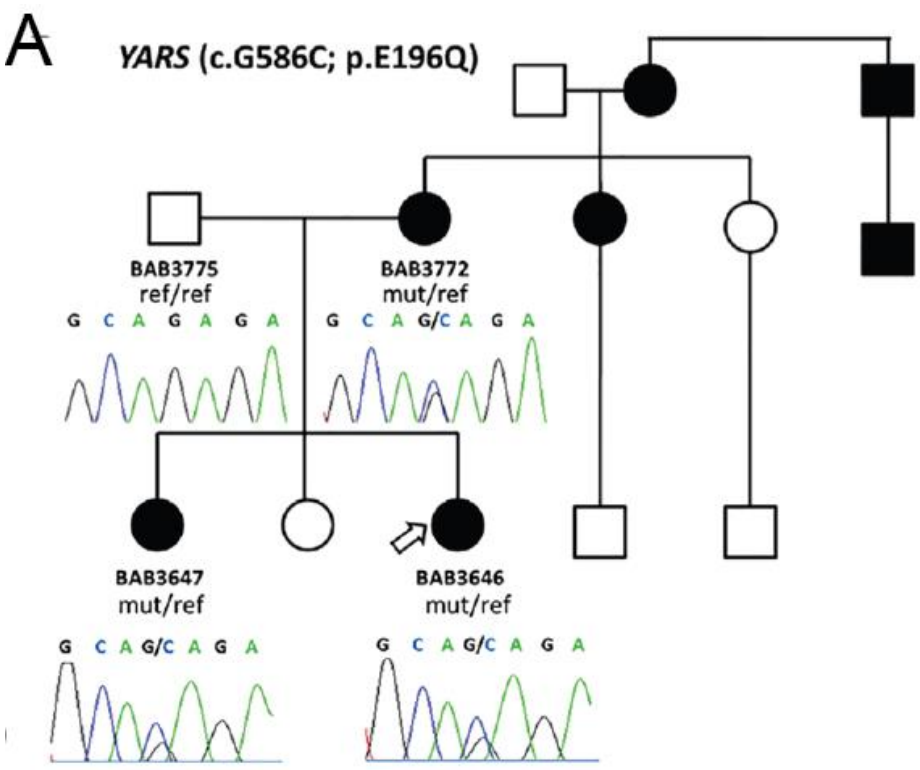
RESULTS

Identification and functional characterization of ARS variants

High-throughput sequencing of DNA from patients with dominant peripheral neuropathy is rapidly identifying variants in ARS genes. Recently, novel variants in *YARS*, *AARS*, and *HARS* were identified; however these variants remained classified as variants of unknown significance until genetic and functional studies were performed to assess their role in disease pathogenesis.

YARS *mutation identified in a family with DI-CMT.*

E196Q [c.568G>C; p.Glu196Gln] *YARS* segregates in a three-generation family affected by autosomal dominant intermediate CMT (DI-CMT; CMTDIC; MIM #608323) (Figure 3.1A).



B

YARS E196Q

Human	RKIFTFA**E**KYLPALG
mouse	RKIFTFA**E**KYLPALG
worm	RKIFILA**E**EQLPKLK
yeast	RKIFVLA**E**ENLPSLG
bacteria	----------

Figure 3.1. Segregation and conservation of E196Q *YARS*. (A) Pedigree of a the family affected with dominant intermediate CMT disease is depicted. All three affected family members available for genetic testing carry the E196Q *YARS* variant. Traces from Sanger sequencing confirm the presence of E196Q *YARS* mutation in affected individuals. Empty symbols indicate unaffected individuals and filled symbols indicate affected individuals. (B) The YARS protein sequence fragment surrounding residue E196 is aligned using multiple, evolutionarily diverse species. The amino acid change is listed at the top. The affected amino acid is in bold, red text.

E196Q *YARS* affects a residue in the catalytic domain of YARS that is conserved between human and yeast (Figure 3.1B) and was not identified in any general population databases.

E196Q YARS *demonstrates decreased function in yeast.*

Two alleles, G41R and E196K *YARS*, were previously associated with DI-CMT and were characterized using biochemical and yeast complementation assays as a functionally null and hypomorphic allele, respectively (Jordanova et al. 2006). The newly identified mutation reported here, E196Q *YARS*, affects the same residue as E196K *YARS*. To determine the effects of E196Q *YARS in vivo*, we developed a haploid Δ*TYSI* (the yeast ortholog of *YARS*) yeast strain bearing a *URA3*-maintenance vector containing wild-type *TYS1* to use in yeast complementation assays. We modeled G41R, E196K, and E196Q *YARS* mutations in the corresponding orthologous residues in *TYS1* (Table 3.2) in a *LEU2*-bearing vector and transformed yeast with wild-type or mutant *TYS1* or an empty vector control. Our results indicate that the G41R variant is a loss-of-function allele, as previously shown (Jordanova et al. 2006), while both variants affecting E196, the previously reported E196K and our newly identified E196Q variant, are hypomorphic alleles (Figure 3.2). Specifically, the G41R allele is unable to complement loss of endogenous *TYS1*, similar to what is observed for yeast transformed with an empty vector ('Empty pRS315') (Figure 3.2). Furthermore, compared to yeast expressing wild-type *TYS1*, expression of E196K and E196Q *TYS1* are unable to fully complement deletion of *TYS1*, resulting in a reduction but not complete abrogation of cell growth. E196Q expressing yeast grew less robustly than E196K expressing yeast (Figure 3.2). These data indicate that E196K and E196Q are both hypomorphic alleles and are consistent with the previous functional characterization of E196K (Jordanova et al. 2006). To further quantify the decreased growth observed for yeast expressing E196Q *TYS1*,

Table 3.2. Human *YARS* Variants Modeled in the Yeast Ortholog *TYS1*

Human YARS[a]	Yeast TYS1[b]
G41R	G45R
E196K	E200K
E196K	E200Q

[a]Amino-acid coordinates correspond to GenBank accession number AAH16689.1.
[b]Amino-acid coordinates correspond to GenBank accession number AAB59329.1.

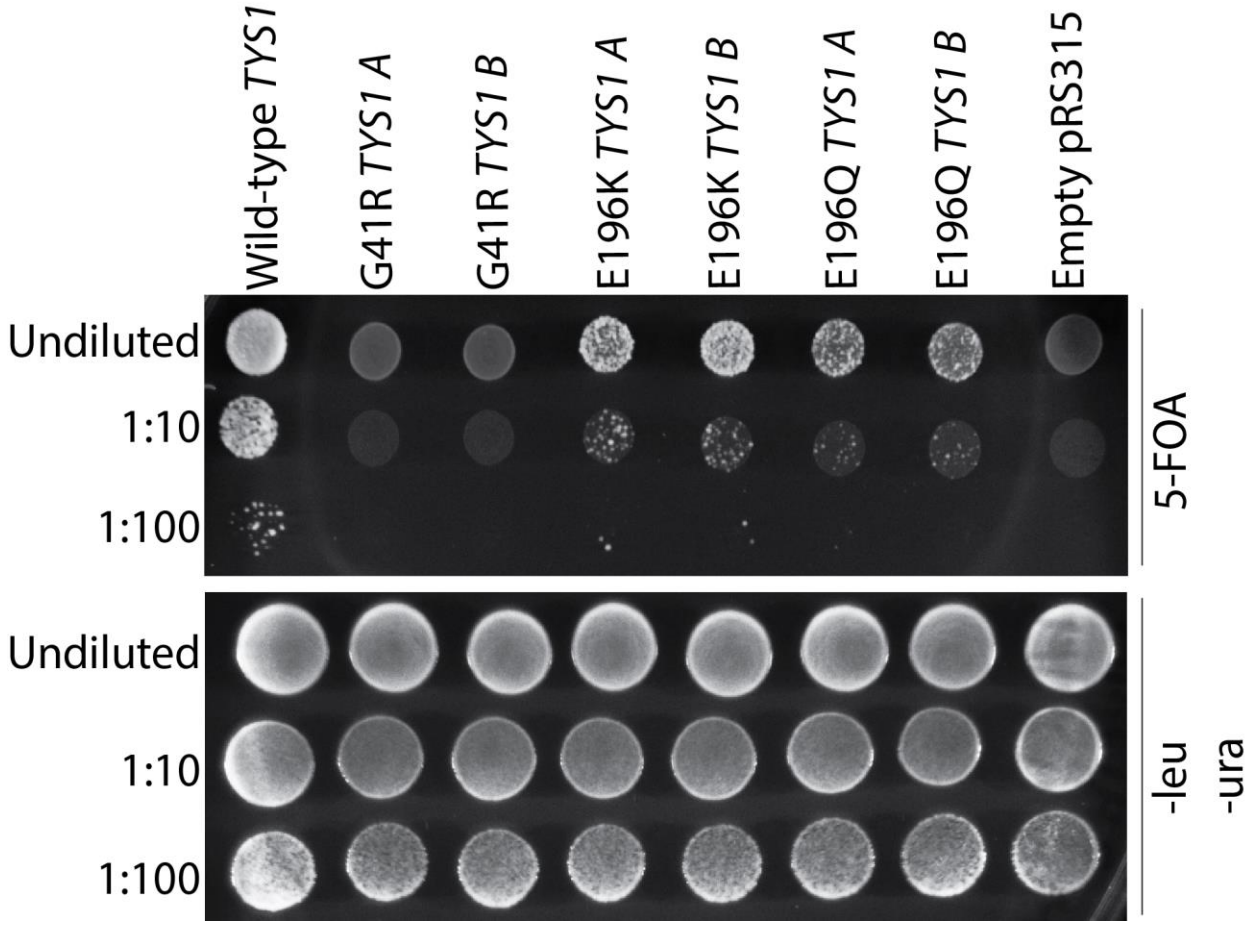


Figure 3.2. Yeast complementation assay for E196Q *TYS1*. A haploid yeast strain with a deletion of *TYS1* and bearing a *URA3* maintenance vector was transformed with a *LEU2*-bearing pRS315 vector containing wild-type *TYS1*, the indicated mutant form of *TYS1*, or no insert ('Empty pRS315') (labeled along the top). Cultures for each strain were grown for 2 days and plated on medium containing 5-FOA to determine if the *TYS1* alleles complement loss of endogenous *TYS1* or on –leu -ura medium as a positive growth control. Two independently generated constructs (indicated as 'A' and 'B' across the top) were evaluated for each mutation. Note the severe reduction of growth for E196K and E196Q at 1:10 and 1:100 dilutions.

growth curve assays were performed in liquid 5-FOA selective media comparing wild-type, G41R, and E196Q *TYS1* and empty vector containing strains. Growth curves of yeast expressing G41R *TYS1* did not grow in liquid 5-FOA media, consistent with the G41R *TYS1* yeast phenotype observed on solid media (Figure 3.3). Yeast carrying the E196Q *TYS1* allele demonstrated slower growth than wild-type *TYS1*, with a delay of ~15 hours between the mid-log phase of wild-type and E196Q yeast strains (Figure 3.3). These studies indicate that E196Q *YARS* is a functional hypomorph and, in conjunction with the segregation studies, supports a pathogenic role for E196Q *YARS* in DI-CMT disease.

AARS mutation identified in a family with myeloneuropathy.

Four siblings with clinical and electrophysiologic evidence of mild, inherited peripheral neuropathy were found to be heterozygous for the missense mutation G102R [c.304G>C; p.Gly102Arg] *AARS* (Figure 3.4A and B). Interestingly, 3 of 4 siblings had lower extremity hyperreflexia indicative of myelopathy, a neurologic deficit involving the spinal cord. A sural nerve biopsy from one sibling demonstrated evidence of chronic axonal loss. Importantly, the G102R *AARS* allele was not identified in control patients or in general population databases. G102R *AARS* is located in the activation domain of the AARS enzyme and is conserved from human to bacteria (Figure 3.4C).

G102R AARS *does not complement in yeast.*

To determine if G102R *AARS* alters the function of the AARS protein, we employed a yeast complementation assay previously used to assess the activity of CMT-associated R329H and N71Y *AARS* mutations (McLaughlin et al. 2011). Using this assay, we modeled G102R and R329H *AARS* in the yeast ortholog *ALA1* (Table 3.3) and transformed the Δ*ALA1* yeast strain,

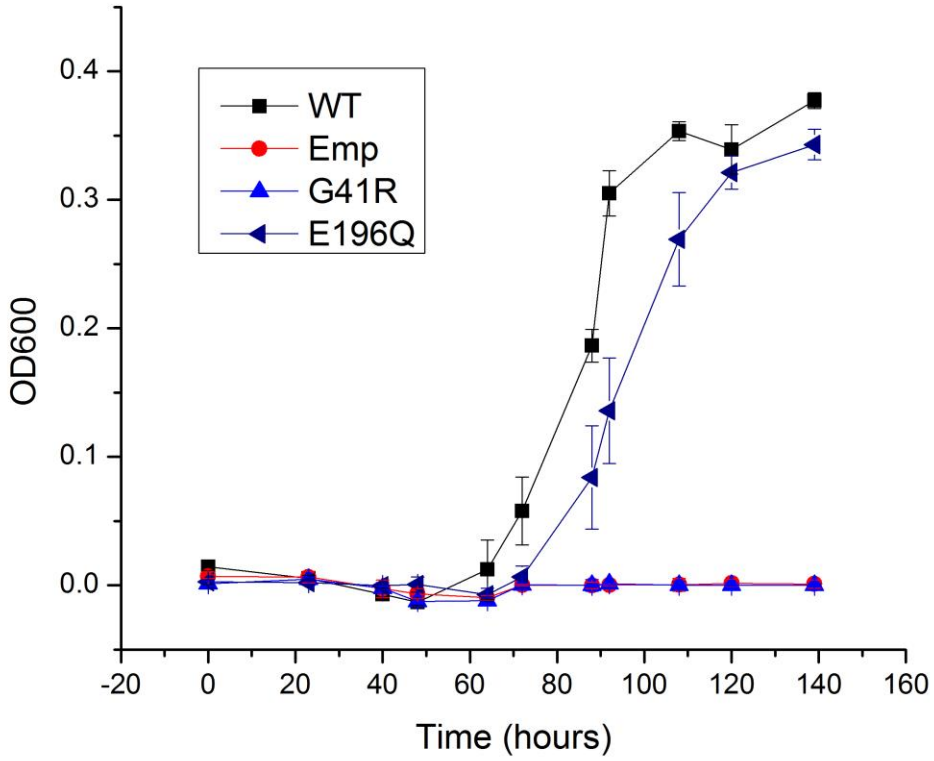


Figure 3.3. Growth curves of yeast expressing E196Q *TYS1*. Pseudo-heterozygous yeast carry both the wild-type *TYS1* gene on a *URA3*-bearing vector and wild-type, G41R, or E196Q *TYS1* or an empty *LEU2*-bearing vector were normalized and grown at 30°C in 0.1% FOA liquid media to select for loss of the *URA3* maintenance vector (Boeke et al. 1987). Growth was assessed via OD_{600} absorbance readings measured at indicated time points. Each time point represents the average of three biological replicates. Error bars indicate standard deviation. Note the ~15 hour lag between mid-log phase for yeast carrying E196Q *TYS1* compared to yeast carrying the wild-type *TYS1* allele.

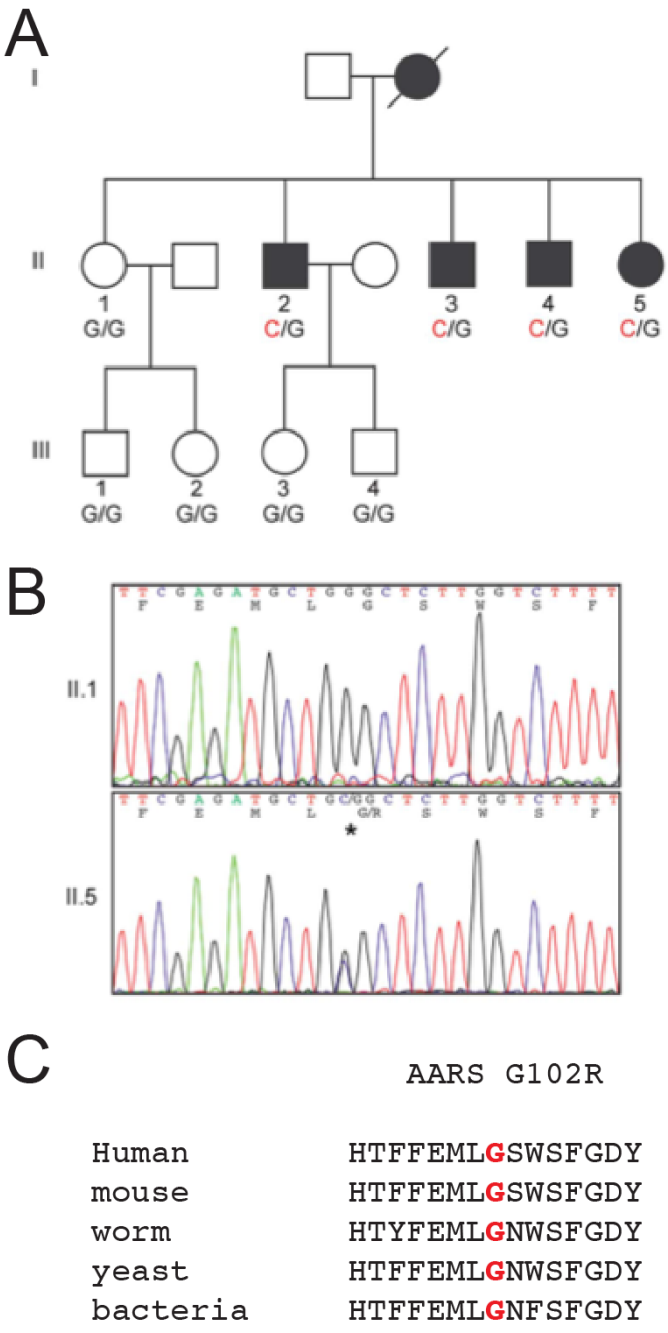


Figure 3.4. Segregation and conservation of G102R *AARS*. (A) Pedigree of a family with myeloneuropathy. All four affected siblings are heterozygous for G102R *AARS*, as indicated by the red base below their respective symbol. Empty symbols indicate unaffected individuals and filled symbols indicate affected individuals. (B) Representative traces from Sanger sequencing confirm the presence of G102R *AARS* in an affected individual (II.5) and the absence of the mutation in her unaffected sister (II-1). (C) A fragment of the activation domain sequence that contains G102 (red, bold text) is aligned in divergent species.

Table 3.3. Human *AARS* Variants Modeled in the Yeast Ortholog *ALA1*

Human AARS[a]	Yeast HTS1[b]
G102R	G106R
R329H	R329H

[a]Amino-acid coordinates correspond to GenBank accession number BAA06808.1.
[b]Amino-acid coordinates correspond to GenBank accession number EDV1089.

which has the endogenous *ALA1* gene deleted and a vector that expresses wild-type *ALA1* and *URA3* to maintain viability, with a *LEU2*-bearing vector containing either wild-type, G102R, or R329H *ALA1* or no insert ('Empty pRS315'). When plated on selection media containing 5-FOA to select for yeast that have spontaneously lost the maintenance vector (Boeke et al. 1987), the wild-type *ALA1* expression vector sustained yeast viability, while the empty vector was unable to complement the knockout allele, consistent with *ALA1* being an essential gene (Figure 3.5). Yeast expressing G102R or R329H *ALA1* were unable to survive on 5-FOA media (Figure 3.5). These results illustrate that G102R *ALA1* is unable to complement loss of endogenous *ALA1* and suggests that G102R *AARS* is a loss-of-function *AARS* allele.

HARS *mutations were identified in seven families with peripheral neuropathy.*

Genetic studies have identified seven *HARS* mutations that each segregate in seven unrelated families with dominantly inherited peripheral neuropathy. T132I [c.395C>T; p.Thr132Ile] *HARS* was identified in a German family (Figure 3.6, Family A) with phenotypic and electrophysiological characteristics consistent with axonal neuropathy. Interestingly, this mutation affects the same amino acid residue as a rare single nucleotide polymorphism (SNP), T132S *HARS* (rs143473232). T132S is present in 1/13,006 alleles surveyed in the Exome Variant Server (EVS). A Moroccan family (Figure 3.6, Family B) was diagnosed with intermediate CMT due to MNCVs that correspond to axonal CMT, demyelinating CMT, or both axonal and demyelinating CMT disease in individual family members. All affected individuals, regardless of the type of CMT disease indicated by electrophysiologic studies, were found to be heterozygous for the same *HARS* variant, P134H [c.401C>A; p.Pro134His]. Two additional mutations, D175E [c.525T>G; p.Asp175Glu] *HARS* and D364Y [c.1090G>T; p.Asp364Tyr] *HARS* were identified in a Czech-Belgian (Figure 3.6; Family C) and Belgian family (Figure 3.6;

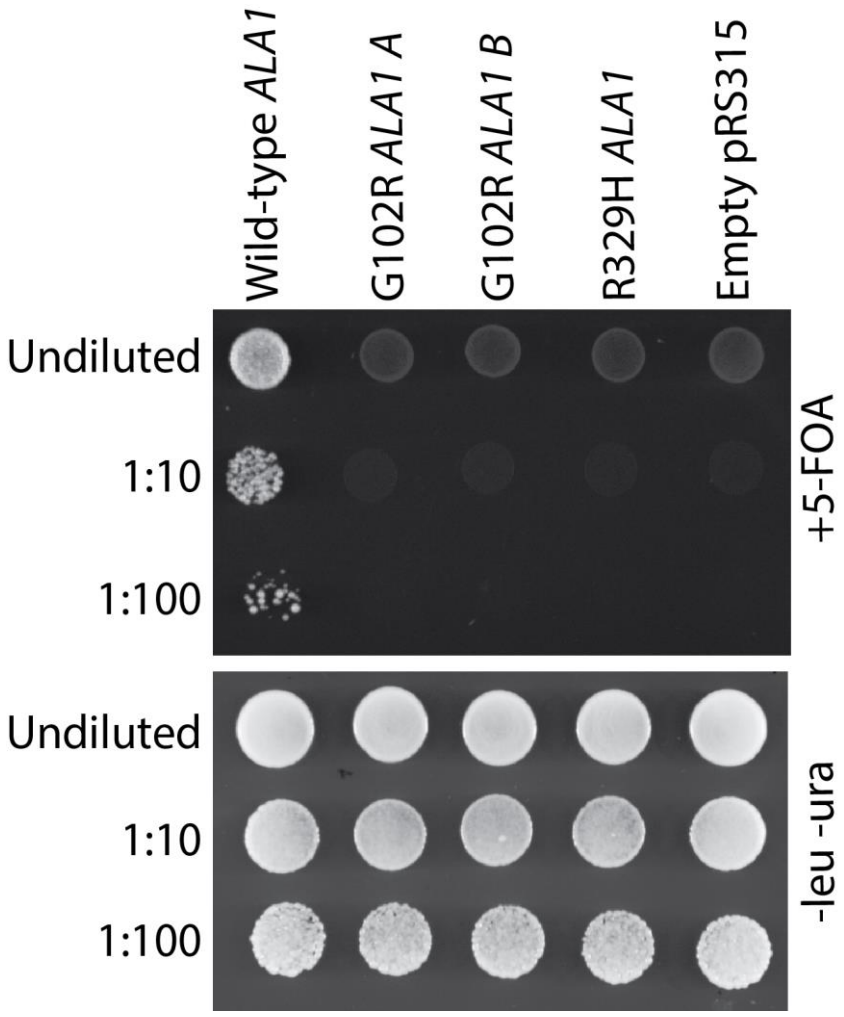


Figure 3.5. Yeast complementation assays for G102R *ALA1*. Haploid Δ*ALA1* yeast strains were transformed with a vector containing no insert (pRS315 Empty) or an insert to express wild-type, G102R or R329H *ALA1* (labeled across the top panel). Cultures resulting from each transformation condition are spotted undiluted and diluted (1:10 and 1:100) on plates containing 5-FOA complete medium or SD -leu -ura growth medium. Experiments were performed using two independently generated *ALA1* expression constructs (labeled A and B) for G102R.

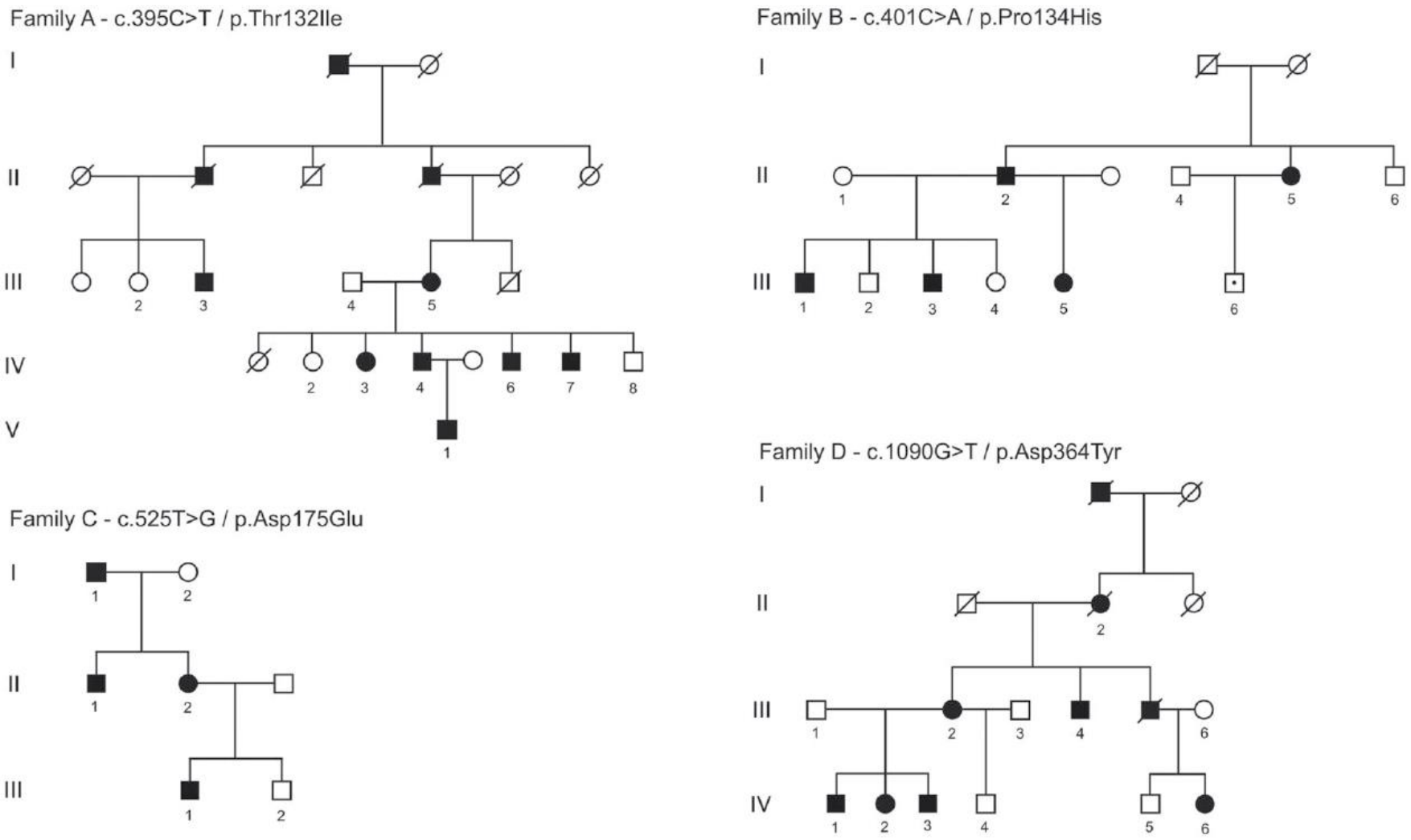


Figure 3.6. Segregation of *HARS* mutations in families with peripheral neuropathy. The mutation identified in each family is listed above the corresponding pedigree. Circles indicate female family members, while male family members are represented in squares. Empty symbols represent unaffected individuals, while filled symbols indicate affected individuals. Generations are indicated in Roman numerals. Arabic numerals are used to identify each genotyped family member. In all pedigrees, the mutation of interest was identified in all affected members and was absent in unaffected individuals.

Family D), respectively, each with family history and electrophysiologic studies consistent with dominantly inherited axonal CMT disease. V155G [c.464A>C is p.Val155Gly] *HARS* was identified in five affected individuals in a three generation family with a pure motor neuropathy and evidence of myelopathy (pedigree structure currently unavailable). Y330C [c.989A>G is p.Tyr330Cys] *HARS* was identified in a mother-son pair with sensory and motor neuropathy with normal MNCVs and decreased amplitudes consistent with a diagnosis of axonal CMT disease. Finally, S356N [c.1067G>U is p.Ser356Asn] *HARS* was identified in a woman with peripheral neuropathy and her mother who is unaffected, indicating a potential for incomplete penetrance in this family. All seven affected residues are conserved between human and yeast, with Y330C and D364Y being conserved to bacteria (Figure 3.7).

The majority of disease-associated HARS *mutations demonstrate loss of function in yeast.*
As a pattern of loss-of-function characteristics is becoming increasingly evident for CMT-associated ARS mutations (Wallen and Antonellis 2013). Yeast complementation assays can be applied to functionally assess variants of unknown significance in ARS genes not previously implicated in disease to determine if these variants demonstrate characteristics consistent with disease-associated alleles. This approach was used to study R137Q *HARS*, a variant that was identified in a patient with sporadic axonal CMT and demonstrated loss-of-function similar to other CMT-associated ARS alleles (Vester et al. 2012). To test the functional consequences of the seven conserved *HARS* missense variants identified in the families described above, we modeled these missense variants in the yeast ortholog *HTS1* (Table 3.4). Additionally, we modeled the rare variant T132S *HARS*, which is listed in dbSNP but has no known association with peripheral neuropathy. This variant affects the same residue as T132I *HARS* that was identified in this study. We independently tested each mutation for the ability to support yeast

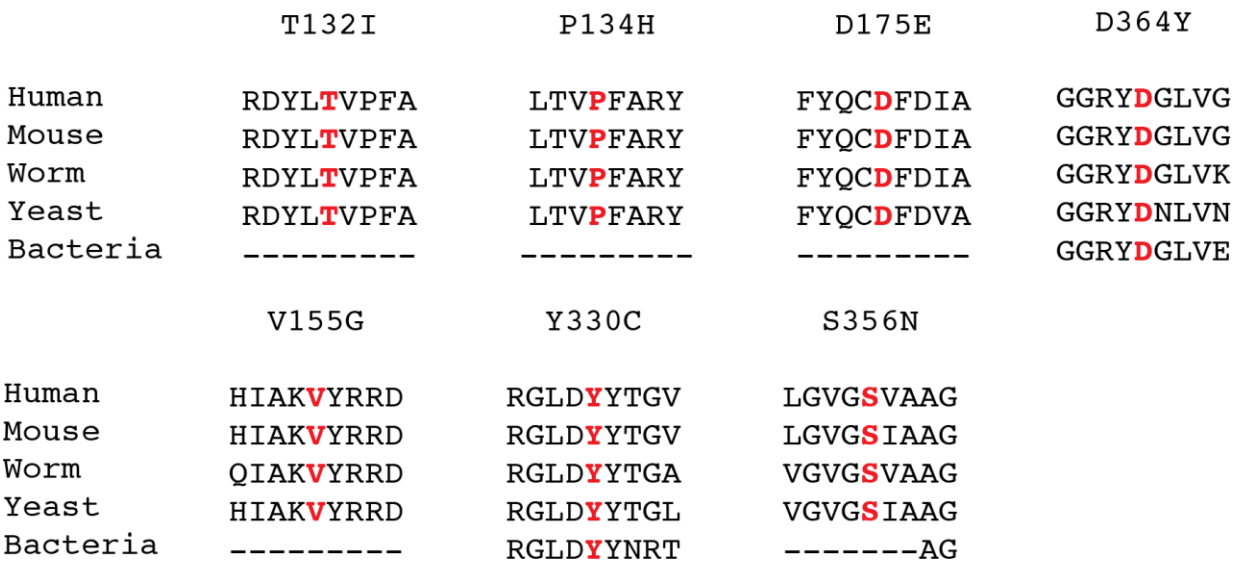


Figure 3.7. Conservation of *HARS* mutations. Protein sequence fragments from multiple, evolutionarily diverse species were aligned to assess for conservation of the amino acid residues affected by *HARS* mutations. The amino acid alteration is listed above the protein fragment. The affected amino acid is in bold, red text.

Table 3.4. Human *HARS* Variants Modeled in the Yeast Ortholog *HTS1*

Human HARS[a]	Yeast HTS1[b]
T132I	T131I
T132S	T131S
P134H	P133H
V155G	V154G
D175E	D174E
Y330C	Y330C
S356N	S370N
D364Y	D378Y

[a]Amino acid coordinates correspond to GenBank accession number NP_002100.2.
[b]Amino acid coordinates correspond to GenBank accession number AAA34696.1.

cell growth compared to wild-type *HTS1* or an empty vector. Briefly, a haploid yeast strain (with the endogenous *HTS1* locus deleted and a maintenance vector to express wild-type *HTS1* and *URA3*) was transformed with either a pRS315 vector with no insert ('Empty pRS315') or a pRS315 vector harboring wild-type, T132I, T132S, P134H, D175E, D364Y, V155G, Y330C, or S356N *HTS1*. Yeast cells were then selected on media containing 5-FOA, which is toxic to yeast carrying the *URA3*-bearing maintenance vector (Boeke et al. 1987). Thus, only yeast cells expressing a functional *HTS1* allele from the pRS315 vector will grow in this assay.

Yeast transformed with a wild-type *HTS1* expression vector demonstrated significant growth, while those transformed with an empty vector did not (Figure 3.8; Figure 3.9), consistent with *HTS1* being an essential gene (Vester et al. 2012). Regarding the newly identified *HARS* mutations described here, yeast expressing T132I, P134H, D175E, or D364Y *HTS1* were unable to grow on 5-FOA media (Figure 3.8) indicating that these are complete loss-of-function alleles. Yeast expressing Y330C or S356N *HTS1* had a pronounced decrease, but not complete abrogation, of growth at 1:10 and 1:100 dilutions compared to wild-type *HTS1* (Figure 3.9), indicating that these are partial loss-of-function alleles. In contrast, V155G *HTS1* was able to support yeast cell growth to levels comparable to wild-type *HTS1* (Figure 3.9). Finally, unlike the disease-associated variant T132I, the rare variant T132S *HTS1* supported yeast growth to the same extent as wild-type *HTS1* (Figure 3.8) indicating that this variant has no significant effect on *HARS* activity in yeast. With the exception of V155G *HARS*, these data suggest that all *HARS* variants that segregate with disease are loss-of-function mutations.

DISCUSSION

This chapter describes the identification and functional characterization of *YARS*, *AARS*, and *HARS* mutations identified in patients with in peripheral neuropathy. The genetic and functional

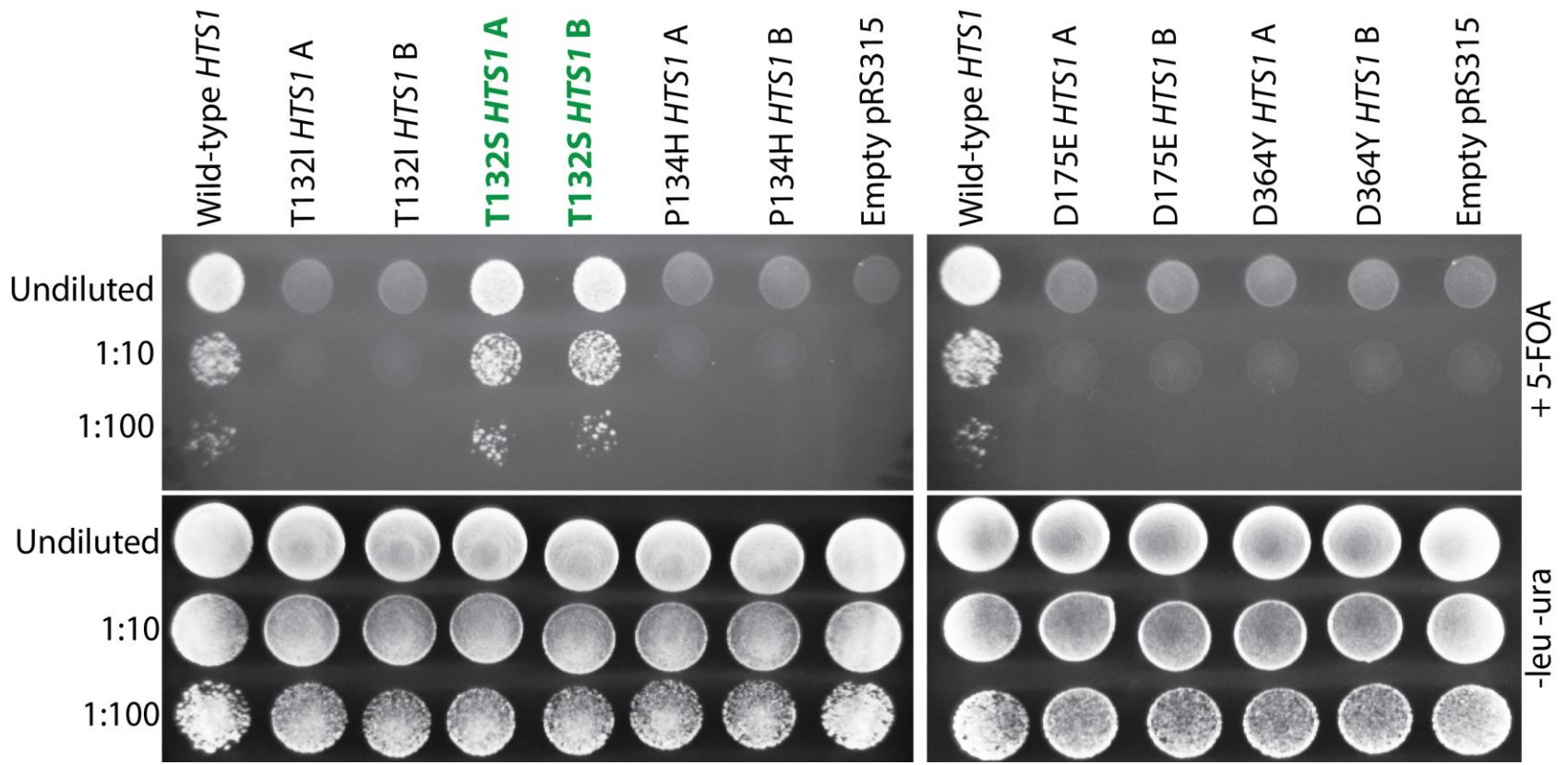


Figure 3.8. Yeast complementation assay of *HTS1* variants. Haploid *ΔHTS1* yeast strains were transformed with a vector containing no insert (pRS315 Empty) or an insert to express wild-type, T132I, T132S, P134H, D175E, or D364Y *HTS1*. Two colonies (indicated by 'A' and 'B') from transformations with T132I, T132S, P134H, D175E, or D364Y *HTS1* are shown. Resulting colonies (undiluted, diluted 1:10, or diluted 1:100) were grown on agar plates containing 0.1% 5-FOA. Note the wild-type level of growth associated with the non-disease associated allele T132S *HTS1* (green).

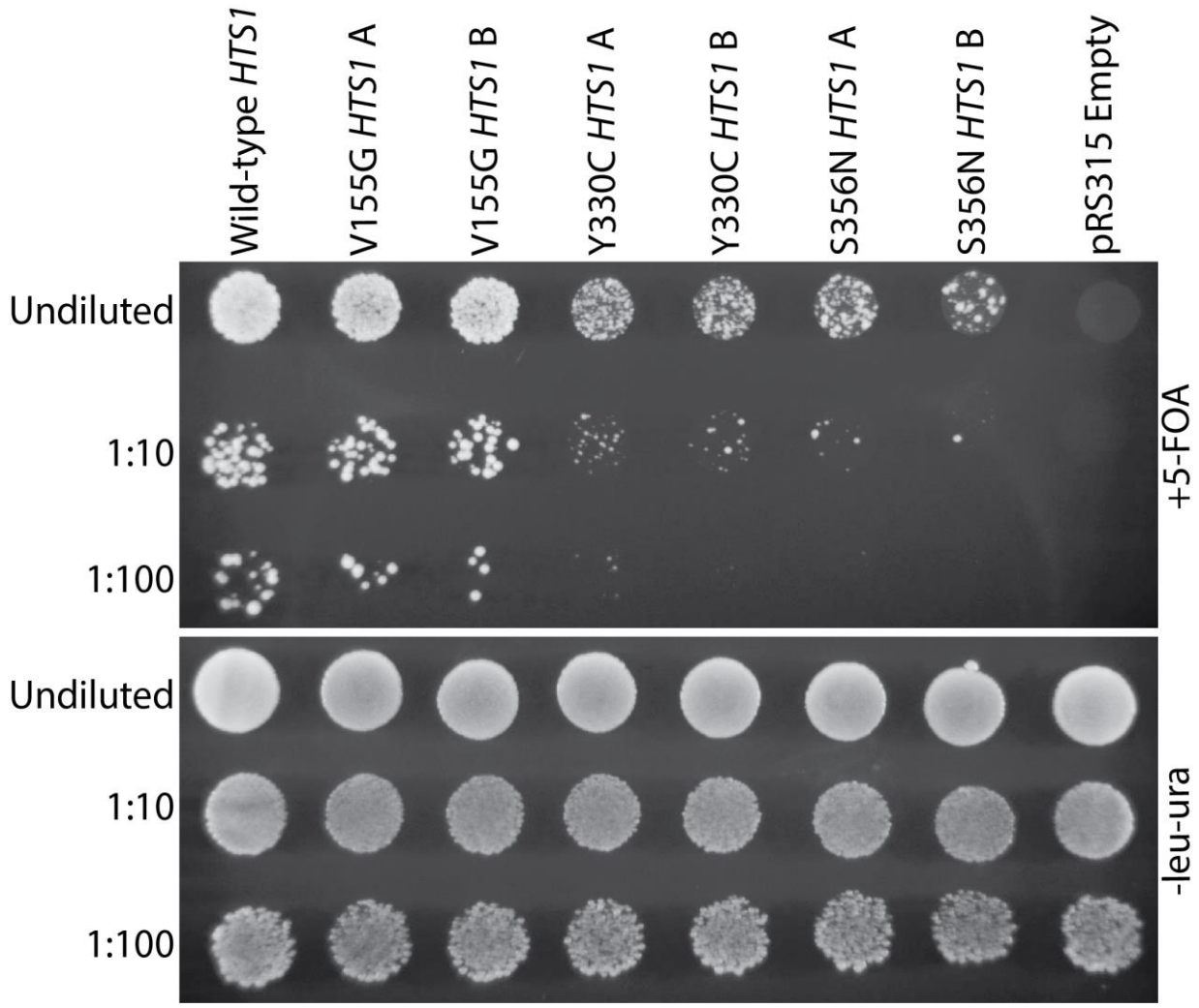


Figure 3.9. Yeast complementation assay of *HTS1* variants. Haploid Δ*HTS1* yeast strains were transformed with a vector containing no insert (pRS315 Empty) or an insert to express wild-type, V155G, Y330C, or S356N *HTS1*. Two colonies (indicated by 'A' and 'B') from transformations with V155G, Y330C, or S356N *HTS1* are shown. Resulting colonies (undiluted, diluted 1:10, or diluted 1:100) were grown on agar plates containing 0.1% 5-FOA or media lacking uracil and leucine. Note the severe depletion of growth associated with Y330C and S356N *HTS1* at 1:10 and 1:100 dilutions.

data presented here implicate nine previously unreported variants in CMT disease. The majority of tested variants demonstrate characteristics typical of disease-associated ARS alleles, specifically conservation between diverse species and loss-of-function in yeast complementation assays. Specifically, genetic data suggest that E196Q *YARS* results in DI-CMT, while G102R *AARS* causes axonal myeloneuropathy. Importantly, the identification of multiple *HARS* variants that segregate in families with a spectrum of neuropathy phenotypes, including axonal and intermediate forms of CMT disease, provides the necessary genetic evidence to fully implicate *HARS* as a CMT-associated locus. Taken together, we conclude that loss-of-function is a common characteristic of disease-associated ARS mutations and an important component of ARS-mediated peripheral neuropathy.

The number of ARS mutations being identified in individuals with peripheral neuropathy has increased rapidly over the past few years. In this chapter alone, we have described and functionally characterized nine newly identified mutations in *YARS*, *AARS,* and *HARS* mutations. The identification of new variants provides the unique opportunity to understand both the clinical and allelic heterogeneity associated with ARS mutations. The clinical heterogeneity of ARS mutations is only beginning to be appreciated. The first mutations in *GARS* were identified by pooling linkage data from five families who suffered from upper limb predominant peripheral neuropathy (Antonellis et al. 2006); however, phenotypes in these families varied and included both motor and sensory (CMT2D) and exclusively motor (dSMA-V) disease. In contrast, *YARS* and *AARS* mutations have only been implicated in lower limb predominant motor and sensory peripheral neuropathy (Jordanova et al. 2006; Latour et al. 2010). While the only phenotype reported for *YARS* mutations, including E196Q described here, remains DI-CMT, this study expands the phenotypic spectrum observed for *AARS* mutations. R329H and N71Y *AARS* were

previously implicated in CMT2N, an axonal CMT phenotype (Latour et al. 2010; Lin et al. 2011; McLaughlin et al. 2011). Although evidence of myelopathy was reported in two families with *GARS* mutations (Sivakumar et al. 2005), the G102R *AARS* mutation represents the first example of myelopathy associated with an *AARS* mutation. Studies at other CMT-associated loci have demonstrated that identical coding mutations can result in variable phenotypes (Thomas et al. 1997). In the studies described here, we observed variable phenotypes for individuals heterozygous for G102R *AARS* (*i.e.* neuropathy versus myeloneuropathy), as well as in family members carrying P134H *HARS* (*i.e.* axonal versus demyelinating versus intermediate CMT disease). This further illustrates the observation that members of the same family who carry an identical coding mutation can be differently affected in regards to age of onset, inclusion of upper motor neuron defects, and axonal versus demyelinating electrophysiologic findings, to name a few examples (Thomas et al. 1997). These variations may be due to genetic modifiers or environmental exposures that vary between individuals and may be interesting avenues for future studies.

Expanding the allelic heterogeneity of ARS alleles through the identification of new ARS variants, both rare variants associated with disease and those in the general population, provides opportunities to identify distinguishing characteristics of disease-associated mutations. Cataloging the functional consequences will not only provide a pipeline for characterization of variants of unknown significance, but also will provide insight into the mechanism of CMT disease. In this chapter, we have shown that eight of nine disease-segregating mutations described here result in loss-of-function in yeast complementation studies, demonstrating the utility of the yeast assay in predicting pathogenicity. Interestingly, *in silico* prediction programs classified both T132I (the disease-associated allele) and T132S *HARS* (an allele with no known

role in disease) as pathogenic variants. When modeled in yeast, however, the disease-associated allele T132I *HTS1* does not support yeast viability, while the T132S variant supports growth to wild-type levels. This illustrates that the yeast assay may be sensitive enough to distinguish between pathogenic and non-pathogenic amino acid changes at the same residue in some cases. In addition we were able to ascertain subtle differences in activity between variants that affect E196 *YARS.* Our results indicate that alteration of this glutamate residue to either lysine or glutamine results in generation of a hypomorphic allele; however, E196K appears to retain slightly more activity than E196Q when comparing yeast growth. While currently no correlation has been made between disease severity and aminoacylation activity of ARS variants, further study into the kinetic differences of these two variants may provide important insights into the structure-function relationship of *YARS*.

Yeast complementation studies allow for *in vivo* assessment of ARS function. All mutations that demonstrate loss-of-function in yeast also demonstrate dramatic reduction in aminoacylation assays *in vitro* (Wallen and Antonellis 2013). Although the majority of disease-associated variants demonstrate loss of function in yeast, some disease-associated variants and all non-pathogenic variants tested to date can complement in this assay (McLaughlin et al. 2011). Thus growth cannot be used to rule out the pathogenicity of a variant. As performed presently, yeast complementation assays utilize the yeast ortholog, not the human protein, to model mutations. In contrast, the aminoacylation assay tests the human enzyme. Thus, it is important to further characterize mutations that grow in yeast using aminoacylation studies to determine if they result in a defect in tRNA charging that is not appreciated when modeled in yeast. Of the mutations studied in this chapter, only two variants, V155G and T132S *HARS*, supported yeast cell growth. Aminoacylation assays are the logical next step to completing the functional assessment for these

variants. In fact, evaluation of V155G *HARS* was recently performed and demonstrated impaired tRNA charging compared to wild-type *HARS* (Chris Franklyn, University of Vermont, personal communication). This datum classifies V155G *HARS* as a loss-of-function mutation and indicates V155G is likely pathogenic.

Previously, R137Q *HARS* was shown to result in abnormalities of the nervous system when overexpressed in *C. elegans* GABA neurons, indicating dominant toxicity for this allele (Vester et al. 2012). Asim Beg at the University of Michigan assessed the D367Y HARS variant described in this chapter for dominant toxicity in worm. Similar to R137Q, D367Y *HARS* overexpression resulted in a dramatic effect on GABA neuron morphology. Worms expressing the D367Y *HARS* allele showed dorsal and ventral nerve gaps, axonal blebbing, and severely aberrant axonal processes. Thus D367Y *HARS* was deemed to be dominantly toxic in this model, consistent with the inheritance pattern observed for the D367Y *HARS*-associated peripheral neuropathy. This model system should be applied to test the dominant pathogenicity of mutations presented in this chapter that have weak or no genetic evidence for their role in dominant disease. Specifically, assessing for dominant toxicity is necessary before implicating V155G and S356N *HARS* mutations in dominantly inherited CMT disease and prior to classifying the T132S *HARS* variant as non-pathogenic.

The identification of novel disease associated ARS alleles in this chapter not only provides the necessary evidence to implicate the *HARS* locus in CMT disease, but also expanded our understanding of the clinical and allelic heterogeneity resulting from ARS mutations. These data illustrate a role for loss of ARS function in the pathomechanism of *YARS*, *AARS*, and *HARS*-mediated CMT disease. This indicates that a common pathogenic mechanism is likely in all ARS-mediated CMT disease. Future work, both included in this thesis (see Chapters 5 and 6)

and studies to be subsequently performed, should now focus on elucidating the pathomechanism of ARS-mediated CMT so that we may develop appropriate therapies to treat patients in the future.

CHAPTER 4

Impaired function is a common characteristic of previously identified *GARS* mutations

INTRODUCTION

Functional studies performed on a subset of CMT-associated ARS mutations have revealed loss-of-function characteristics (Antonellis et al. 2006; Chihara et al. 2007; Nangle et al. 2007; Xie et al. 2007; Stum et al. 2011). As discussed in chapter 3, all of the *YARS*, *AARS*, and *HARS* mutations tested to date cause a severe decrease in enzyme activity and/or demonstrate an inability to rescue deletion of the endogenous ARS gene in yeast complementation assays (Jordanova et al. 2006; McLaughlin et al. 2011; Vester et al. 2012; Gonzaga-Jauregui et al. 2015; Motley et al. 2015; Safka Brozkova et al. 2015). However, the majority of CMT-associated ARS mutations have been identified in *GARS*, and there are conflicting data regarding the presence or absence of loss-of-function characteristics of these alleles (Wallen and Antonellis 2013).

The functional consequences of human *GARS* mutations have been assessed in aminoacylation assays, yeast and fly complementation assays, and localization studies in cultured neurons. While many of the mutations demonstrate a loss-of-function effect in at least one of these assays (Antonellis et al. 2006; Chihara et al. 2007; Nangle et al. 2007; Xie et al. 2007; Stum et al. 2011), nine of the thirteen identified *GARS* mutations have not been rigorously tested (see Table 1.4) (Wallen and Antonellis 2013). Indeed, this includes five mutations (A57V, D146N, S211F,

P244L, and I280F) (Lee et al. 2004; James et al. 2006; Rohkamm et al. 2007; Abe and Hayasaka 2009) that have not been evaluated in any assay. Characterizing the full panel of *GARS* mutations will be necessary to determine whether a loss-of-function mechanism is a component of CMT2D and dSMA-V. These data will be critical for a better understanding of the molecular pathology of ARS mutations and for developing effective therapies for patients with ARS-related CMT disease.

This chapter focuses on the genetic, enzyme kinetic, yeast complementation, and protein localization studies that elucidate the functional consequences of CMT-associated *GARS* mutations. Consistent with studies on *YARS, AARS,* and *HARS* mutations, we hypothesize that disease-associated *GARS* mutations will demonstrate loss-of-function characteristics in at least one assay, while non-pathogenic variants will show wild-type activity in all assays.

The data presented in this chapter were previously published in *Human Mutation* (Griffin et al. 2014). Permission was requested to reproduce figures and tables. The author performed all the studies in this chapter with the following exceptions: (1) Aminoacylation studies were performed by Reiko Sakaguchi and Ya-Ming Hou at Thomas Jefferson University; (2) The identification of the S581L *GARS* mutation in two unrelated probands was performed by Michael Gonzalez and Stephan Zuchner at the University of Miami and Charles Searby at the University of Iowa; (3) Sequencing reactions were performed at the University of Michigan Sequencing Core.

MATERIALS AND METHODS

Computational assessment of GARS *mutations*

GARS protein sequences were collected from the NCBI Protein Database (http://www.ncbi.nlm.nih.gov/protein/) for the indicated species using the following accession

numbers: human (*Homo sapiens*, accession number AAA57001.1), mouse (*Mus musculus*, accession number AAH21747.1), zebrafish (*Danio rerio*, accession number XP 692410.4), roundworm (*Caenorhabditis elegans*, accession number NP 498093.1), and baker's yeast (*Saccharomyces cerevisiae*, accession number NP 009679.2). Multiple species amino acid sequence alignments were then generated using ClustalW2 software (Larkin et al. 2007).

GARS *and* GRS1 *expression constructs*

DNA constructs for aminoacylation, yeast complementation, and localization studies were generated using Gateway cloning technology (Invitrogen, Carlsberg, CA). Briefly, the human *GARS* open reading frame was amplified from a human cDNA sample, and the *S. cerevisiae GRS1* locus was amplified from *S. cerevisiae* genomic DNA. Primers were designed with flanking Gateway sequences attB1 (forward) and attB2 (reverse; primer sequences in Appendix A). Entry clones were generated by recombining PCR-purified amplicons into the pDONR221 vector using BP clonase per manufacturer's specifications. After transformation into *E. coli*, DNA from individual entry clones was isolated and subjected for DNA sequencing analysis to confirm the presence of the appropriate wild-type gene. For each mutation studied here, mutation-containing oligonucleotides were generated, and the QuickChange II XL Site-Directed Mutagenesis Kit (Stratagene, Santa Clara, CA) was used as per the manufacturer's instructions. After transformation into *E. coli*, DNA from individual clones was purified and sequenced to confirm the presence of each mutation and the absence of any cycle-induced errors. Subsequently, validated entry clones were purified and recombined into the appropriate Gateway-compatible vector using LR clonase and the manufacturer's specifications. The vectors included pET-21a(+) (aminoacylation assays), pEGFP-N2 (for expressing GARS with a C-terminal EGFP tag), and pRS315 (yeast complementation assays). DNA from the resulting

expression constructs was purified and digested with the restriction enzyme *Bsr*GI (New England Biosystems, Ipswich, MA) to confirm the presence of the appropriate insert.

Aminoacylation assays

Wild-type and mutant human GARS proteins were expressed in E. coli Rosetta 2 (DE3) pLys cells with a C-terminal in-frame 6xHis tag and purified with nickel affinity resin according to the manufacturer's protocol (Novagen, Rockland, MA). The T7 transcript of human $tRNA^{Gly/UCU}$ (UCU, anticodon) was prepared and purified as previously described (Hou et al. 1993), heat denatured at 85°C for 3 min, and annealed at 37°C for 20 min before use. Steady-state aminoacylation assays were monitored at 37°C in 50 mM HEPES (pH 7.5), 20 mM KCl, 10 mM $MgCl_2$, 4 mM DTT, 2 mM ATP, and 50 mM ^{3}H-glycine (Perkin Elmer, Waltham, MA) at a specific activity of 16,500 dpm/pmole. The reaction was initiated by mixing GARS enzyme (20–600 nM) with varying concentrations of $tRNA^{Gly/UCU}$ (0.3–20 μM). Aliquots of a reaction mixture were spotted on filter paper, quenched with 5% trichloroacetic acid, washed, dried, and measured for radioactivity by a liquid scintillation counter (LS6000SC; Beckman Coulter Inc., Fullerton, CA). The amount of radioactivity retained on filter pads was corrected for quenching effects to determine the amount of Gly-$tRNA^{Gly}$ present on the filter paper. Steady-state kinetics (K_m and k_{cat}) were determined by fitting the initial rate of aminoacylation as a function of tRNA concentration to the Michaelis–Menten equation (Schreier and Schimmel 1972).

Yeast complementation assays

The RJT3/II-1 haploid yeast strain [*MATα grs1::HIS3, his3Δ200, leu2Δ1, lys2Δ202, trpΔ63, ura3–52*, pTsscII-maint (*cen, GRS1, URA3*)] carrying a deleted endogenous *GRS1* allele and wild-type *GRS1* on a *URA3*-bearing pRS316 maintenance vector was previously generated

(Turner et al. 2000). The Δ*GRSI* haploid yeast strain was transformed with a 200 ng of *LEU2*-bearing pRS315 vector containing wild-type or mutant *GRS1* (described above) or pRS315 with no insert as previously reported (Antonellis et al. 2006). Each transformation was performed at least four times with at least three independent plasmid DNA preparations. Four colonies were selected from each transformation for additional analysis. Each colony was grown to saturation in media lacking leucine and uracil (-leu -ura) for 48 hours. Next, 10 µl of undiluted and diluted (1:10 and 1:100 in H_2O) samples from each culture were spotted on plates containing 0.1% 5-FOA (Boeke et al. 1987) complete medium, yeast extract– peptone–glycerol (YPG) plus 0.1% 5-FOA medium, or SD -leu -ura growth medium (Teknova, Hollister, CA) and incubated at 30°C or 37°C for 72 hr. Yeast cell growth was determined by visual inspection.

Cell culture and protein localization studies

The mouse motor neuron, neuroblastoma fusion cell line (MN-1) (Salazar-Grueso et al. 1991) was cultured and transfected with constructs to express wild-type or mutant GARS in-frame with a C-terminal enhanced green fluorescent protein (EGFP) tag as previously described (Antonellis et al. 2006). Cells were counted using the Countess Automated Cell Counter (Invitrogen, Calsbad, CA) and ~1.25 X 10^5 of MN-1 cells were plated in each well of a 4-well polystyrene tissue culture treated glass slide (BD Biosystems, Sparks, MD). Slides were incubated overnight at 37°C in 5% CO_2 in Dulbecco's modified eagle medium (DMEM) supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 100 U/ml penicillin, and 50 µg/ml of streptomycin (Invitrogen, Carlsbad, CA). Each well was transfected using the lipofectamine 2000. Specifically, for each transfection reaction, 125 µl of OptiMEM I minimal growth medium and 1.25 µl of Lipofectamine 2000 were mixed and incubated together for 10 min at room temperature. 1.5 µg of pEGFP-N2 plasmid was combined with 125 µl of OptiMEM I.

Subsequently, the Lipofectamine-OptiMEM I mixture was combined with the plasmid-OptiMEM I solution to yield 250 μl of the final transfection solution and incubated at room temperature for 20 min. Each well of the 4-well slide was washed with 500 μl of 1X PBS. The entire transfection solution was then added directly to the flask and incubated at 37°C in 5% CO_2 for 4 hours. After 4 hours, the transfection solution was removed and replaced with 500 ml of complete DMEM media (described above). Cells were allowed to recover for 2 days at 37°C in 5% CO_2. After 48 hours, growth medium was removed, and cells were washed in 1× PBS and then incubated in 1× PBS/0.4% paraformaldehyde for 10 min at room temperature. Cells were washed in 1× PBS, co-stained with 300 nM DAPI for 5 min, washed again in 1× PBS, and finally coated with Pro-Long antifade reagent (Invitrogen). Images were obtained with an IX71 Inverted Microscope using cellSens Standard image software (Olympus, Center Valley, PA).

Population screening and segregation analysis

The presence of each *GARS* variant in the general population was assessed using the Exome Variant Server (snp.gs.washington.edu/EVS), the 1000 Genomes Project database (www.1000genomes.org), and dbSNP (www.ncbi.nlm.nih.gov/snp). Additionally, a search of the GEM.app exome-sequencing database was performed for *GARS* variants in patients with CMT disease (genomics.med.miami.edu). For S581L *GARS* segregation studies, oligonucleotide primers were designed to amplify *GARS* exon 16, which harbors the second and third nucleotides of the serine 581 codon (primer sequences in Appendix A). Patient DNA samples were subjected to PCR amplification under standard conditions along with a "no DNA" PCR-negative control reaction. Subsequently, PCR products were purified on Mini Spin Columns (Epoch Life Science, Inc., Missouri City, TX), eluted, and then subjected to Sanger DNA sequencing analysis using the "forward" and "reverse" amplification primers in separate reactions. The appropriate review

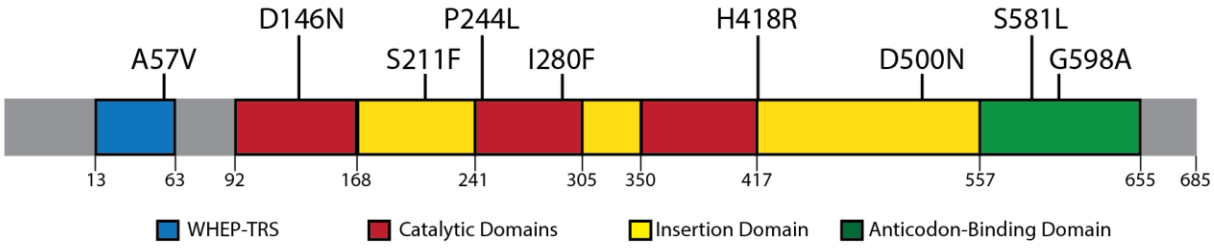
boards of each participating institution approved these studies.

RESULTS

Localization and conservation of human GARS *mutations*

GARS mutations found in patients with CMT disease are distributed throughout the primary structure of the protein (Antonellis et al. 2003). Notably, the mutations affect conserved residues, which are often conserved between human and yeast, suggesting a critical role in protein function (Antonellis et al. 2003). To determine if newly reported *GARS* mutations demonstrate a similar level of conservation, we mapped the residues onto the human protein and generated an alignment with evolutionarily diverse organisms (Figure 4.1A and B). Mutations were evenly distributed throughout the major functional domains of the protein (Xie et al. 2007) (Figure 4.1A): (1) A57V affects a residue in the disordered WHEP-TRS domain (residues 13–63; pfam00458 in the NCBI Conserved Domain Database), which is putatively involved in tRNA binding and interacting with other aminoacyl-tRNA synthetases in enzyme complexes (Ray et al. 2011), but is not critical for GARS function (Xie et al. 2007); (2) D146N, P244L, and I280F affect amino acids in the catalytic core (residues 92–168, 241–305, and 350-418; pfam 00587) responsible for ligation of the amino acid to tRNA; (3) S211F, H418R, and D500N affect amino acids within insertion domains of uncertain function; and (4) S581L and G598A reside in the anticodon-binding domain (residues 557–655; pfam03129) that binds to glycine-specific tRNAs. The evolutionary conservation of mutated GARS residues varies greatly (Figure 4.1B). D146, P244, I280, and H418 are identical from human and yeast, while S211 and G598 are conserved, but to a lesser degree (*e.g.*, from human through worm). In contrast, A57 is conserved among mammals and worm, but not zebrafish, and D500 and S581 are conserved only among

A



B

	A57V	D146N	S211F
Human	LEAKEL**A**LQPKDD	HVDKFA**D**FMVKDV	VNYNVK**S**PITGND
Mouse	LEAKEL**A**LQPKDD	HVDKFA**D**FMVKDV	VNYNVK**S**PTTGND
Zebrafish	LEAKELSLQPKDD	HVDKFA**D**YMVKDV	VQYNVK**S**PSTGND
Worm	LEDTEI**A**LAPKEA	HVDRFA**D**WMVKDM	TRFNFK**S**PITGND
Yeast	----------	HVDKFS**D**WMCRDL	EKYDIGNPVTGET

	P244L	I280F	H418R
Human	MPGYLR**P**ETAQGI	NSFRNE**I**SPRSGL	CYDLSC**H**ARATK
Mouse	MPGYLR**P**ETAQGI	NSFRNE**I**SPRSGL	CYDLSC**H**ARATK
Zebrafish	MQGYLR**P**ETAQGI	NSFRNE**I**SPRSGL	CYDLLC**H**ARATK
Worm	FKAFLR**P**ETAQGI	LGFRNE**I**SPRQGL	CYDLQQ**H**YKATN
Yeast	LKGYLR**P**ETAQGQ	KSFRNE**I**SPRAGL	AYDLTV**H**SKKTK

	D500N	S581L	G598A
Human	TFQLTK**D**MINVKR	PFVKEL**S**EALTRH	KVDDSS**G**SIGRRYA
Mouse	TFQLTK**D**MVSVKR	PFVKEL**S**EALTRN	KVDDSS**G**SIGRRYA
Zebrafish	TFKLTK**D**MVNIKR	PFVREL**S**EALTRN	KVDDSS**G**SIGRRYA
Worm	QNYALTPEHLNIK	PVMDAVKEELSRF	KVDDSS**G**TIGRRYA
Yeast	DGEVELDDKLVKI	PVHHEVAKILRKS	KIDDSGVSIGKRYA

Figure 4.1. Localization and conservation of *GARS* variants. (A) *GARS* mutations (top) are depicted on a cartoon of the known functional and structural domains of GARS including the WHEP-TRS domain (blue), catalytic domain (red), insertion domains (yellow) and anticodon-binding domain (green). The amino-acid positions for each domain are indicated below the cartoon. (B) GARS protein sequence alignments from multiple, evolutionarily diverse species are depicted. The amino acid change is listed at the top of each protein fragment. The affected amino acid is highlighted in red, bold text.

vertebrates. Thus D146, S211, P244, I280, H418, and G598 are predicted to be critical for enzyme function, while A57, D500, and S581 are less likely to be essential for GARS activity.

The majority of GARS *mutations reduce aminoacylation activity.*

The primary function of ARS enzymes is to charge tRNA molecules with cognate amino acids (Delarue 1995) in a two-step aminoacylation reaction: (***i***) binding and activation of the amino acid with ATP and (***ii***) conjugation of the activated amino acid to the 3' end of the appropriate tRNA (Antonellis and Green 2008). The aminoacylation assay evaluates the combined reactions of both steps and has been used to establish if CMT-associated ARS mutations disrupt this canonical function (Nangle et al. 2007; McLaughlin et al. 2010; Froelich and First 2011; McLaughlin et al. 2011). Here, we assessed previously untested *GARS* variants (A57V, D146N, S211F, P244L, I280F, H418R, and G598A) for the ability to charge human cytoplasmic $tRNA^{Gly}$ with tritium-labeled glycine. For comparison purposes, we included two *GARS* mutations previously shown to impair enzyme kinetics: G240R and G526R (Nangle et al. 2007). With regard to the untested variants, D146N, S211F, P244L, I280F, H418R, and G598A GARS demonstrate less than 10% aminoacylation activity compared to wild-type GARS (Table 4.1). In contrast, A57V displayed the ability to charge $tRNA^{Gly}$ molecules with a much higher efficiency than other mutant proteins, albeit with a 50% reduction relative to wild-type GARS.

The majority of testable GARS *mutations dramatically reduce yeast cell viability.*

Yeast complementation assays have been performed to determine the functional consequences of CMT-associated ARS mutations in the context of a living cell (Wallen and Antonellis 2013). Mutations are modeled in the *S. cerevisiae GARS* ortholog *GRS1*, and cellular growth is used as a proxy for enzyme function. This assay employs a haploid yeast strain with the endogenous *GRS1* locus deleted and viability maintained via expression of wild-type *GRS1* on a *URA3*-bearing

Table 4.1. Aminoacylation kinetics of GARS protein variants

Variant	Km (μM)	kcat (s-1)	kcat/Km (μM-1s-1)	Relative to WT
Wild-type	1.03 ± 0.03	0.14 ± 0.02	0.14 ± 0.02	1
A57V	0.92 ± 0.22	0.07 ± 0.01	0.07 ± 0.02	1 / 2
D146N	0.6 ± 0.1	0.006 ± 0.002	0.01 ± 0.01	1 / 14
S211F	-	-	(3.2 ± 1.5) x 10-5	1 / 4,400
G240R	3.2 ± 0.6	(4.2 ± 0.3) x 10-3	(1.3 ± 0.4) x 10-3	1 / 110
P244L	-	-	undetectable	undetectable
I280F	-	-	(7.9 ± 2.0) x 10-5	1 / 1,700
H418R	0.004 ± 0.001	(3.9 ± 0.1) x 10-5	0.01 ± 0.01	1 / 16
G526R	7.1 ± 1.3	(2.3 ± 0.4) x 10-5	(3.2 ± 2.0) x 10-6	1 / 44,000
G598A	1.4 ± 0.3	(1.1 ± 0.1) x 10-3	(7.9 ± 3.0) x 10-4	1 / 180

Amino-acid coordinates correspond to GenBank accession number AAA57001.1

± indicates standard deviations

- indicates that kcat and km are not directly measured; the value for kcat/Km (where applicable) was measured under an approximate condition where the substrate concentration is below an estimated Km.

“Undetectable" indicates that values are below the limits of detection

maintenance vector (Turner et al. 2000; Antonellis et al. 2006; Stum et al. 2011). Based upon ClustalW2 alignments (Figure 4.1B), we identified the orthologous GRS1 amino acid of each previously untested *GARS* mutation and performed site-directed mutagenesis to model the human mutations in yeast. We were able to confidently model D146N, P244L, and I280F *GARS* at the orthologous residues in *GRS1* based on homology (Table 4.2); A57V, S211F, D500N, and S581L are not conserved between human and yeast and thus could not be modeled in this assay. For clarity *GRS1* mutations are referred by the human mutation nomenclature.

Yeast cells were transformed with the experimental alleles on a *LEU2*-harboring vector and plated on solid growth medium containing 0.1% 5-Fluoroorotic Acid (5-FOA; Figure 4.2). 5-FOA is toxic to yeast cells expressing *URA3* and thus selects for cells that have spontaneously lost the maintenance vector (Boeke et al. 1987). Only yeast cells containing a functional *GRS1* allele on the *LEU2*-bearing experimental plasmid are able to complement the deleted endogenous *GRS1* allele and survive on 5-FOA plates. Our wild-type *GRS1* expression construct allowed robust yeast cell growth and our experimental vector without a *GRS1* gene ('Empty' in Figure 4.2) was unable to complement the null allele in yeast cells (Figure 4.2A). These results are consistent with the construction of a functional *GRS1* experimental system and with *GRS1* being an essential gene, respectively (Turner et al. 2000; Antonellis et al. 2006). With regard to the modeled mutations, I280F *GRS1* was able to maintain yeast viability (Figure 4.2A) in a comparable manner to wild-type *GRS1*. In contrast, D146N and P244L *GRS1* resulted in a severe reduction of yeast cell growth compared to wild-type *GRS1* (Figure 4.2A). We also evaluated yeast cells incubated at 37°C to assess for temperature-sensitive effects of the studied mutations. These analyses did not reveal any further depletion of growth associated with the modeled mutations compared to wild-type *GRS1* (Figure 4.2B). Finally, since GARS is a bi-functional

Table 4.2. Human *GARS* Variants Modeled in the Yeast Ortholog *GRS1*

Human GARS[1]	Yeast GRS1[2]
D146N	D96N
P244L	P227L
I280F	I255F

[1]Amino-acid coordinates correspond to GenBank accession number AAA57001.1.
[2]Amino-acid coordinates correspond to GenBank accession number NP_009679.2.

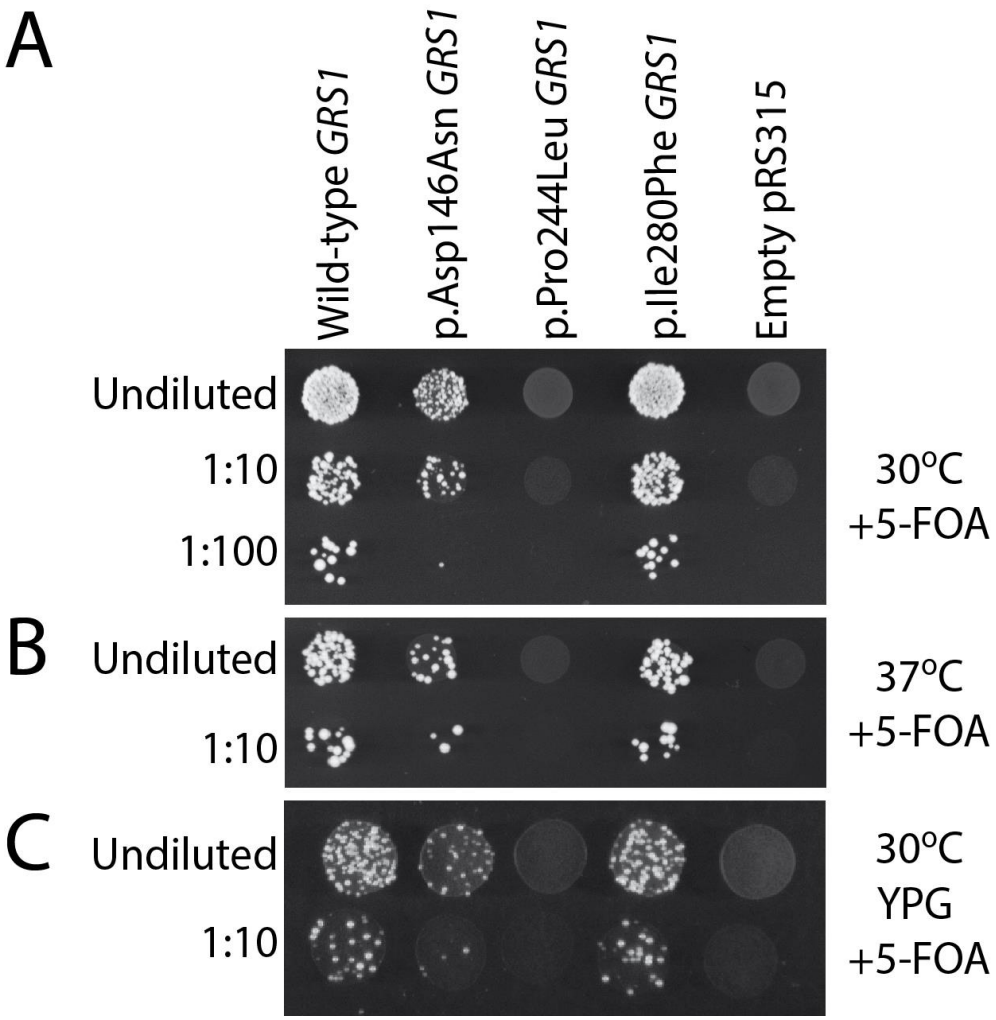


Figure 4.2. Characterization of yeast expressing wild-type and mutant *GRS1*. Each yeast strain was transformed with a *LEU2*-bearing pRS315 vector containing wild-type *GRS1*, the indicated mutant form of *GRS1*, or no insert ('Empty pRS315'). Cultures for each strain (labeled along the top for each panel in A, B, and C) were grown for two days in liquid media and spotted on selective solid growth medium directly or after dilution as indicated. Strains were plated on medium containing 5-FOA to determine if the *GRS1* alleles complement loss of endogenous *GRS1* gene at 30°C (A) and 37°C (B). (C) Strains were plated on solid medium containing 5-FOA and glycerol (YPG) and grown at 30°C to assess for an effect on mitochondrial GARS function.

ARS enzyme responsible for charging tRNA in both the cytoplasm and mitochondria (Antonellis and Green 2008), yeast were grown on medium containing glycerol and 5-FOA at 30°C to require yeast to rely solely on mitochondrial respiration, which allows the detection of impaired mitochondrial function. Similar to the studies at 37°C, the relative pattern of viability remained consistent when comparing mutant and wild-type *GRS1* and therefore did not reveal any further mutation-dependent depletion in yeast cell growth (Figure 4.2C). Thus, two of the three human *GARS* mutations modeled in *GRS1* dramatically decrease cell viability *in vivo*, consistent with a loss-of-function effect.

The majority of GARS *mutations reduce punctate localization in cultured neurons.*

The wild-type GARS protein localizes to 'puncta' within axons in the human peripheral nerve where it likely functions in local protein translation (Antonellis et al. 2006). Certain *GARS* mutations impair the punctate localization in cultured neurons, suggesting that reduced localization to axons is a component of CMT2D and dSMA-V disease pathogenesis (Antonellis et al. 2006). However, not all *GARS* mutations have been tested for an effect on puncta formation in neuronal cells. We therefore expressed wild-type and the eight untested mutant GARS-EGFP fusion proteins in a mouse motor neuron, neuroblastoma fusion cell line (MN-1 cells) (Salazar-Grueso et al. 1991; Antonellis et al. 2006). Fluorescence microscopy revealed that wild-type, A57V, D146N, D500N, and S581L GARS-EGFP all demonstrated a punctate localization pattern in MN-1 cells (Figure 4.3A, B, C, G, and H). In contrast, S211F, P244L, I280F, and G598A GARS-EGFP proteins did not associate with puncta in MN-1 cells (Figure 4.3D, E, F, and I). Importantly, for the mutant GARS-EGFP proteins that did form puncta, these structures were also observed in MN-1 neurite projections, consistent with previous observations of wild-type GARS localization (data not shown) (Salazar-Grueso et al. 1991; Fallini et al. 2012).

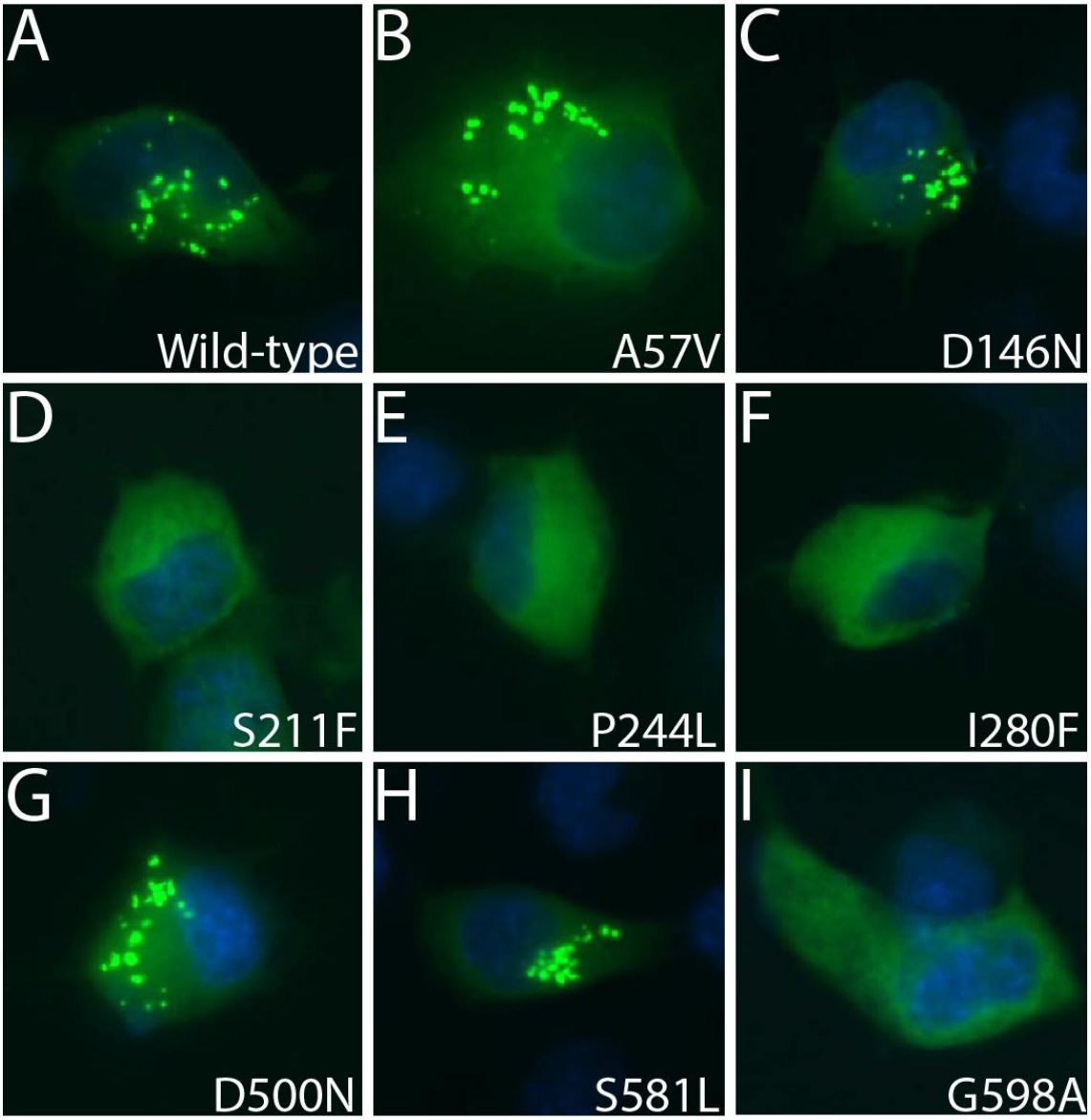


Figure 4.3. Expression of wild-type and mutant GARS in mouse motor neuron-derived cells. MN-1 cells expressing EGFP-tagged wild-type (A), A57V (B), D146N (C), S211F (D), P244L (E), I280F (F), D500N (G), S581L (H), or G598A GARS (I) were evaluated for the presence or absence of granules by fluorescence microscopy.

S581L GARS *does not segregate with disease in newly discovered pedigrees.*

The A57V variant was identified in a single individual with weakness and muscle wasting in the upper extremities (Rohkamm et al. 2007) and S581L was identified in a single proband with an extensive family history of lower limb predominant axonal CMT disease (James et al. 2006). No further segregation analyses were performed on either mutation. Furthermore, ethnically-matched control populations were not screened for the absence of A57V or S581L *GARS*. Therefore, we searched genome variant databases for the occurrence of these two mutations. A57V and S581L were both identified by the NHLBI GO Exome Sequencing Project database at a low frequency of 1/11,842 (rs370531212:C>T) and 16/119,342 (rs201358272:C>T), respectively. Of the 11 additional *GARS* mutations associated with disease, D500N was identified at a frequency of 3/112,608 in control or general populations, while the remaining 10 have not been reported in control or general populations.

Our on-going mutation analyses revealed S581L *GARS* in two unrelated probands with dominant axonal CMT disease. To determine if this mutation is associated with disease in the two respective families, we performed segregation studies. DNA sequence analysis of additional affected individuals revealed that S581L *GARS* does not segregate with disease in either family (Figure 4.4A and B). Specifically, clinically affected individuals in both families did not carry S581L *GARS*. Combined with the identification of this variant in the general population, these data indicate that S581L *GARS* should not be considered a disease-causing mutation.

DISCUSSION

Prior to the completion of this work, thirteen GARS mutations had been associated with CMT disease; however, only four of the 13 reported *GARS* mutations had previously been rigorously

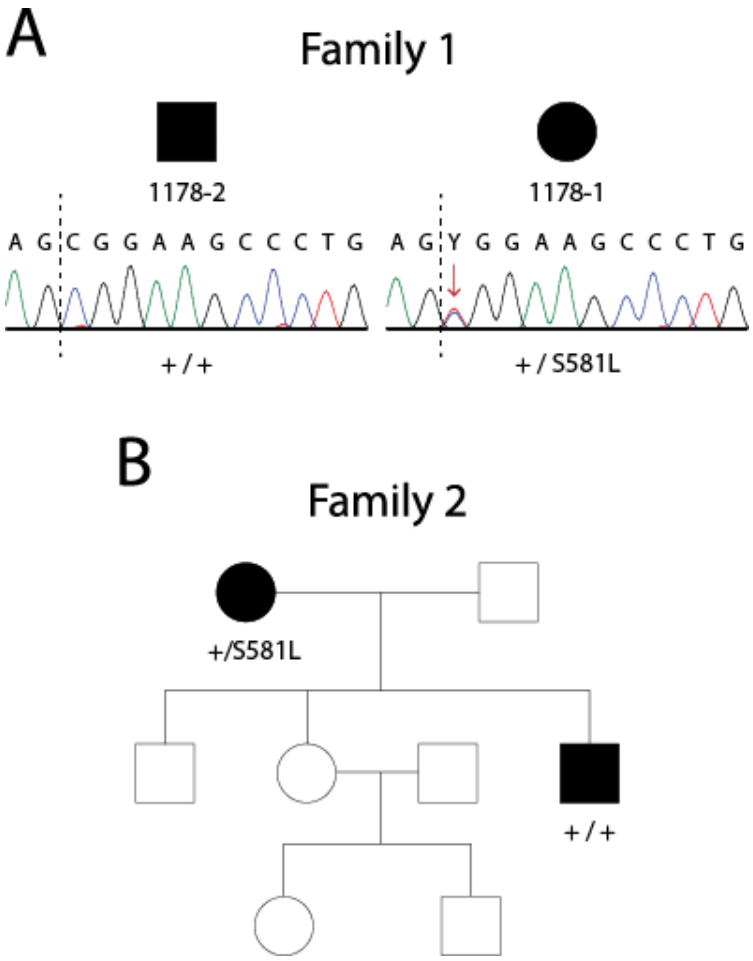


Figure 4.4. S581L *GARS* does not segregate with CMT disease. DNA samples from affected individuals from two presumably unrelated families with autosomal dominant, axonal CMT disease were genotyped for S581L *GARS*. Female patients are indicated with circles and male patients are indicated with squares. Filled symbols represent affected individuals with a diagnosis of dominant axonal CMT disease and empty symbols indicate unaffected individuals. Where applicable, an individual's genotype is indicted with + (for the wild-type allele) or S581L (for the variant allele). (A) Samples from individuals in Family 1 were anonymized resulting in loss of the pedigree structure for this family; however, individuals 1178-1 and 1178-2 are directly related, both affected, and reside in a pedigree originally described as dominant axonal CMT disease. Representative chromatographs from sequencing analysis of genomic DNA from affected individuals are shown. The sequence includes the AG splice site acceptor at the 3' end of intron 15 followed by the first 12 nucleotides of exon 16 (intron/exon junction marked by dashed line). The S581 codon (TCG) is encoded by the last nucleotide of exon 15 (T; not shown) and the first two nucleotides (CG) of exon 16. Individual 1178-1 is heterozygous for a C>T mutation (red arrow) in the first base of exon 16 resulting in the S581L allele. 1178-2 is homozygous for the wild-type genotype at this nucleotide (C). (B) The pedigree for Family 2 illustrates the affected status of each family member and the S581L genotypes of two affected individuals with dominant axonal CMT disease.

tested (Wallen and Antonellis 2013). The four mutations that had been previously characterized all demonstrated loss-of-function in at least one assay. To determine the functional consequences of *GARS* mutations, we evaluated the remaining nine *GARS* mutations in aminoacylation, yeast complementation, and localization studies and report a high correlation between impaired enzyme function and CMT disease. We determined that the D146N, H418R, and G598A *GARS* alleles encode a protein product with a severe defect in tRNA charging. Additionally, S211F, P244L, I280F, and G598A GARS fail to form puncta that were previously associated with wild-type GARS in human neuronal tissues and cultured cells (Antonellis et al. 2006). Importantly, our genetic analyses rule out S581L *GARS* as a pathogenic *GARS* mutation. Combined with previous reports, the data presented here reveal that 10 of the 12 disease-associated *GARS* mutations identified to date demonstrate loss-of-function characteristics (Table 4.3) (Antonellis et al. 2006; Nangle et al. 2007; Xie et al. 2007; Stum et al. 2011). The two mutations that do not meet this criterion, A57V and D500N *GARS*, are discussed in detail below. This work furthers our understanding of the allelic spectrum of *GARS*-associated CMT disease and strongly suggests that impaired GARS function is an important component of the axonal peripheral neuropathy.

The *in vitro* aminoacylation assay is performed with tRNA and a single purified enzyme to determine the activity of wild-type or mutant GARS proteins (Hou et al. 1993). This assay employs catalytic amounts of the enzyme (nM) with saturating amounts of substrate tRNA (in mM). Combining previous studies, data presented here, and recent aminoacylation data collected that demonstrates >95% reduction in aminoacylation for E71G *GARS* (Ya-ming Hou, personal communication), it is now known that 10 of the 12 disease-associated *GARS* mutation lead to at least a 90% decrease in tRNA charging capacity (Table 4.3) However, while aminoacylation

Table 4.3. Summary of functional studies on *GARS* variants

Variant	Enzyme Activity	Yeast Viability	Granule Formation	References
Wild-type	Normal	Viable	Yes	Antonellis et al. 2006, this study
A57V	Normal[1]	NA	Yes	This study
E71G	Reduced	Viable	Yes	Antonellis et al. 2006, Nangle et al. 2006, Hou, unpublished
L129P	Reduced	Reduced	No	Antonellis et al. 2006, Nangle et al. 2006
D146N	Reduced	Reduced	Yes	This study
S211F	Reduced	NA	No	This study
G240R	Reduced	Viable	No	Antonellis et al. 2006, Nangle et al. 2006
P244L	Reduced	Lethal	No	This study
I280F	Reduced	Viable	No	This study
H418R	Reduced	Lethal	No	Antonellis et al. 2006, Nangle et al. 2006, and this study
D500N	Normal	NA	Yes	Nangle et al. 2006, and this study
G526R	Reduced	Lethal	Yes	Antonellis et al. 2006, Nangle et al. 2006, and Xie et al. 2007
S581L	Normal	NA	Yes	Nangle et al. 2006, Cader et al. 2007, and this study
G598A	Reduced	Viable	No	Stum et al. 2011, and this study

Amino-acid coordinates correspond to GenBank accession number AAA57001.1

Red text indicates a result consistent with a loss-of-function effect.

[1]This variant was associated with ~50% activity compared to wild-type GARS and we considered a >90% reduction in activity as dramatically 'Reduced'.

NA – not applicable; mutation could not be modeled in yeast ortholog

assays provide key information on enzyme function, the reaction is performed outside the context of a living cell. Therefore, it may not faithfully represent the ability of mutant enzymes to meet *in vivo* requirements for tRNA charging. In contrast, yeast complementation assays provide an *in vivo* system to characterize the functional consequences of *GARS* mutations (Turner et al. 2000; Antonellis et al. 2006). Importantly, all *GARS* mutations that fail to complement yeast cell growth also have a severe reduction in aminoacylation activity. However, certain mutations, such as I280F, show reduced aminoacylation activity *in vitro* but complement yeast cell growth in a manner similar to the wild-type enzyme (Table 4.3). One possible explanation for the discrepancy between aminoacylation and yeast complementation studies is that bacterially produced proteins may lack post-translational modifications important for enzyme function in human cells or for proper activity in specific cellular contexts. Additionally, other cellular components may stabilize the mutant protein product in the yeast system, allowing for improved function and viability. The yeast assay could not be used to evaluate all *GARS* mutations due to a lack of conservation with human residues. Alternative assays (see below) should be explored to determine the *in vivo* consequences of such mutations. While yeast offers insight into the functional consequences of *GARS* mutations *in vivo*, it does not reveal how *GARS* mutations affect axons. The function of GARS and other ARS enzymes in axons remains unclear; however, there is strong evidence that they provide tRNA charging for local protein translation (Ingoglia et al. 1983). One possibility is that *GARS* mutations reduce enzyme activity below some required threshold of tRNA charging necessary for axon function (see chapter 5). Indeed, mutations that reduce catalytic activity but sustain yeast cell viability, such as I280F, may not have enough activity to sustain axon-specific translation requirements.

Proper localization within a cell can be critical for protein function, with impaired localization

resulting in a cell-compartment specific loss of function. Wild-type endogenous GARS forms discrete puncta in cultured neurons and in peripheral neurons and axons from healthy human subjects (Antonellis et al. 2006). Data collected to date show that seven of the 12 disease-associated *GARS* mutations give rise to proteins that are unable to associate with granules in cultured neurons (Table 4.3). In another study, Nangle *et al* also showed that mutant GARS proteins mislocalize (Nangle et al. 2007). However, their experiments did not reveal discrete GARS puncta, which is likely due to differential placement of the fluorescent protein tag (*i.e.* N-terminal versus C-terminal tag). The impaired association with puncta observed for certain GARS mutant proteins may be due to altered protein structure and/or reduced binding to specific protein or RNA partners required for puncta formation. It will be critical to identify proteins and/or RNAs that interact with GARS in these structures to further understand puncta function and how disruption of these structures may lead to CMT disease (see Chapter 6).

Although impaired function is correlated with *GARS*-mediated CMT disease, two disease-associated mutations have not demonstrated impaired function in any of the assays used in this study: A57V and D500N (Table 4.3). D500N *GARS* segregates with disease in a large pedigree with dominantly inherited, upper limb predominate axonal neuropathy (Del Bo et al. 2006), providing strong genetic and phenotypic evidence supporting a role in CMT2D disease pathogenesis. It is important to note that this variant has been detected at a frequency of 3/112,608 in the ExAC database. However, since individuals included in the ExAC database are not subjected to a neurologic assessment and since CMT is often a late onset disease, inclusion in this database does not exclude D500N from having a causative role in CMT disease. In contrast, A57V has weak genetic evidence for a role in CMT disease. First, this variant was identified in a single individual with upper limb predominant CMT disease and no additional family members

were available for segregation studies (Rohkamm et al. 2007). Second, ethnically matched controls were not tested to determine the presence of this allele in appropriate populations. Additionally, A57V was identified by the NHLBI GO Exome Sequencing Project (rs370531212:C>T) at a low frequency (1/11841). Since the role of A57V in disease is still unclear and D500N has not demonstrated loss of function characteristics, we propose that both *GARS* alleles should be evaluated in the fly neuron complementation assay to assess impaired function that cannot be detected by other assays (Chihara et al. 2007) and in worm neurotoxicity assays to assess for dominant toxicity (Vester et al. 2012; Safka Brozkova et al. 2015). These efforts will complete the functional evaluation of these variants, reveal if they have a potential role in dominantly inherited CMT disease, and, if so, determine if loss of function is an important component of the pathogenic mechanism.

With the completion of the human genome project and the development of rapid, affordable sequencing technologies, DNA sequencing is increasingly employed for research and diagnostic purposes. Once a gene has been implicated in disease, patients with a similar phenotype can be quickly screened for variants at that locus. These studies provide a large cache of information that allows for a better understanding of the locus and allelic heterogeneity of human genetic disease. However, rare variants identified in individual patients or small pedigrees are difficult to implicate in disease phenotypes due to the absence of strong genetic evidence (MacArthur et al. 2014). Recent studies have utilized computational, functional, and clinical data to determine that a subset of variants reported in disease-associated genes are likely to be functionally neutral and not responsible for disease. For example, Easton *et al.* computationally evaluated 1,433 variants of unknown significance identified in the *BRCA1* gene and predict that only 20% of those variants are deleterious based on genetic and clinical assessments compared to known disease-

causing mutations (Easton et al. 2007). In a separate study, evaluation of 149 variants in the *CFTR* gene identified in patients with cystic fibrosis revealed that 12 do not meet clinical and functional criteria necessary to cause disease (Sosnay et al. 2013). These results reiterate the need for clinical and functional assessments when determining the role of variants in human disease in the absence of strong genetic evidence.

Here, we demonstrate that S581L *GARS*, a variant previously implicated in CMT disease, does not segregate with disease in two newly identified families with dominant axonal CMT disease and is present at a low frequency in the general population (16/119,342). This variant was originally identified in individuals with lower limb predominant CMT disease (James et al. 2006), which is uncharacteristic of *GARS*-associated CMT2D. Interestingly, S581L has normal aminoacylation activity and localization in neurons, as demonstrated by its ability to charge tRNA at the same rate as wild-type GARS (Nangle et al. 2007) and to form puncta similar to wild-type GARS (Figure 4.3). Taken together, these data strongly indicate that this variant is not disease causing and warrant removing it from the list of disease-associated variants. Instead, S581L *GARS* should be tested alongside disease-associated alleles as a control in mechanistic studies.

In conclusion, our data indicate that loss of function is highly correlated with *GARS*-mediated CMT disease and that loss of ARS function is likely a necessary first step for disease pathogenesis. However, additional research is needed to determine the mechanism by which impaired tRNA charging leads CMT2D disease (see Chapter 5). Importantly, determining the precise mechanism of mutant *GARS* pathogenicity in axons will be essential for designing effective therapies for patients with CMT2D.

CHAPTER 5

Evaluating the pathogenic mechanism of ARS-mediated CMT disease

INTRODUCTION

A growing body of evidence, both previously published and included in this work, demonstrates that disease-associated ARS mutations result in impaired enzyme function (Antonellis et al. 2003; Jordanova et al. 2006; Latour et al. 2010; McLaughlin et al. 2011; Vester et al. 2012; Gonzalez et al. 2013). In total, 30 dominant disease-associated ARS alleles have been functionally characterized revealing that: (i) 15 result in impaired tRNA charging in biochemical assays (Jordanova et al. 2006; Nangle et al. 2007; McLaughlin et al. 2010; McLaughlin et al. 2011; Griffin et al. 2014); (ii) 19 impair yeast cell growth (Antonellis et al. 2006; Jordanova et al. 2006; McLaughlin et al. 2010; McLaughlin et al. 2011; Vester et al. 2012; Gonzalez et al. 2013; Griffin et al. 2014; Motley et al. 2015; Safka Brozkova et al. 2015); (iii) nine mislocalize in cultured neurons (Antonellis et al. 2006; Jordanova et al. 2006; Nangle et al. 2007) (Griffin et al. 2014); and (iv) two are unable to complement deletion of the fly ortholog (Chihara et al. 2007). Together, 28 of 30 ARS variants show reduced function in at least one assay, indicating that impaired ARS function is a component of ARS-mediated CMT disease. Interestingly, the two mutations that have not shown loss-of-function characteristics, A57V and D500N *GARS*, are both reported in general populations and additionally, A57V has weak genetic evidence for a role in dominant CMT disease. However, the precise mechanism by which impaired ARS function

leads to CMT disease is unknown. Thus, studies focused on determining the pathology of ARS-mediated CMT disease should focus on investigating mechanisms of impaired function in dominant disease.

There are two major mechanisms by which loss-of-function mutations lead to dominant disease: haploinsufficiency and a dominant-negative effect. Mice heterozygous for a gene-trap null *Gars* allele are phenotypically normal despite having a ~50% reduction in *Gars* protein levels (Seburn et al. 2006). This suggests that haploinsufficiency is not the pathological mechanism in *GARS*-mediated CMT disease. In a dominant-negative mechanism, a loss-of-function mutant protein binds the wild-type gene product and antagonizes its function, depleting overall enzyme activity in the cell to ~25-33% of normal levels. Lymphoblastoid cell lines from patients with *GARS*-associated CMT disease (+/G240R) have an ~87% reduction in aminoacylation *in vitro* compared to control cell lines (Ya-Ming Hou, personal communication). A dominant-negative mechanism is further supported by data that mice heterozygous for *Gars* loss-of-function missense mutations (C201R and P234KY) demonstrate a dominant axonal neuropathy, (Seburn et al. 2006; Achilli et al. 2009), while mice heterozygous for a null gene-trap *Gars* allele are phenotypically normal, indicating that expression of a mutant protein is required for *GARS*-mediated CMT disease. Furthermore, ARS null alleles (*e.g.*, deletions and nonsense mutations) in patients with ARS-mediated CMT disease (Wallen and Antonellis 2013) have not been identified. In fact, all ARS missense mutations tested to date result in stable proteins. To exert a dominant-negative effect, the mutant gene product must dimerize with the wild-type gene product. All ARS enzymes with strong genetic evidence for a role in CMT disease (*GARS*, *YARS*, *AARS*, and *HARS*) function as oligomers (Antonellis et al. 2003; Jordanova et al. 2006; McLaughlin et al. 2011; Vester et al. 2012) and mutant GARS and YARS proteins maintain the

ability to form wild-type/mutant heterodimers (Jordanova et al. 2006; Nangle et al. 2007). Importantly, there is direct experimental evidence in support of a dominant-negative mechanism in *YARS*-mediated CMT disease. Pseudo-heterozygous yeast expressing wild-type and mutant *YARS* alleles demonstrated decreased growth (Jordanova et al. 2006). However, *GARS, AARS,* and *HARS* mutations when co-expressed with wild-type ARS have not demonstrated a dominant-negative effect in yeast. Still, the experimental evidence warrants a thorough investigation of a dominant-negative mechanism in ARS-mediated CMT-disease.

In this chapter we will investigate a dominant-negative mechanism for *GARS* mutations using a *trans*-complementation assay in yeast. Data suggests that dramatically reduced levels of ARS activity may be sufficient to support yeast cell growth (McLaughlin et al. 2010). Therefore, co-expression of wild-type and mutant ARS may not be sufficient to deplete ARS activity to below the threshold for yeast viability. We hypothesized that a dominant-negative effect would be observed by expressing mutant ARS in a sensitized yeast strain whose ARS activity is already greatly reduced. To test this hypothesis, we will express loss-of-function, missense GRS1 alleles in a hypomorphic *GRS1* yeast strain and assess for effects on yeast viability.

The author performed all the studies in this chapter with the following exception: (1) The University of Michigan Sequencing core performed the DNA sequencing reactions; and (2) Δ*GRS1* haploid yeast strain was previously generated by the Schimmel Laboratory at Scripps Institute (Turner et al. 2000).

MATERIALS AND METHODS

GRS1 expression constructs

GRS1/pDONR221 construct was generated as described in chapter 4. Briefly, the yeast *GRS1*

locus was amplified with primers containing Gateway adapter sequences (primer sequences included in Appendix A), cloned into pDONR221 using BP clonase, and sequence verified. Mutation-containing oligonucleotides were generated, and the QuickChange II XL Site-Directed Mutagenesis Kit (Stratagene, Santa Clara, CA) was used as per the manufacturer's instructions to induce the desired mutations in the pDONR construct (primer sequences included in Appendix A). After transformation into *E. coli*, DNA from individual clones was purified and sequenced to confirm the presence of each mutation and the absence of any cycle-induced errors. Subsequently, validated entry clones were purified and recombined into the appropriate Gateway-compatible vector using LR clonase per manufacturer's specifications. pRS315 (*LEU2*-bearing vector) and pRS414 (*TRP2*-bearing vector) were used for the yeast complementation assays. DNA from the resulting expression constructs was purified and digested with the restriction enzyme *Bsr*GI (New England Biosystems) to confirm the presence of the appropriate insert.

Yeast complementation assays

Yeast complementation assays were used to test the activity of *GRS1* alleles (L129P, H418R, G526R, L129P:H418R *cis*, and L129P:H526R *cis*) as previously described in chapter 4 (Antonellis et al. 2006). The Δ*GRS1* haploid yeast strain [harboring a pRS316 maintenance vector to express wild-type *GRS1* and *URA3* (Turner et al. 2000)] was transformed with 200 ng of wild-type or mutant *GRS1* in a *LEU2*-bearing pRS315 vector and selected on medium lacking uracil and leucine (Teknova, Hollister, CA, USA). For each transformation, two colonies were selected for further analysis. Each colony was grown to saturation in –leu-ura selection medium for 48 hr. Next, 10 μl of undiluted and diluted (1:10 and 1:100 in H_2O) samples from each culture were spotted on plates containing 0.1% 5-FOA complete medium or SD –leu -ura growth

medium (Teknova) and incubated at 30°C for 72 hr. Yeast cell growth was determined by visual inspection.

Yeast trans-*complementation assays*

The Δ*GRS1* haploid yeast strain [harboring a pRS316 maintenance vector to express wild-type *GRS1* and *URA*] (described above) (Turner et al. 2000; Antonellis et al. 2006) was transformed with a 200 ng of a *TRP2*-bearing pRS414 vector containing wild-type, L129P, H418R, or G526R *GRS1* or pRS414 with no insert and were selected on plates lacking uracil and trytophan. The resulting pseudo-heterozygous strains (containing both the *URA3*- and *TRP2*-bearing vectors) were transformed with 200 ng of a *LEU2*-bearing vector containing a *GRS1* allele (wild-type, L129P, H418R, G526R, L129P:H418R *cis*, or L129P:G526R *cis*) or no insert. Yeast were grown on media lacking uracil, leucine, and tryptophan for 3 days at 30°C to select for yeast harboring all three vectors. Resulting yeast were grown to saturation for 2 days in media lacking leucine, uracil, and tryptophan at 30°C and subsequently spotted undiluted or diluted (1:10 or 1:100 in H_2O) on plates containing 5-FOA and lacking leucine and tryptophan to select for the spontaneous loss of the *URA3*-bearing maintenance vector (Boeke et al. 1987) and the presence of the two experimental vectors or on media lacking leucine, uracil, and tryptophan as a control. Yeast were grown for 3 days at 30°C and visual inspection of growth was used to determine the effect of co-expression of mutant *GRS1* alleles.

RESULTS

A subset of GRS1 *alleles complement in* trans *to support yeast viability.*

A body of evidence suggests that a dominant-negative mechanism may be involved in ARS-mediated CMT disease. In fact, previous studies have demonstrated that co-expression of mutant and wild-type *YARS* lead to decreased viability in yeast (Jordanova et al. 2006). We hypothesized

that if a dominant-negative effect is involved in *GARS*-mediated CMT disease, the growth of a hypomorphic strain would be further suppressed by the co-expression of a loss-of-function, missense ARS allele. To assess *GARS* mutations for a dominant-negative effect *in vivo*, we utilized a yeast strain that expresses the hypomorphic *GARS* allele L129P (Antonellis et al. 2003). When modeled in yeast (Table 5.1), L129P results in decreased cell growth compared to wild-type *GARS* expression (Figure 5.1). We selected two disease-associated *GARS* mutations, H418R and G526R, which were previously shown to be unable to support yeast cell viability (Antonellis et al. 2006) (Table 5.1; Figure 5.1), to co-express in our sensitized L129P *GRS1* yeast strain. Importantly, H418R and G526R *GARS* are known to retain dimerization ability (Nangle et al. 2007; Xie et al. 2007). A *LEU2*-bearing vector containing wild-type, L129P, H418R or G526R *GRS1* or no insert was introduced into the L129P strain or a wild-type control strain (Figure 5.2A). After selection on 5-FOA media, yeast expressing wild-type *GRS1* from two vectors (wild-type/wild-type) showed robust growth, while yeast containing an L129P-expressing vector and an empty vector (L129P/Empty) showed intermediate growth, consistent with previous observations for L129P yeast strains (Figure 5.2B) (Antonellis et al. 2006). Expression of L129P from two plasmids (L129P/L129P) improved yeast viability, but did not completely rescue yeast growth to wild-type levels, indicative of a gene dosage effect of L129P *GRS1* on yeast viability. Surprisingly, co-expression of L129P and either H418R (L129P/H418R) or G526R (L129P/G526R) resulted in robust growth similar to wild-type levels (Figure 5.2A). To determine if H418R and G526R retained some basal level of activity that was not sufficient to sustain yeast cell growth independently but that could increase overall enzyme activity in an additive manner based on gene dosage, yeast strains expressing H418R or G526R from two plasmids (H418R/H418R or

Table 5.1. Human *GARS* variants modeled in the yeast ortholog *GRS1*

Human GARS[a]	Yeast GRS1[b]
L129P	L79P
H418R	H402R
G526R	G521R

[a]Amino acid coordinates correspond to GenBank accession number AAA57001.1.
[b]Amino acid coordinates correspond to GenBank accession number NP_009679.2.

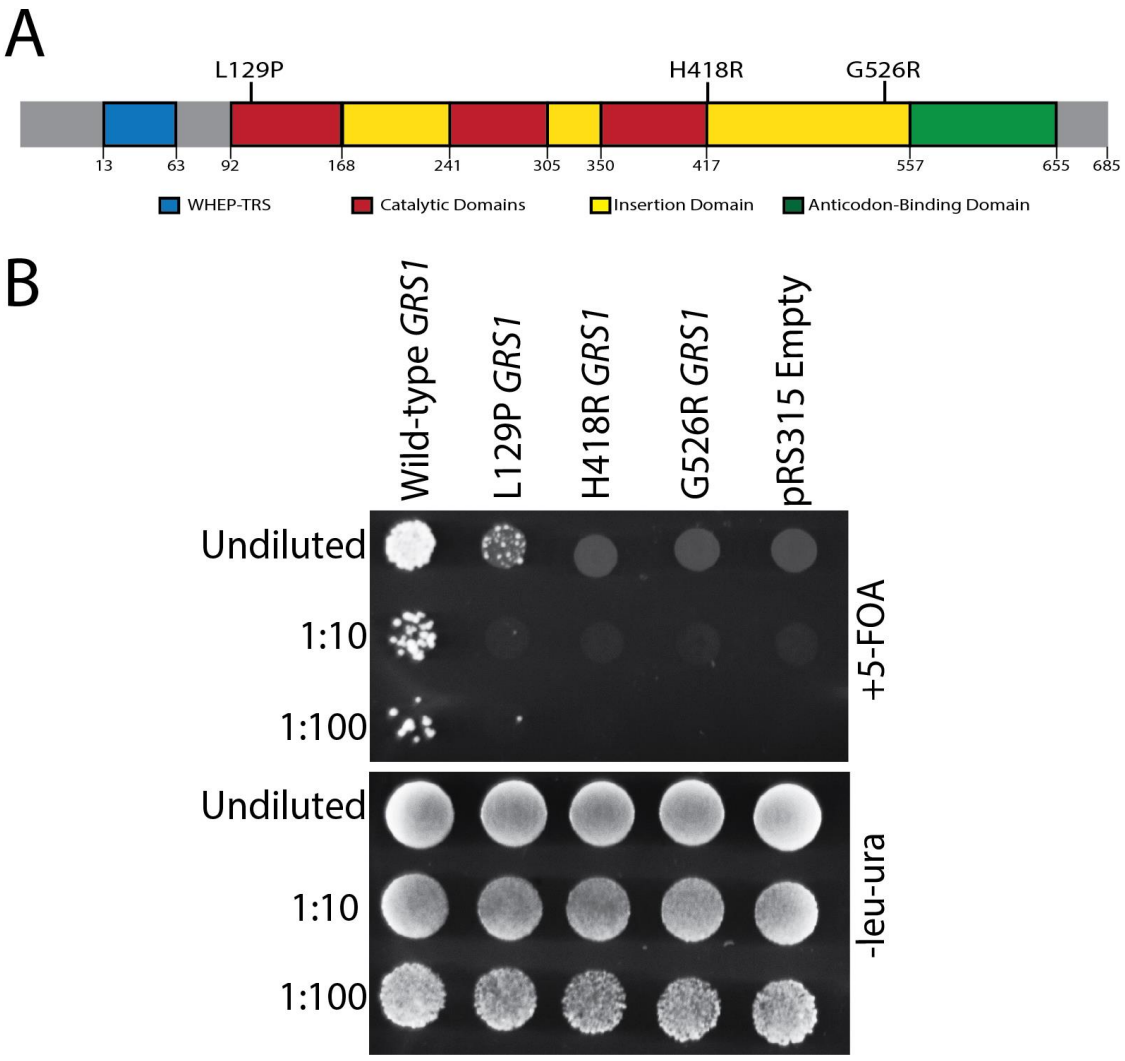


Figure 5.1. Confirmation of mutant *GRS1* phenotypes in yeast. (A) L129P, H418R, and G526R *GARS* are mapped (top) on a cartoon depiction of the GARS protein domains that includes the WHEP domain (blue), catalytic domains (red), insertion domains (red), and anticodon-binding domain (green). The amino acid positions of the domains are listed below the cartoon. (B) L129P, H418R, and G526R *GRS1* were transformed into haploid Δ*GRS1* yeast and selected on 5-FOA media to confirm the phenotypes previously reported in yeast for these alleles (Antonellis et al., 2006). Note the intermediate growth pattern for yeast expressing L129P *GRS1*.

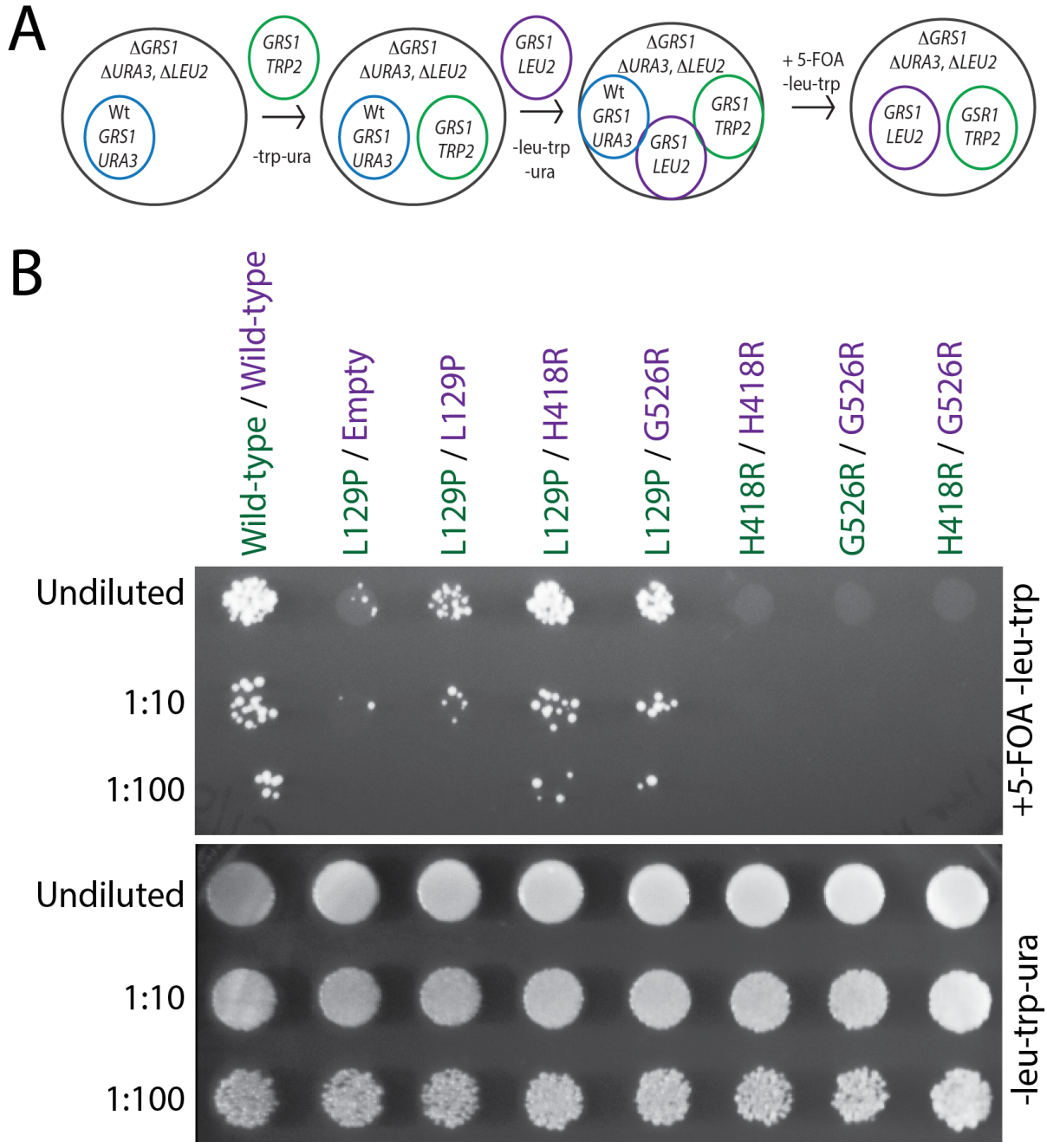

A
ΔGRS1
ΔURA3, ΔLEU2
Wt
GRS1
URA3
GRS1
TRP2
-trp-ura
GRS1
LEU2
-leu-trp
-ura
+ 5-FOA
-leu-trp
GSR1
TRP2
B
Wild-type / Wild-type
L129P / Empty
L129P / L129P
L129P / H418R
L129P / G526R
H418R / H418R
G526R / G526R
H418R / G526R
Undiluted
1:10
1:100
+5-FOA -leu-trp
-leu-trp-ura

Figure 5.2. Yeast *trans*-complementation assay of *GRS1* alleles. (A) The cartoon depicts the yeast *trans*-complementation assay used in this study. A haploid Δ*GRS1* yeast strain expressing wild-type *GRS1* from a *URA3*-bearing vector (blue) is transformed with *TRP2*-bearing pRS414 plasmid containing either wild-type, L129P, H418R, or G526R *GRS1* (purple) or no insert and is selected on media lacking tryptophan and uracil. The resulting yeast are subsequently transformed with a *LEU2*-bearing pRS315 vector containing a wild-type, L129P, H418R, or G526R *GRS1* allele or no insert ('Empty') and yeast are selected on media lacking uracil, tryptophan, and leucine. Individual yeast colonies are then grown to saturation in media lacking uracil, tryptophan, and leucine (not shown) and spotted on plates lacking tryptophan and leucine and containing 5-FOA to select for yeast containing the two experimental *GRS1* alleles and lacking the *URA3* maintenance vector (Boeke et al. 1987). (B) Yeast carrying all three vectors were spotted undiluted or diluted (1:10 and 1:100) on media lacking uracil, tryptophan, and leucine (–leu-trp-ura) or media lacking tryptophan or leucine and containing 5-FOA (+5-FOA-leu-trp). The allele on the pRS414 vector is in green text, while the allele on the pRS315 vector is in purple text. Note the robust growth for yeast co-expression L129P and either H418R (L129P/H418R) or G526R (L129P/G526R).

G526R/G526R) or co-expressing H418R and G526R (H418R/G526R) were generated. In contrast to L129P/H418R or L129P/G526R expressing yeast, H418R/H418R, G526R/G526R, and H418R/G526R yeast strains were unable to grow on 5-FOA (Figure 5.2). Taken together, these data suggest that L129P and H418R or G526R can *trans*-complement in yeast to improve enzymatic activity.

One explanation for the *trans*-complementation observed is intragenic complementation (Yook 2005). Intragenic complementation occurs when two mutant alleles are in *trans* in the same gene locus and each allele itself produces a defective gene product; however, when the two defective products associate, a product with increased activity is produced. We hypothesized intragenic complementation was occurring in our systems via one of two mechanisms. In the first mechanism, the H418R or G526R gene product stabilizes the L129P product to improve L129P catalytic function (*i.e*. structural complementation). Alternatively, wild-type motifs on the opposing subunits could compensate in *trans* for the lack of function in the opposing GARS subunit (*i.e.* functional complementation). All three of the tested mutations affect residues in different motifs of GARS (Figure 5.1). L129P GARS is located in motif 1 of the catalytic domain, while H418R is located in insertion domain 3, and G526R resides in motif 3 of the catalytic domain (Xie et al. 2007).

To distinguish the mechanism by which *trans*-complementation was occurring in our yeast model, we utilized a genetic approach. We hypothesized that if the mechanism was structural complementation, the H418R or G526R mutations would have the same stabilizing effect on the L129P mutation even when present in *cis* (*i.e.* L129P:H418R *cis* homodimers would be functional). Conversely, if the mechanism was functional complementation by opposing

domains, mutations in *cis* would ablate the synergistic growth effect when expressed in *trans* with L129P (*i.e.* L129P/L129P:H418R heterodimers would support yeast growth only to levels equivalent to L129P homodimers). First we generated and assessed the function of the L129P: H418R and L129P:G526R *cis* alleles. When expressed in the Δ*GRS1* yeast strain, neither the *GRS1* L129P:H418R *cis* allele, nor the L129P:G526R *cis* allele was able to complement deletion of endogenous *GRS1* (Figure 5.3). This result indicated that the presence of H418R or G526R in the dimer were not sufficient to stabilize and improve the function of L129P on the opposing subunit. To test for functional *trans*-complementation between opposing subunits, we co-expressed L129P and either the L129P:H418R *cis* or L129P:G526R *cis* allele. Important to note, in yeast co-expressing L129P and a *cis* allele, three dimer states exist. For example, co-expression of L129P and L129P/H418R would results in L129P/L129P homodimers, L129P/ L129P:H418R heterodimers, and L129P:H418R/ L129P:H418R homodimers. Since we already know that L129P:H418R/ L129P:H418R homodimers are non-functional and L129P/L129P homodimers have an intermediate level of activity, if complementation is occurring in the heterodimer, we would expect an increase in growth compared to the presence of L129P/L129P dimers only. Yeast co-expressing L129P and a *cis* allele (L129P/L129P:H418R or L129P/L129P:G526R) grew to comparable levels as L129P expression alone (Figure 5.4), indicating that the L129P/*cis* allele heterodimers are not functional. Taken together, we conclude that the mechanism of intragenic complementation is the result of functional complementation by wild-type motifs on opposing mutant subunits.

DISSCUSSION

In this chapter, we tested *GARS* mutations for a dominant-negative effect. We developed a yeast *trans*-complementation assay to assess for the ability of null *GRS1* alleles to suppress activity of

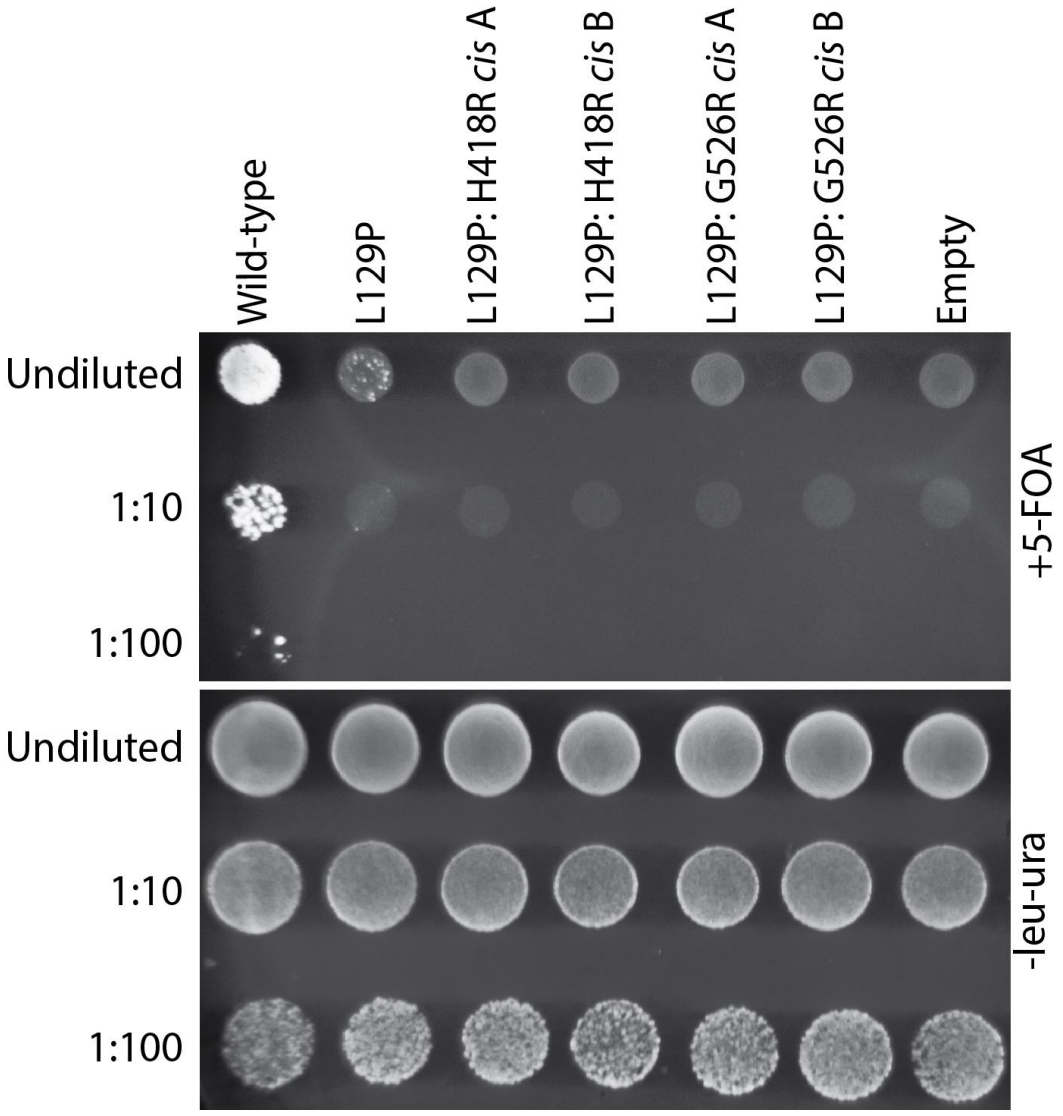


Figure 5.3. Yeast complementation assays using *GRS1 cis* alleles. To test the effects of null mutations on the stabilization of the L129P, *cis* alleles containing both L129P and either H418R (L129P:H418R) or G526R (L129P:G526R) were generated. Yeast transformations of the haploid Δ*GRS1* yeast strain were performed to introduce a *LEU2*-bearing vector containing wild-type, L129P, L129P:H418R, L129P:G526R or no insert ('Empty') and resulting yeast were spotted on 5-FOA or –leu-ura plates to assess the complementation ability of the *cis* alleles. Two independently generated plasmids of each *cis* allele were used (A and B).

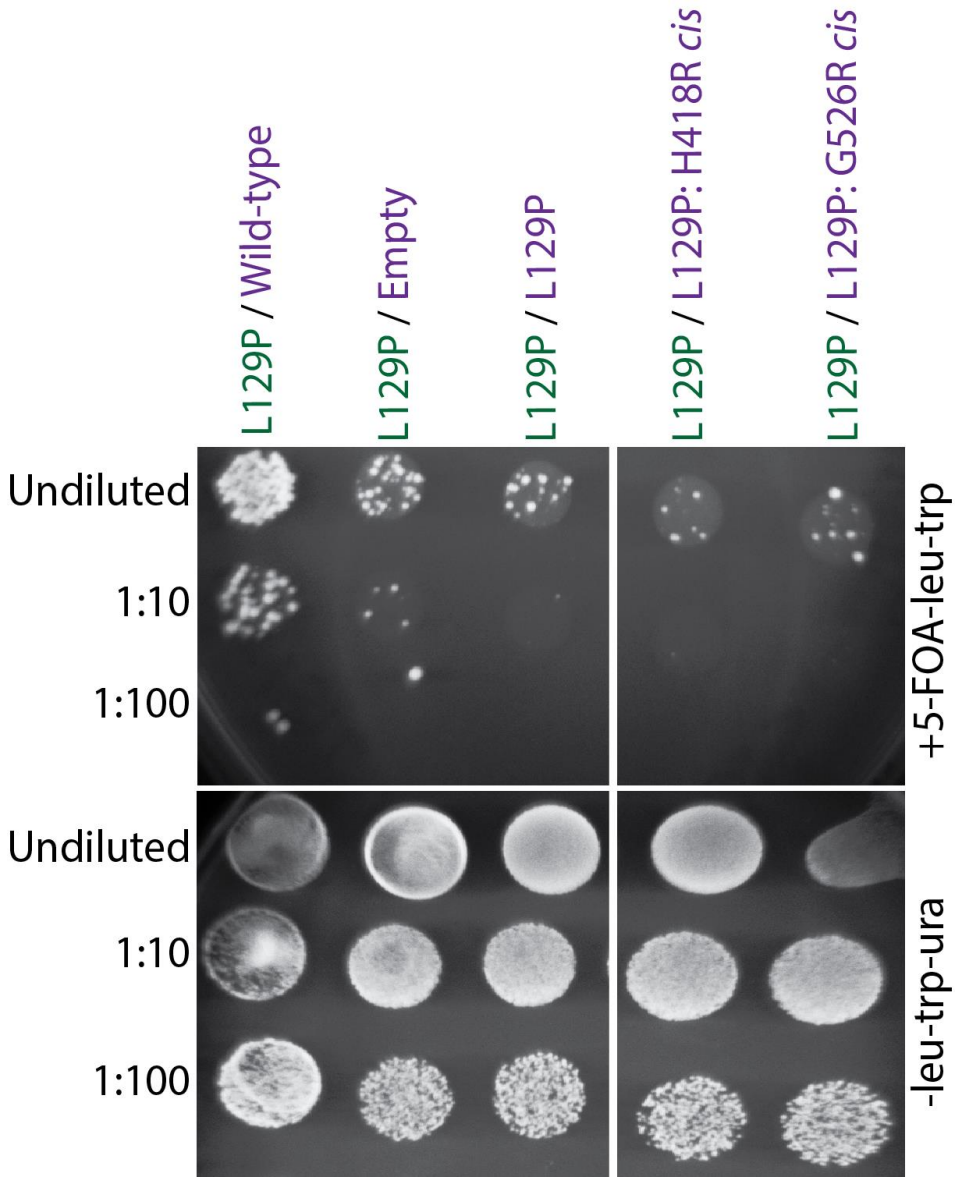


Figure 5.4. Yeast *trans*-complementation assays with *GRS1 cis* alleles. A haploid Δ*GRS1* yeast strain containing a *URA3*-maintenance vector was transformed with a L129P *GRS1* allele on the *TRP2*-bearing pRS414 vector (green) and either wild-type, L129P, L129P:H418R, L129P:G526R or no insert on a *LEU2*-bearing pRS315 vector (purple). Yeast carrying all three vectors were spotted undiluted or diluted (1:10 and 1:100) on media lacking uracil, tryptophan, and leucine (–leu-trp-ura) or media lacking tryptophan or leucine and containing 5-FOA (+5-FOA-leu-trp) media.

a hypomorphic *GRS1* allele. Contrary to our hypothesis that antagonization would occur, co-expression of the hypomorphic L129P *GRS1* allele and the null H418R or G526R allele improved yeast viability to wild-type levels. A series of genetic experiments testing mutations in *cis* and *trans*, revealed that the intragenic complementation was due functional motifs on each subunits compensating for the lack of function on the opposing subunit to improve heterodimer activity.

This is not the first report of intragenic complementation of *GARS* mutations. In mouse models, homozygosity for either C201R or P234KY *Gars*, missense *Gars* alleles that result in a dominant peripheral neuropathy in heterozygous animals, results in reduced viability (Seburn et al. 2006; Achilli et al. 2009); however, animals that are compound heterozygous for these two alleles have normal viability, but decreased lifespan, indicating that C201R/P278KY heterodimers have increased function compared to homodimers of either mutation (Motley et al. 2011). Importantly, C201R and P234KY *Gars* affect different functional motifs in GARS, insertion domain I and motif 2 of the catalytic domain, respectively (Xie et al. 2007). Although viable, compound heterozygous mice have a more severe disease phenotype and a decreased life span compared to mice heterozygous for either missense allele, indicating that GARS activity in compound heterozygotes was not increased to levels equivalent to those in heterozygous cells (Motley et al. 2011). The observation of intragenic complementation in two *in vivo* systems raises the hypothesis that for each domain of the GARS dimer, only one wild-type subunit is needed for normal GARS activity levels. Thus wild-type/mutant heterodimers would be hypothesized to be fully functional. This would refute the dominant-negative hypothesis as overall GARS activity in heterozygous cells would be 66-75% of wild-type, an overall activity greater than in the cases of haploinsufficiency. This level of activity is unlikely as experimental evidence from G240R

GARS heterozygous lymphoblastoid cells shows an 87% percent decrease in aminoacylation compared to wild-type cells (Ya-ming Hou, personal communication). To further explore this observation, collaborators are assessing GARS activity in tissues from *Gars* heterozygous mice (Seburn et al. 2006; Achilli et al. 2009). Importantly, these studies will not only inform us regarding the dominant negative hypothesis, but may also provide valuable insight into the tissue specificity of ARS-mediated disease. Furthermore, our yeast data actually argue against the hypothesis that only one wild-type motif per dimer is necessary for dimer function. Specifically, the observation that when expressed in *trans*, H418R/G526R dimers do not increase yeast viability. These data indicate that the presence of a wild-type motif on the opposing subunit is not sufficient for complementation in all cases. These studies have neither implicated nor refuted the dominant negative hypothesis and thus efforts should continue to understanding how and why an expressed stable protein is necessary for disease pathogenesis.

Recently, collaborators in Dimitris Beis' laboratory at the Biomedical Research Foundation Academy of Athens (Athens, Greece) identified a mutation in zebrafish *gars*, T209K, which results in abnormal cardiac development and embryonic lethality in homozygous zebrafish. Interestingly, heterozygous animals are phenotypically normal (Malissovas et al., in press). Studies on the zebrafish T209K Gars protein revealed that T209K is unable to dimerize with itself or wild-type Gars. We performed functional studies and determined this mutation has impaired function in yeast complementation and localization assays. Therefore, this mutation provided an excellent opportunity to experimentally assess if dimerization was critical in ARS-mediated disease. To do so, zebrafish embryos were injected with zebrafish *gars* mRNA harboring the T209K mutation or the CMT-associated mutation G526R, a loss-of-function allele that has increased dimerization capacity (Xie et al. 2007). Overexpression of G526R *gars*, but

not T209K *gars*, resulted in dominant neurotoxicity. Together, these data suggest that both loss-of-function and an ability to dimerize are important components of *GARS*-mediated neurotoxicity in zebrafish.

Zebrafish and mouse studies argue for a dominant-negative mechanism in *GARS*-mediated CMT while our yeast studies offer some evidence to refute this hypothesis. This may be due to the use of yeast *GRS1* versus vertebrate *GARS* orthologs in the respective studies. We know that CMT-associated *GARS* mutations modeled in *GRS1* do not always recapitulate the functional consequences of those mutations in the human protein. For example, G240R and I280F *GRS1* complement loss of endogenous *GRS1,* suggesting these mutations are functional in *vivo;* however, when modeled in the truncated human *GARS* open reading frame that has been shown to complement in yeast (Chien et al. 2014), both G240R and I280F *GARS* fail to support yeast growth indicating loss of function (Stephanie Oprescu, personal communication). Therefore, we must repeat the dominant negative studies in the humanized yeast model to test a dominant negative effect using human wild-type and mutant *GARS*. Importantly, *GARS* variants described in this thesis, such as T209K (dimerization mutant) and S581L (functional, non-pathogenic mutant), would serve as excellent controls for these and other mechanistic studies.

In conclusion, in this chapter we provide evidence in yeast that some mutant alleles can functionally complement one another, casting doubt on the role of a dominant-negative effect in ARS-mediated CMT disease. However, understanding the effects of the human wild-type/mutant heterodimer remains an open question. We are optimistic that the recent developments in the yeast tools and variant identification are finally sufficient to properly address the dominant-negative hypothesis for ARS-mediated disease.

CHAPTER 6

Investigating the role of GARS puncta in axonal health and disease

INTRODUCTION

Although ARS enzymes are ubiquitously expressed, essential enzymes, heterozygosity for loss-of-function ARS mutations leads to an axon-specific neuropathy. A growing body of literature demonstrates that disruption of RNA metabolism or protein translation leads to neurodegenerative diseases (Wang et al. 2007; Scheper et al. 2007a; Liu-Yesucevitz et al. 2011). For example, mutations in *CLP1*, a RNA kinase involved in tRNA splicing, results in a neurological syndrome affecting both the central and peripheral nervous system, including axonal sensory and motor neuropathy (Hanada et al. 2013; Karaca et al. 2014; Schaffer et al. 2014). One possible explanation for this observation is that neurons, especially those in the peripheral nervous system (Antonellis and Green 2008), are uniquely sensitive to impaired protein translation (Scheper et al. 2007b). These latter cells are long-lived, terminally differentiated, and harbor axons that must be supported over long distances and time. It is known that mRNA is transported to axons (Ingoglia et al. 1983) and proteins are translated locally in response to cell signals (Liu-Yesucevitz et al. 2011). Local protein translation is highly regulated and disruption of multiple aspects of this process, from mRNA transport (*e.g. FUS* mutations in ALS (Vance et al. 2009)) to translation repressors (*e.g. FMR1* mutations in Fragile X Syndrome (Garber et al. 2006)) have been implicated in neurological disease. Why mutations in different components of this process lead to dramatically different neurologic phenotypes is currently

unknown.

Interestingly, endogenous wild-type GARS proteins localize to 'puncta' within axons in the human peripheral nerve (Antonellis et al. 2006). When modeled *in vitro*, seven of the 12 *GARS* mutations impair the punctate localization in cultured neurons, suggesting that reduced localization to axons is a component of CMT2D and dSMA-V disease pathogenesis (Antonellis et al. 2006; Antonellis and Green 2008; Wallen and Antonellis 2013) (see chapter 4). The remaining five mutant GARS proteins form puncta in MN-1 cell bodies and neurite projections, consistent with previous observations of wild-type GARS localization to axons (Antonellis et al. 2006; Fallini et al. 2012; Griffin et al. 2014). Since proper localization within a cell can be critical for protein function, impaired localization can result in a cell compartment-specific loss of function (reviewed in (Hung and Link 2011)). Interestingly, while no strong correlation between genotype and phenotype has been appreciated for *GARS*-mediated peripheral neuropathy, G598A and 245-248ΔETAQ, the two GARS mutations that cause infantile spinal muscular atrophy (SMA), both fail to form puncta, indicating loss-of-puncta formation may correlate with a more severe phenotype. The molecular components and function of GARS puncta is unknown. Therefore, it is critical to identify proteins that interact with GARS in puncta to help define the function of these structures and begin to understand how disruption may lead to axonal dysfunction. Furthermore, these studies may be important in determining potential drug targets for the treatment of *GARS*-mediated CMT disease.

Prior to initiation of this project, our lab performed proteomic analysis on lymphoblastoid cell lines derived from patients heterozygous for G240R *GARS*, a mutation that disrupts puncta formation (Antonellis and Green 2008). These studies revealed reduced levels of 22 proteins with known roles in neuron function, including DYNC1H1 (a molecular motor involved in

retrograde axonal transport (Michaelevski et al. 2010)) and RAB7 (a protein that regulates vesicle trafficking (Cantalupo et al. 2001)). Interestingly, mutations in *DYNC1H1* and *RAB7* are associated with CMT2O (Weedon et al. 2011) and CMT2B (Spinosa et al. 2008), respectively. mRNA for both genes localizes to axons, suggesting that they are translated locally (Gumy et al. 2011). Furthermore, recent studies have demonstrated decreased protein translation in the motor neurons of *gars* mutant flies (Niehues et al. 2015), but the mechanism by which this occurs remains to be elucidated. In archaea, AARS and SARS have been shown to associate with ribosomes, presumably to facilitate tRNA recycling and translation efficiency (Godinic-Mikulcic et al. 2014). Thus, CMT-associated *GARS* mutations may affect the axonal translation efficiency of key proteins by inhibiting association with ribosomal proteins and/or RNAs critical for efficient translation in axons. Determining if GARS interacts with ribosomal proteins and if mutant GARS disrupts these interactions will help us investigate the observation of decreased protein translation in patient cells and fly neurons.

The studies presented in this chapter focus on understanding the differential interacting partners of wild-type and mutant GARS that are important for puncta formation and/or failure to form these structures. Importantly, these studies have the potential to identify both loss- and gain-of-function interactions of mutant GARS. Identifying proteins that differentially interact with wild-type and mutant GARS will provide insight into basic GARS biology, ARS-mediated disease mechanisms, and potential avenues for therapeutic development.

Recently, 245-248ΔETAQ *GARS* was identified in a patient with severe early onset spinal muscular atrophy (SMA). Due to the severity of the phenotype, 245-248ΔETAQ *GARS* is an excellent candidate for initial immunoprecipitation / mass spectrometry studies comparing wild-type and mutant interacting partners. We hypothesize that this mutation will demonstrate the

functional characteristics appreciated of other CMT-associated *GARS* mutations, including decreased aminoacylation, failure to complement in yeast, and mislocalization in neurons, and will have altered protein interactions in MN-1 cells compared to wild-type GARS.

The author performed all the work presented in this chapter with the following exceptions: (1) Clinical characterization and mutation identification was performed by collaborators; (2) Aminoacylation assays were performed by Ya-ming Hou at Thomas Jefferson University; (3) Mass spectrometry was performed at the Proteomics Core at the Fred Hutchinson Cancer Institute in Seattle, Washington and the Proteomics Resource Facility at the University of Michigan; and (4) Bioinformatics of the mass spectrometry data was performed by Dattatreya Mellacheruvu in Alexey Nesvizhskii's laboratory at the University of Michigan.

MATERIALS AND METHODS

GARS *and* GRS1 *expression constructs*

DNA constructs were generated for aminoacylation, yeast complementation, localization, and immunoprecipitation studies. Vectors used included pET-21a(+) (aminoacylation assays), pEGFP-N2 (for expressing GARS with a C-terminal EGFP tag), pRS315 (yeast complementation assays), and pTM3xFLAG_GW (for expressing GARS with a C-terminal 3xFLAG tag for immunoprecipitations). For immunoprecipitation studies, pTM3xFLAG_GW vector was generated using the Gateway Vector Conversion System (Invitrogen) per manufacturer's instructions to insert Gateway Cassette B in frame with the 3xFLAG tag in the pTM3xFLAG vector (generously provided by Tomo-ichiro Miyoshi, Moran Laboratory, University of Michigan). Briefly, pTM3xFLAG was digested with XhoI and AfeI, blunted with the Quick Blunt enzyme kit (New England Biolabs, Ipswich, MA), gel purified, and ligated with Gateway Conversion Cassette B. Sequencing using primers designed to the backbone of

pTM3xFLAG upstream and downstream of XhoI and AfeI sites, respectively, were used to sequence verify the proper orientation of the gateway cassette (primer sequences in Appendix A).

To model 245-248ΔETAQ *GARS*, mutation-containing oligonucleotides were generated and the QuickChange II XL Site-Directed Mutagenesis Kit (Stratagene, Santa Clara, CA) was used as per the manufacturer's instructions (primer sequences in Appendix A) to mutate *GRS1*/pDONR221 and *GARS*-C-term/pDONR221. *GARS*-C-term/pDONR221 contains the open reading frame of human *GARS* but lacks a stop codon and is used to generate C-terminal fusion proteins. Native *GARS*/pDONR221, which contains the stop codon, was used to generate the untagged native *GARS* control plasmid for IP/mass spec studies (see Chapter 4 for cloning details of pDONR221 constructs). Sequence verification was performed to confirm the presence of each mutation and the absence of any cycle-induced errors. Subsequently, validated entry clones were purified and recombined into the appropriate Gateway-compatible vector (see above) using LR clonase per the manufacturer's specifications. DNA from the resulting expression constructs was purified and digested with the restriction enzyme *Bsr*GI (New England Biosystems) to confirm the presence of the appropriate insert and sequencing was performed to confirm the presence of in frame tags when appropriate.

Aminoacylation assays

Wild-type and 245-248ΔETAQ GARS human proteins were expressed in E. coli Rosetta 2 (DE3) pLys cells with a C-terminal in-frame His tag and purified with nickel affinity resin according to the manufacturer's protocol (Novagen, Rockland, MA). The T7 transcript of human $tRNA^{Gly/UCU}$ (UCU, anticodon) was prepared and purified as previously described (Hou et al. 1993), was heat denatured at 85°C for 3 min, and annealed at 37°C for 20 min before use.

Steady-state aminoacylation assays were monitored at 37°C in 50 mM HEPES (pH 7.5), 20 mM KCl, 10 mM $MgCl_2$, 4 mM DTT, 2 mM ATP, and 50 mM ^{3}H-glycine (Perkin Elmer, Waltham, MA) at a specific activity of 16,500 dpm/pmole. The reaction was initiated by mixing GARS enzyme (20–600 nM) with varying concentrations of tRNA (0.3–20 μM). Aliquots of a reaction mixture were spotted on filter papers, quenched by 5% trichloroacetic acid, washed, dried, and measured for radioactivity by a liquid scintillation counter (LS6000SC; Beckman Coulter Inc., Fullerton, CA). The amount of radioactivity retained on filter pad was corrected for quenching effects to determine the amount of synthesis of Gly-$tRNA^{Gly}$. Steady-state kinetics was determined by fitting the initial rate of aminoacylation as a function of tRNA concentration to the Michaelis–Menten equation (Schreier and Schimmel 1972).

Yeast complementation assays

The RJT3/II-1 haploid yeast strain [*MATα grs1::HIS3, his3Δ200, leu2Δ1, lys2Δ202, trpΔ63, ura3–52*, pTsscII-maint (*cen*, *GRS1, URA3*)] carrying a deleted endogenous GRS1 allele and wild-type GRS1 on a *URA3*-bearing pRS316 maintenance vector was previously reported (Turner et al. 2000; Antonellis et al. 2006). RJT3/II-1 was transformed with a *LEU2*-bearing pRS315 vector containing wild-type or *GARS* 245-248ΔETAQ *GRS1* (described above) or pRS315 with no insert as previously reported (Antonellis et al. 2006). Each transformation was performed at least three times with at least three independent plasmid DNA preparations. Four colonies were selected from each transformation for additional analysis. Each colony was grown to saturation in –leu -ura selection medium for 48 hours. Next, 10 μl of undiluted and diluted (1:10 and 1:100) samples from each culture were spotted on plates containing 0.1% 5-fluoroorotic acid (5-FOA) (Boeke et al. 1987) complete medium or SD -leu -ura growth medium (Teknova, Hollister, CA) and incubated at 30°C for 72 hr. Yeast cell growth was determined by

visual inspection.

Cell culture and protein localization studies

The mouse motor neuron, neuroblastoma fusion cell line (MN- 1) was cultured and transfected with constructs to express human wild-type or 245-248ΔETAQ *GARS* in-frame with a C-terminal enhanced green fluorescent protein (EGFP) tag as previously described in chapter 4 (Salazar-Grueso et al. 1991; Antonellis et al. 2006). After 48 hours, growth medium was removed, and cells were washed in 1× PBS and then incubated in 1× PBS/0.4% paraformaldehyde for 10 min at room temperature. Cells were washed in 1× PBS, co-stained with 300 nM DAPI for 5 min, washed again in 1× PBS, and finally coated with Pro-Long antifade reagent (Invitrogen). Images were obtained with an IX71 Inverted Microscope using cellSens Standard image software (Olympus, Center Valley, PA).

MN-1 transfections for immunoprecipitation

MN-1 cells were grown at 37°C in 5% CO_2 to 60% confluency in T-175 flasks in Dulbecco's modified eagle medium (DMEM) supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 100 U/ml penicillin, and 50 μg/ml of streptomycin (Invitrogen, Carlsbad, CA). Each flask was transfected using the lipofectamine 2000. Specifically, for each transfection reaction, 5 ml of OptiMEM I minimal growth medium and 50 μl of Lipofectamine 2000 were mixed and incubated together for 10 min at room temperature. 15 μg of wild-type, 245-248ΔETAQ or native *GARS* pTM3xFLAG plasmid was combined with 5 ml of OptiMEM I. Subsequently, the Lipofectamine-OptiMEM I mixture was combined with the plasmid-OptiMEM I solution to yield 10 ml of the final transfection solution and incubated at room temperature for 20 min. The 60% confluent T-175 flask of cells was washed in 1X PBS. The entire transfection solution was then added directly to the flask and incubated at 37°C in 5% CO_2 for 4 hours. After 4 hours, the

transfection solution was removed and replaced with 20 ml of complete DMEM media (described above). Cells were allowed to recover for 2 days at 37°C in 5% CO_2. After 2 days, cells were harvested using trypsin, centrifuged at 2000 rpm for 2 minutes, and washed twice with 1X PBS. For our initial mass spectrometry studies performed at the Proteomics Core at the Fred Hutchinson Cancer Institute, one T-175 flask per sample was used. For mass spectrometry studies performed at the University of Michigan Proteomics Resource Facility, four T-175 flasks were pooled prior to cell lysis for each experimental plasmid and used for immunoprecipitations performed according to the protocol described below that was appropriately scaled 4X to account for the increase in cellular protein input.

Immunoprecipitation

25 µl of Dynabeads Protein G Magnetic beads (Invitrogen) were used for each reaction (i.e. per cell lysate from one T175 flask). Using a magnetic rack, beads were washed twice in 1000 µl of wash solution (0.5% bovine serum albumin and 0.1% Triton X-100 in PBS) and subsequently resuspended in 1000 µl of the wash solution. 2 µg anti-Flag M2 antibody (Sigma, F3165) was added to the beads and incubated rocking at 4˚C overnight. Harvested transfected cells (see above) were resuspended in 1.5 ml lysis buffer [20 mM Tris-HCl (pH7.5) 2.5mM $MgCl_2$, 300mM KCl, 0.1% NP-40, 1mM DTT, 0.2mM PMSF, 1x Halt™ Protease Inhibitor Cocktail (EDTA free) (Thermo Scientific)] and incubated gently rocking at 4°C for 90 min. Cell lysate was centrifuged at 12,000 rpm for 5 min at 4°C. Protein containing supernatant was removed and utilized as the input for the subsequent IP reaction in a clean tube. Protein was quantified using the Pierce™ BCA Protein Assay Kit (Thermo Scientific) and 1 mg of protein was used as input for each sample. Western blot with anti-Flag M2 antibody (Sigma, F3165) (1:2500 dilution) was used to confirm expression of 3xFLAG tagged proteins in input samples. Immediately before

use, the αFlag-conjugated magnetic beads were washed twice with 500µl of lysis buffer using a magnetic rack, resuspended in 25 µl lysis buffer and added directly to the protein sample. The protein and beads were incubated at 4°C for 2 hours with gentle rocking. Using a magnetic rack, supernatant was removed from the beads and discarded. The beads were washed 5 times with 1 ml of wash buffer [20 mM Tris-HCl (pH7.5) 2.5mM $MgCl_2$, 300mM KCl, 0.1% NP-40, 1mM DTT]. Beads were resuspended in 30 µl of wash buffer per 25 µl of starting beads for samples destined for the Proteomics Core at the Fred Hutchinson Cancer Institute and 50 µl of wash buffer per 25 µl of starting beads for samples destined for the University of Michigan Proteomics Resource Facility. Competitive elution of the flag tagged protein was performed using 200 µg/ml of 3xFlag peptide and incubate for 3 hours at 4°C rocking gently. Using a magnetic rack, supernatant was removed and saved for further analysis.

Silver staining for analysis of IP products

To visualize proteins for band excision for mass spectrometry studies performed at the Proteomics Core at the Fred Hutchinson Cancer Institute, IP products were separated on a denaturing 4-20% Tris-Glycine gel and silver stained. Briefly, 30 µl of the IP elution supernatant was boiled with 20 µl of 2X Tris-glycine SDS Sample Buffer (Life technologies, Carlsbad, CA) and 2 µl β-mercaptoethanol. Samples were run on a Novex™ 4-20% Tris-Glycine gel (Life technologies) for 2.5 hours at 100V. Gels were stained with the SilverQuest™ Staining Kit (Invitrogen) per manufacturer's instructions. All gel incubations and washes were performed at room temperature with gentle agitation and all solutions were made using ultra pure H_2O (uH_2O). The gel was rinsed in uH_2O, and subsequently fixed overnight in 100 ml of fix solution (40% ETOH, 10% acetic acid) agitating gently at room temperature. Gel was washed in 100 ml of 30% ETOH for 10 min, incubated in 100 ml of sensitization solution (30% ETOH, 10%

Sensitization Solution), washed in 100 ml of 30% ETOH for 10 min, washed in 100 ml of uH_2O for 10 min, stained in 100 ml of staining solution (1% Staining Solution) for 15 min, washed in 100 ml of uH_2O for 20 sec, and developed in 100 ml of developing solution (10% Developer solution, 1 drop Developer Enhancer) until bands appeared. 10 ml of Stopper solution was then added directly to the developing solution and incubated for 10 min. For mass spectrometry, bands were excised for two independently generated samples sets (A and B in Figure 6.8) and sent to the Proteomics Core at the Fred Hutchinson Cancer Institute.

Trichloroacetic acid (TCA) precipitation

For mass spectrometry performed at the University of Michigan Proteomics Resource Facility, IP products were concentrated using a TCA precipitation protocol. 2 µl of 2% sodium deoxycholate was added to 200 µl of IP product and incubated on ice for 30 min. 20 µl of 100% (w/v) TCA was added and incubated on ice for 1 hour. The IP product was then centrifuged at max speed at 4˚C for 10 min. The supernatant was discarded and the remaining pellet was washed in 500 µl of acetone, and incubated on ice for 10 min. The acetone solution was centrifuged at max speed at 4˚C for 10 min. The resulting pellet was resuspended in 30 µl of a solution containing 20 mM HEPES and 8M Urea (pH=8.0). Three independently generated IP products for each wild-type, 245-248ΔETAQ, and untagged native GARS control sample were sent for analysis.

Mass Spectrometry and bioinformatics

Tandem mass spectrometry was performed using the Orbi-elite Tandem Mass Spectrometer (Thermo Scientific) at Proteomics Core at the Fred Hutchinson Cancer Institute and the Orbitrap Fusion Mass Spectrometer (Thermo Scientific) at the University of Michigan Proteomics Resource Facility. The open source proteomics search engine X!Tandem (The Global Proteome

Machine Organization), was used to match tandem mass spectrometry data with peptide spectra. Peptide Prophet was used to validate peptide assignments (Keller et al. 2002) and Protein Prophet was employed to group peptides into proteins (Nesvizhskii et al. 2003). ABACUS extracted spectral counts for the data for quantitative analysis (Fermin et al. 2011). Finally, interactions were scored to remove background and comparison of interacting proteins was performed to calculate the relative fold changes for protein interactions between wild-type and 245-248ΔETAQ GARS.

RESULTS

245-248ΔETAQ GARS *is a loss-of-function allele.*

A young female child presented with an inability to crawl and impaired hand function at 11 months of age. Clinical evaluation subsequently diagnosed her with infantile spinal muscular atrophy (SMA). Sequencing revealed that she is heterozygous for the *GARS* mutation 245-248ΔETAQ. Her fraternal twin brother is unaffected and is homozygous for wild-type *GARS*. 245-248ΔETAQ *GARS* results in the deletion of four amino acids in the glycine binding pocket that are responsible for amino acid binding in the catalytic domain of GARS (Figure 6.1A) (Freist et al. 1996). Conservation analysis demonstrated that these residues are invariably conserved in all taxa from bacteria through human, illustrating the critical role they play in enzyme function (Figure 6.1B).

Since *GARS* mutations have been implicated in both CMT disease and SMA, including infantile SMA (Antonellis et al. 2006; James et al. 2006), we functionally characterized 245-248ΔETAQ *GARS* to determine if this mutation resulted in impaired function characteristic of disease-

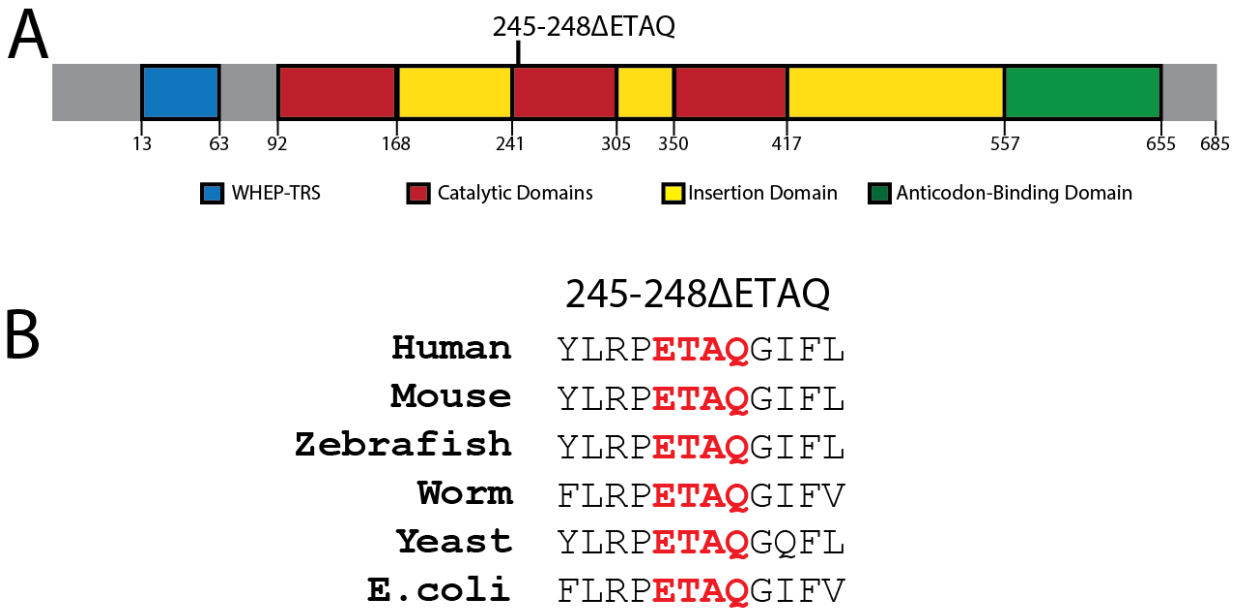


Figure 6.1. Localization and conservation of *GARS* 245-248ΔETAQ. (A) The cartoon displays the functional domains of GARS: WHEP-TRS (blue), catalytic domain (red), insertion domain (yellow), and anticodon-binding domain (green). 245-248ΔETAQ is located in the catalytic domain (top). (B) Alignments of evolutionarily diverse species illustrates that residues 245-248 ETAQ are conserved from human to *E. coli*.

associated *GARS* mutations. Aminoacylation assays revealed that 245-248ΔETAQ *GARS* had almost a complete loss of tRNA charging capacity compared to wild-type *GARS* (1/11,000) (Figure 6.2). When modeled in *GRS1*, the yeast ortholog of *GARS* (Table 6.1), 245-248ΔETAQ *GRS1* was unable to complement deletion of the endogenous *GRS1* allele (Figure 6.3). Finally, when assessed for localization patterns, 245-248ΔETAQ *GARS* failed to form puncta in MN-1 cells (Figure 6.4). Taken together, 245-248ΔETAQ *GARS* demonstrates loss-of-function characteristics in all three assays commonly used to study ARS variant activity and consistent with the characteristics previously described for disease-associated *GARS* mutations. Due to the severity of this mutation, we deemed that 245-248ΔETAQ *GARS* would be an excellent candidate for initial studies comparing protein-binding patterns of wild-type GARS and non-puncta forming mutant GARS.

Immunoprecipitation and mass spectrometry analysis of GARS proteins

To determine the differential protein binding partners of wild-type and 245-248ΔETAQ *GARS,* immunoprecipitation (IP) and mass spectrometry was performed. Specifically, wild-type and 245-248ΔETAQ *GARS* open reading frames lacking the stop codon and a wild-type *GARS* open reading frame containing the stop codon (native *GARS*) were cloned into the pTM3xFLAG_GW mammalian expression vector to generate plasmids for the expression of wild-type GARS-3xFLAG, 245-248ΔETAQ GARS-3xFLAG and untagged native GARS, respectively. MN-1 cells were transfected with each plasmid, grown for two days, and harvested. Protein lysates were isolated and a western blot was performed using an antibody directed at the FLAG tag to confirm the expression of ~87 kDa protein corresponding to GARS-3xFLAG and 245-248ΔETAQ GARS-3xFLAG and absence of a tagged peptide in the untagged native GARS sample (Figure 6.5). An antibody targeting the C-terminus of GARS was used to confirm the

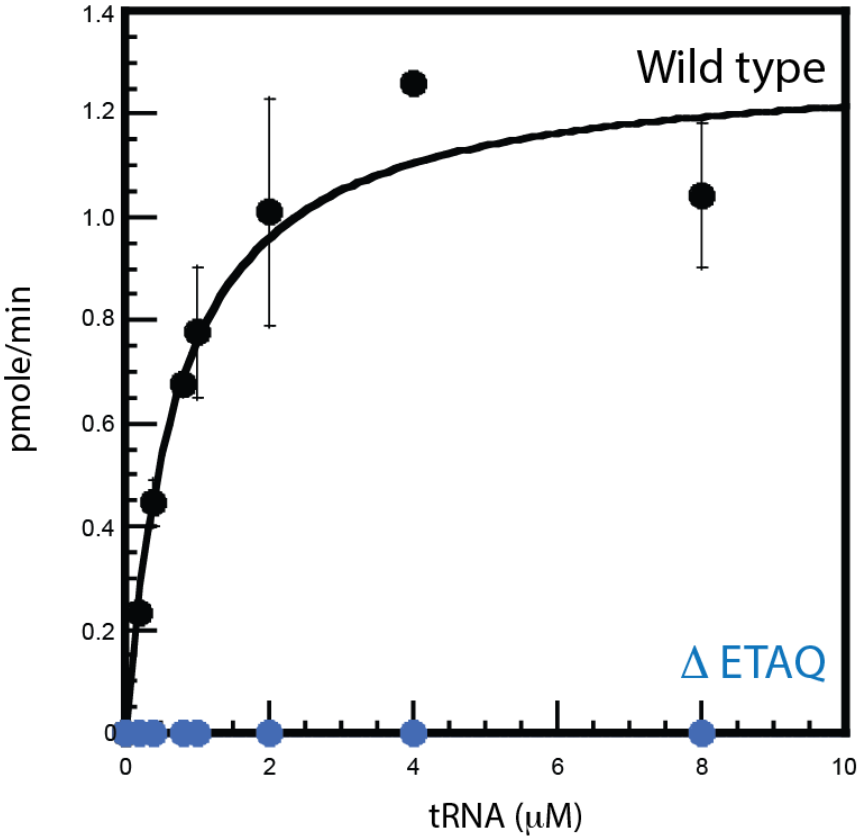


Figure 6.2. Aminoacylation assays of 245-248ΔETAQ *GARS*. Human recombinant wild-type and 245-248ΔETAQ GARS were assessed for tRNA charging activity *in vitro*. Using the Michaelis-Menten equation, the velocity of aminoacylation (pmole/min/pmole of enzyme) of $tRNA^{gly}$ was calculated as a function of tRNA concentration for wild-type GARS (black) and 245-248ΔETAQ GARS (blue). Error bars indicated standard deviation.

Table 6.1. Human 245-248ΔETAQ *GARS* Modeled in the Yeast Ortholog *GRS1*

Human GARS[1]	Yeast GRS1[2]
245-248ΔETAQ	228-331ΔETAQ

[1]Amino-acid coordinates correspond to GenBank accession number AAA57001.1
[2]Amino-acid coordinates correspond to GenBank accession number NP_009679.2

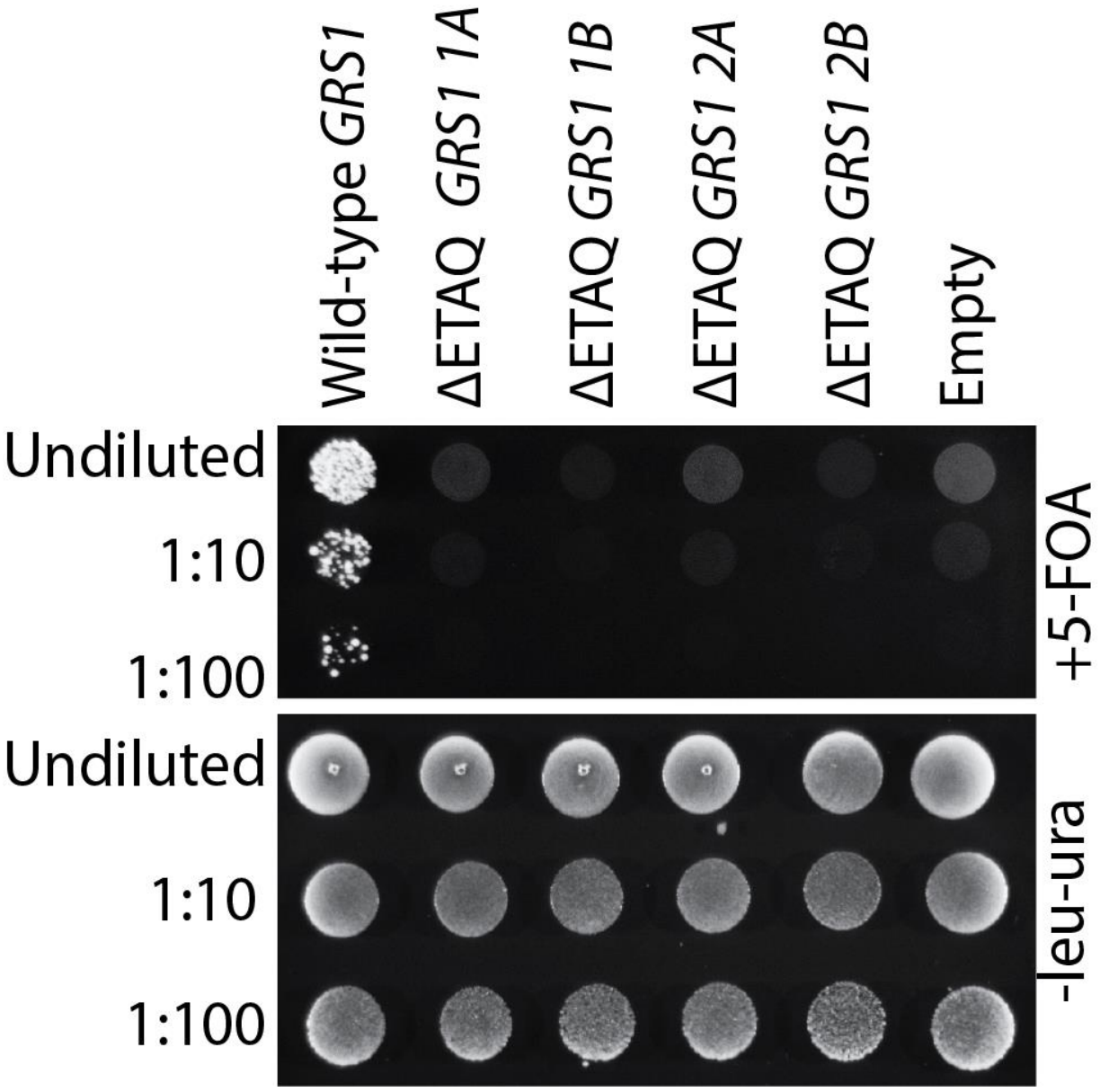


Figure 6.3. Yeast complementation studies of 245-248ΔETAQ *GRS1*. Yeast lacking endogenous *GRS1* were transformed with a *LEU2*-bearing vector containing wild-type or 245-248ΔETAQ GRS1 or no insert ('Empty'). Saturated cultures for each strain (labeled on top) were spotted undiluted or diluted (1:10 and 1:100) on media contain 5-FOA or lacking leucine and uracil, and grown for 72 hours at 30°C to determine if the 245-248ΔETAQ *GRS1* sustained yeast viability. Two independent plasmid preparations of 245-248ΔETAQ *GRS1* (1 and 2) were transformed into yeast and two colonies from each transformation were assessed (A and B).

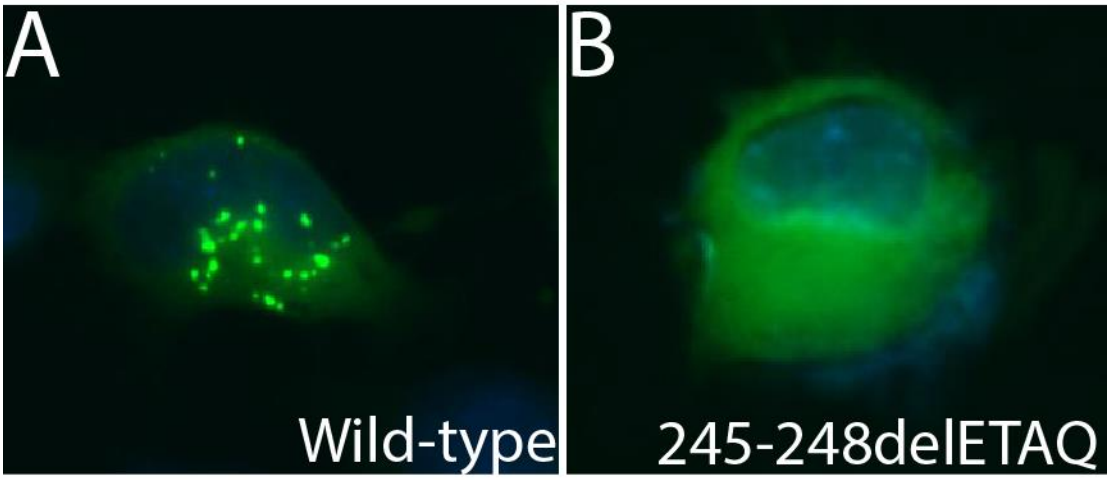


Figure 6.4. Localization studies of 245-248ΔETAQ *GARS*. MN-1 cells were transfected with expression constructs to express wild-type GARS (A) and 245-248ΔETAQ GARS (B) in frame with EGFP. EGFP positive cells were evaluated for presence of GARS puncta.

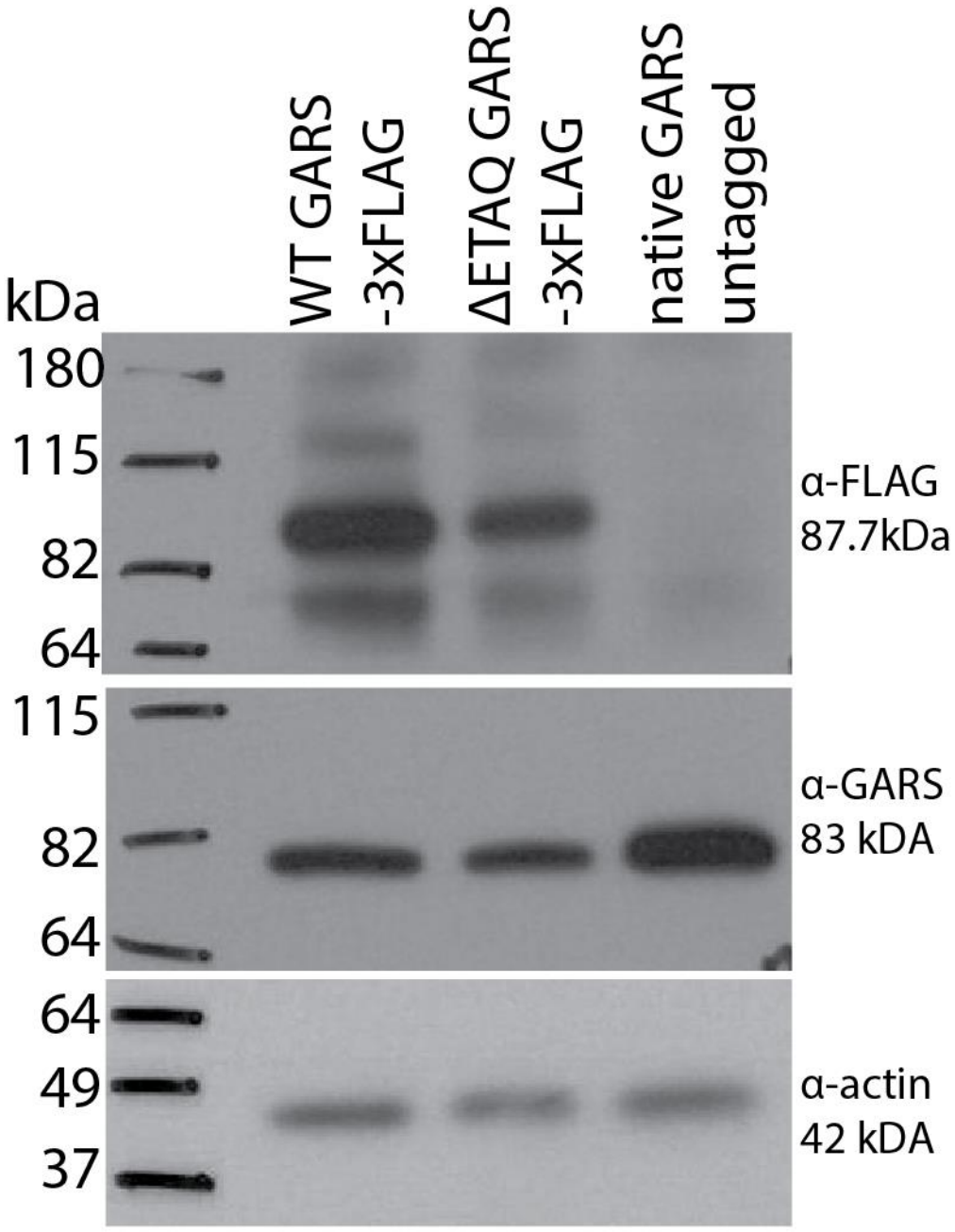


Figure 6.5. Western blot of protein input for immunoprecipitations. Whole cell protein lysates from MN-1 cells transfected vectors to express wild-type or ETAQ GARS-3xFLAG or untagged native GARS (labeled on top) were separated on a denaturing protein gel and blotted with an anti-FLAG antibody to confirm the presence of wild-type and ETAQ GARS-3xFLAG proteins and the absence of a FLAG-tagged peptide in the native GARS untagged control. An antibody targeting the C-terminus of the GARS protein was used to detect expression of endogenous and untagged GARS. Note the increase in GARS levels in the untagged native GARS control. An antibody to actin was used as a loading control.

presence of endogenous GARS. We observed an increased signal with the GARS antibody in the native GARS sample. This indicates that the untagged native *GARS* pTM3xFLAG results in expression of an untagged GARS protein of comparable in size to endogenous GARS (~83 kDa). Since the GARS antibody targets the C-terminus of the protein, we hypothesize that the GARS epitope is blocked by the FLAG tag in the wild-type and 245-248ΔETAQ GARS-3xFLAG samples, explaining the absence of a band at the expected size for the FLAG-tagged GARS protein product.

Equal amounts of total protein for each condition were subjected to IP using anti-FLAG conjugated magnetic beads. Beads were washed and proteins were eluted via a competitive elution with FLAG peptide. The isolated proteins were separated on a denaturing protein gel and either immunoblotted with antibodies to FLAG to confirm successful immunoprecipitations of tagged proteins (Figure 6.6) or silver stained to visualize the protein banding patterns in each sample (Figure 6.7). We classified bands observed on silver stained gels into four categories: (1) non-specific bands present in all samples including the untagged native GARS control; (2) bands that correspond to the approximate sizes of tagged and endogenous GARS, which were present in the tagged GARS samples but absent in the untagged native GARS control sample; (3) bands indicating proteins that bind both wild-type and 245-248ΔETAQ-3xFLAG GARS that were present in both tagged samples, but absent in the untagged control; and (4) protein bands observed exclusively in the 245-248ΔETAQ-3xFLAG sample, which may indicate novel interacting partners of 245-248ΔETAQ GARS. Resulting bands were isolated in four gel slices (Figure 6.8) and subjected to mass spectrometry at the Fred Hutchinson Cancer Institute proteomics core in Seattle Washington. Alternatively, IP products were TCA precipitated

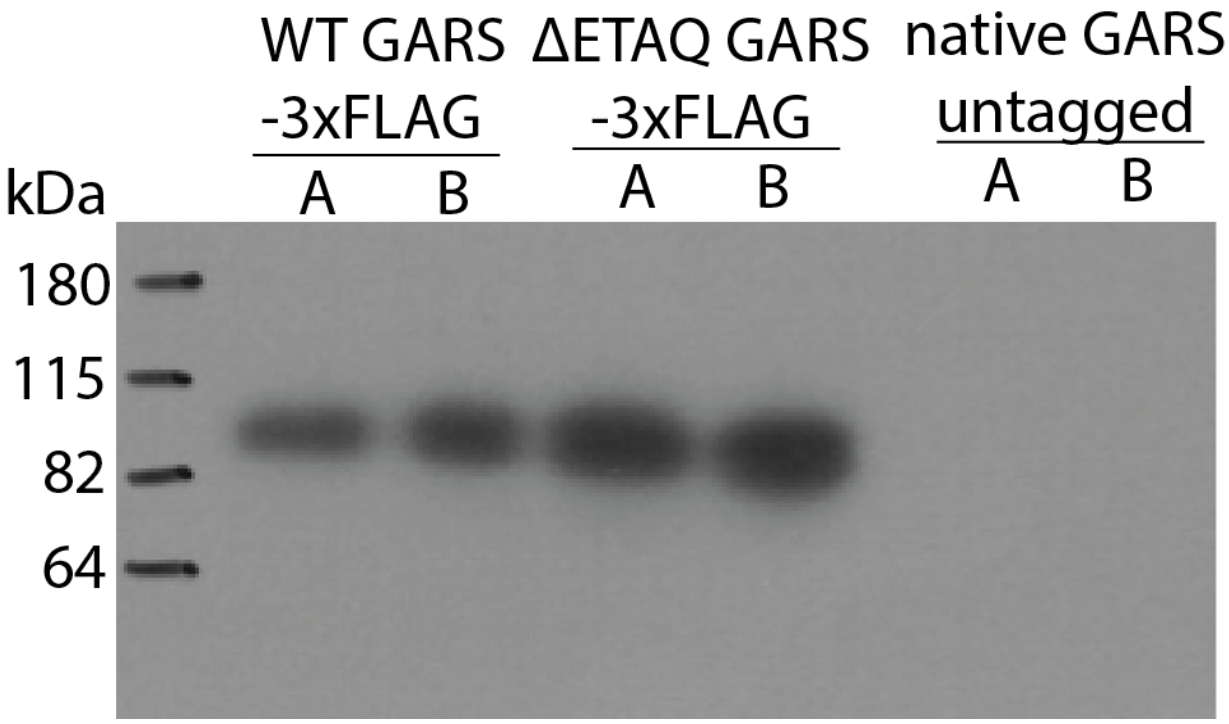


Figure 6.6. Western blot of immunoprecipitation products. Products from immunoprecipitations of protein extracts from MN-1 cells transfected with constructs expressing wild-type (WT) and 245-248ΔETAQ (ΔETAQ) GARS-3xFLAG or untagged native GARS were run on a denaturing gel and immunoblotted with an anti-FLAG antibody to confirm successful immunoprecipitation of the FLAG-tagged peptides.

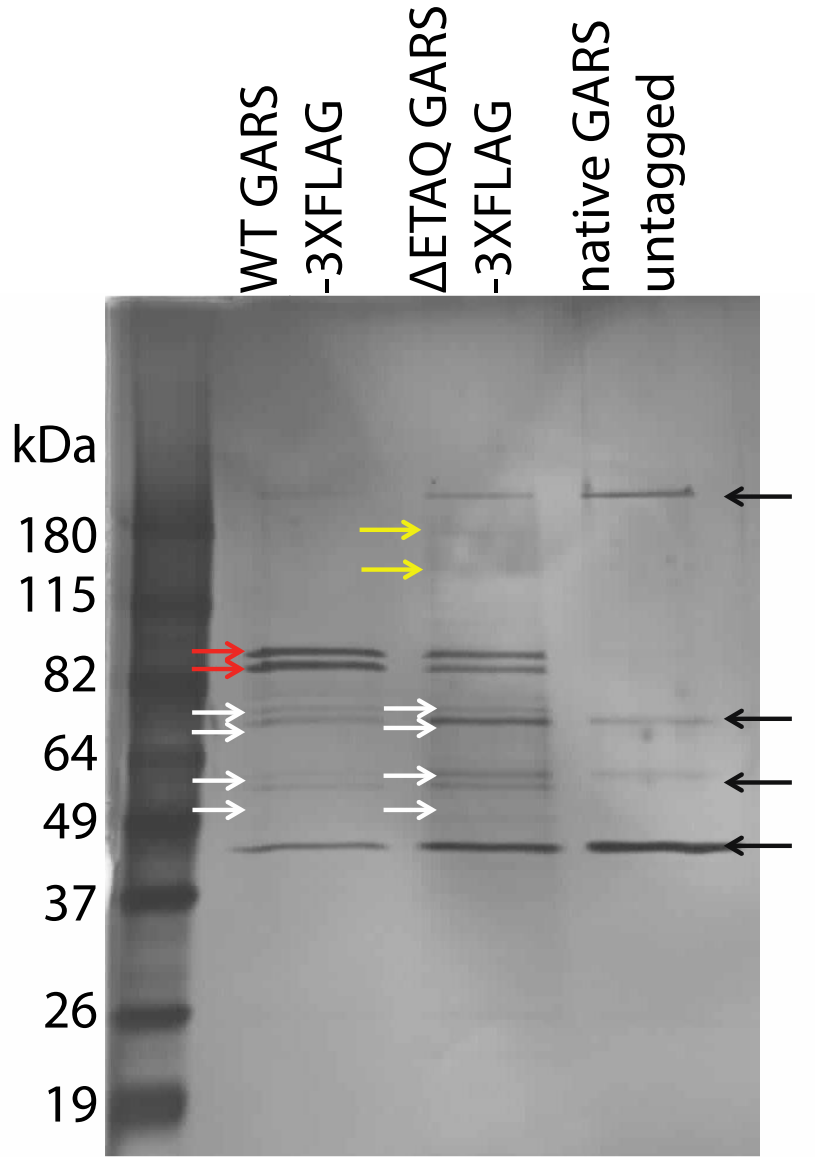


Figure 6.7. Silver stained gel of immunoprecipitation products. Immunoprecipitation reactions using an anti-FLAG antibody were performed on protein lysates isolated from MN-1 cells expressing wild-type (WT) or 245-248ΔETAQ (ΔETAQ) GARS-3xFLAG or native GARS (untagged). Resulting products were separated on a denaturing gel and silver stained for visual inspection of protein bands. Four types of bands were appreciated. Black arrows indicate bands present in all lanes. Red arrows indicate bands corresponding to the approximate sizes of tagged and endogenous GARS. Note the absence of these bands in the native GARS lane. White arrows indicate common interacting proteins between both wild-type and 245-248ΔETAQ-3xFLAG samples. Yellow arrows mark protein bands observed exclusively in the 245-248ΔETAQ-3xFLAG sample.

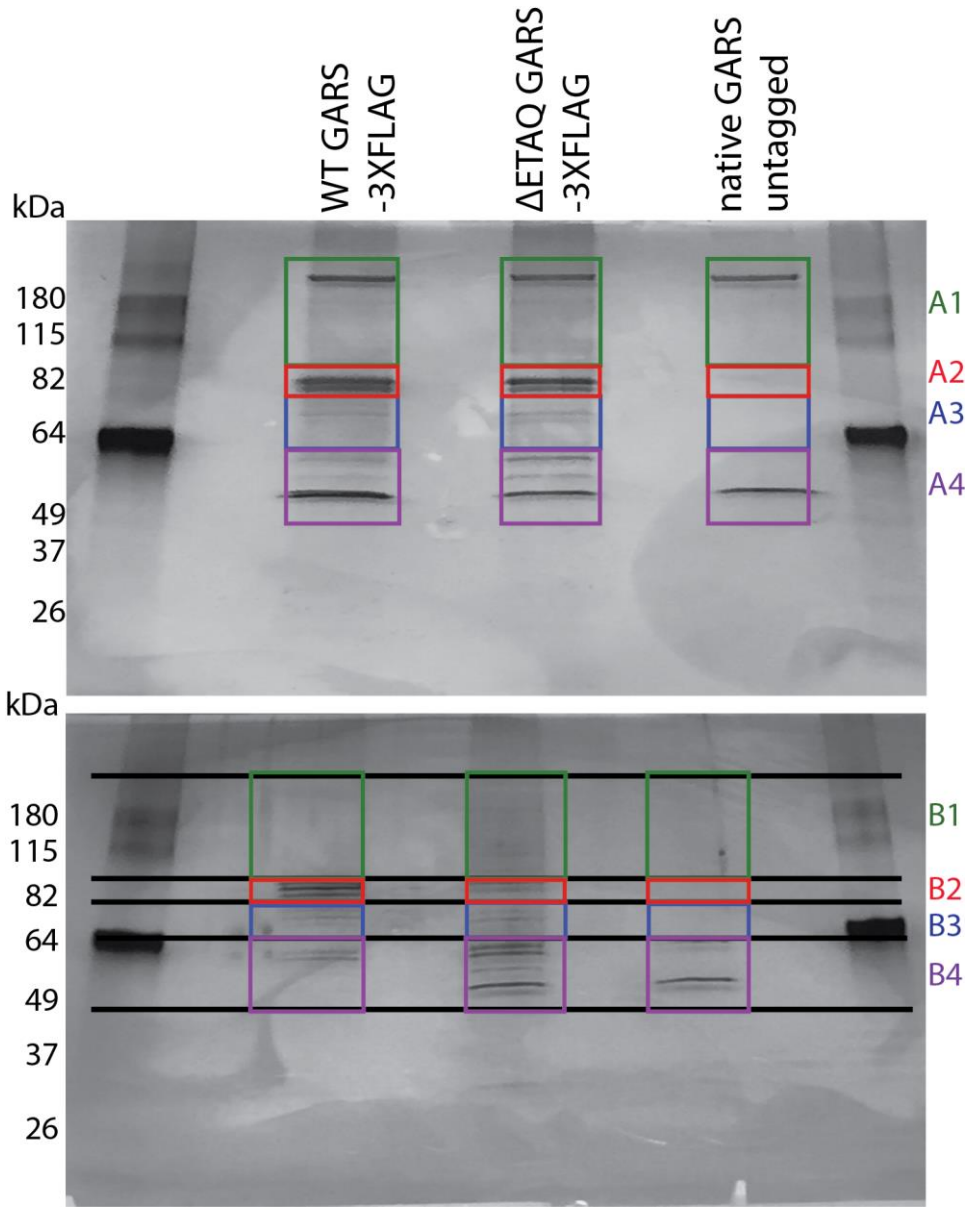


Figure 6.8. Images of silver stained gel samples used for mass spectrometry. Immunoprecipitations products from MN-1 cells expressing with wild-type (WT) or 245-248ΔETAQ (ΔETAQ) GARS-3xFLAG or untagged native GARS from two independent experiments (A and B) were run on a denaturing protein gel and silver stained to visualize bands. Gel lanes were each divided into four gel slices for mass spectrometry analysis as indicated (1-4). Gel slices were sent to the Fred Hutchinson Cancer Institute for mass spectrometry analysis.

immediately after IP and TCA samples were sent to University of Michigan Proteomics Resource Facility for mass spectrometry analysis.

Resulting mass spectrometry data was assessed to determine both the identity of GARS-interacting proteins, as well as to quantitate the abundance of each protein in the sample. For the Fred Hutchinson data set, the resulting quantifications of peptides between the wild-type and 245-248ΔETAQ samples were compared to calculate the relative fold change for each putative interacting partner identified. Importantly, the top two proteins identified were human GARS (our bait protein) and mouse GARS, which is endogenous to the MN-1 cell line used in these studies (Salazar-Grueso et al. 1991). To date, the best characterized interacting partner of GARS is GARS itself, as the protein must homodimerize to form a functional enzyme capable of charging tRNA (Freist et al. 1996). As would be expected based on the estimated protein sizes of 87 kDa and 83 kDa for human GARS-3xFLAG and mouse GARS, respectively, these proteins were present in high abundance in slices A2 and B2 (Figure 6.8) of both the wild-type and 245-248ΔETAQ GARS-3xFLAG samples (Table 6.2). Interestingly, the wild-type GARS sample had a 1.3 fold higher interaction with the mouse GARS compared to the 245-248ΔETAQ GARS-3xFLAG samples. It is important to note that this analysis wasdone using total spectral counts, not unique spectral counts for these two proteins so peptides are double counted.

Data analysis of the Fred Hutchinson data set revealed that both wild-type and mutant GARS bind to many of the same proteins. One example is heat shock protein 90 (HSP90). HSP90 was the second most abundant protein present in both wild-type and 245-248ΔETAQ GARS-3xFLAG samples. It was identified in slices A2 and B2 for both samples (Figure 6.8) as would be predicted by its molecular weight of 90 kDa. HSP90 is a molecular chaperone responsible for

Table 6.2. Spectral counts for GARS in immunoprecipitation samples

Protein	Species	WT Spectral Counts*	ΔETAQ Spectral Counts*	Fold Change
GARS	Human	575	438	1.3
GARS	Mouse	282	204	1.3

*Counts represent total spectral counts, not unique spectral counts. Shared peptides are double counted in this quantification.

assisting in proper protein folding and stabilization of other proteins during heat stress, and is involved in the degradation pathway of proteins (Csermely et al. 1998).

Since our hypothesis for the disruption of GARS puncta was due to differential protein-protein interactions between wild-type and mutant GARS, we calculated the fold change differences for protein interactions in wild-type and 245-248ΔETAQ GARS-3xFLAG samples in the Fred Hutchinson data set and report here the top 10 fold changes in each direction (Table 6.3 and Table 6.4). The top two proteins with increased binding to wild-type GARS compared 245-248ΔETAQ GARS were PLEC (3.29 fold) and ACTN4 (2.67 fold) (Table 6.3). PLEC interlinks intermediate filaments with microtubules and microfilaments (Svitkina et al. 1996), while ACTN4 is an F-actin cross-linking protein that anchors actin intracellular structures (Rual et al. 2005), indicating the interaction with the cytoskeleton may be critical for GARS puncta formation. Conversely, UBA1, the protein that catalyzes the first step in ubiquitin conjugation to mark cellular proteins for degradation (Kudo et al. 1991), demonstrated a 4.8 fold increase in binding to 245-248ΔETAQ compared to wild-type GARS (Table 6.4). Ubiquitination is used to mark misfolded proteins for degradation. This may indicate that 245-248ΔETAQ GARS is misfolded, and may explain the inability of 245-248ΔETAQ GARS to form puncta in cells.

Analysis of TCA precipitated IP products by the University of Michigan Proteomics Resource Facility is still ongoing. However, preliminary results from a single set of samples shows an abundance of GARS protein in the samples. In addition, wild-type GARS and 245-248ΔETAQ GARS were both found to interact with HSPa8, chaperone protein in the heat shock 70 family that has been implicated in neurodegenerative disease (Mayer and Bukau 2005). The remaining samples are currently being processed and will used for comparison with the data set generated at the Fred Hutchinson Institute.

Table 6.3. Protein demonstrating increased binding to wild-type GARS

Protein	Function	Max Fold Change	WT Spectral Counts	ΔETAQ Spectral Counts
PLEC	Interlinks intermediate filaments with microtubules and microfilaments	3.29	19	3
ACTN4	F-actin cross-linking protein anchors actin intracellular structures	2.67	28	8
COR1C	May be involved in cytokinesis, motility, and signal transduction	2.6	35	11
MYPT1	Key regulator of protein phosphatase 1C (PPP1C)	2.6	9	1
SRC8	Contributes to organization of actin cytoskeleton and cell shape	2.47	75	28
NFL	Neurofilament involved in the maintenance of neuronal caliber	2.2	7	1
FLII	May co-activate in transcriptional activation by hormone-activated nuclear receptors (NR)	2.17	9	2
WDR1	Induces disassembly of actin filaments	2.7	17	6
6HBB1	Involved in oxygen transport from the lung to peripheral tissues	2	4	0
CYTSA	Involved in cytokinesis and spindle organization	1.8	5	1

Table 6.4. Protein demonstrating increased binding to 245-248ΔETAQ GARS

Protein	Function	Max Fold Change	WT Spectral Counts	ΔETAQ Spectral Counts
UBA1	Catalyzes the first step in ubiquitin conjugation to mark cellular proteins for degradation	4.8	1	20
RUVB2	Component of the NuA4 histone acetyltransferase complex involved in transcriptional activation	4.4	1	18
VARS	Catalyzes the attachment of valine to $tRNA^{Val}$	4	0	12
RPL4	60S ribosomal protein L4	3.75	0	11
VAT1	Involved in organization of actin cytoskeleton	3.5	0	10
STIP1	Mediates the association of the molecular chaperones HSC70 and HSP90	3.4	1	13
ENPL	Molecular chaperone that functions in the processing and transport of secreted proteins	3.33	5	26
PLIN3	Required for the transport of mannose 6-phosphate receptors (MPR) from endosomes to the trans-Golgi network	3	0	8
GOGA5	Involved in maintaining Golgi structure	3	0	8
SMAD2	Receptor-regulated SMAD	3	0	8

DISCUSSION

In this chapter, we have implicated a *de novo GARS* mutation in the pathogenesis of infantile SMA and have begun to understand GARS protein-protein interactions in neuronal cells. 245-248ΔETAQ *GARS* deletes conserved residues critical for glycine binding and severely affects all aspects of GARS activity tested. Specifically, 245-248ΔETAQ *GARS* dramatically decreases GARS activity to <0.001% that of wild-type GARS in *in vitro* biochemical assays. In yeast, 245-248ΔETAQ *GARS* fails to support cell viability, indicating this variant impairs activity *in vivo*. Finally, this mutation disrupts GARS puncta formation in MN-1 cells. Since 245-248ΔETAQ *GARS* results in a severe phenotype and displays loss-of-function characteristics in all assays commonly used to assess *GARS* mutations (Antonellis et al. 2006; Nangle et al. 2007; Stum et al. 2011), we selected this mutation for IP/mass spectrometry studies. Although still on-going, these IP/mass spectrometry studies have the potential to offer novel insights into both GARS biology and mechanisms of *GARS*-mediated disease.

The presence of GARS puncta was first appreciated over 10 years ago (Antonellis et al. 2006); however, no scientific advances have been published as to the structure, function, or role in disease of GARS puncta. Here we compare the protein-protein interactions of wild-type and 245-248ΔETAQ *GARS* to begin to address these outstanding questions. Our preliminary mass spectrometry data raises some interesting hypotheses. For example, wild-type GARS binds PLEC and ACTN4, two proteins involved in cytoskeleton structure, with greater than 3-fold affinity compared to 245-248ΔETAQ *GARS*. This may indicate that interactions with the cytoskeleton are critical for GARS puncta structure or localization. The puncta may be anchored in specific locations in the cell to facilitate proper activity of GARS or the cytoskeleton itself may be a necessary structural component of GARS puncta. There are examples of cytoskeleton-

dependent granule formation. For example, stress granules cannot form in the presence of nocodazole or vincristine, two pharmacologic agents that destabilize microtubules (Ivanov et al. 2003). In contrast, the microtubule-stabilizing drug paclitaxel and actin filament-disrupting drug latrunculin B enhance stress granule formation. If the cytoskeleton is necessary for the stabilization of GARS puncta, decreased interactions with PLEC and ACTN4 by 245-248ΔETAQ *GARS* may explain why this mutation fails to form puncta. If this interaction is confirmed via co-IP, careful follow-up studies should evaluate drugs that act on microtubules and actin filaments for an effect on wild-type and mutant puncta formation. If pharmacologic stabilization of the cytoskeleton enhances GARS puncta formation, as it does for stress granule formation, this may indicate a therapeutic avenue worth further investigation for GARS-mediated CMT disease.

In contrast to the proteins described above, multiple proteins showed a higher level of interaction with 245-248ΔETAQ GARS than wild-type GARS. Specifically, 245-248ΔETAQ GARS demonstrated a 4-fold higher interaction with UBA1, an E1 ubiquitin ligase that initiates the E1-E2-E3 enzymatic cascade of the Ubiquitin-Proteasome System (UPS) that marks proteins for degradation (Schulman and Harper 2009). One potential explanation for this increased interaction is that the deletion of the ETAQ residues results in improper folding of 245-248ΔETAQ GARS, leading to its ubiquitination by UBA1. Degradation of the mutant GARS gene product may be detrimental to the cell by leading to decreased GARS levels and/or the interaction of mutant GARS and UBA1 itself may be toxic to the cell. UPS and UBA1 activity has been shown to be critical for neuron and synaptic development and health (Jiang et al. 2010; Deglincerti et al. 2015). Mutations in *UBA1* are associated with Infantile-onset X-linked spinal muscular atrophy (SMAX; XL-SMA) a rare form of SMA that has both motor and sensory

involvement and cerebellar abnormalities (Ramser et al. 2008; Dlamini et al. 2013). Although it is not completely clear how *UBA1* mutations lead to disease, it is known that all five *UBA1* mutations cluster in a domain responsible for interacting with gigoxonin, a protein that has been implicated in giant axonal neuropathy (GAN1) (Bomont et al. 2000). The gigaxonin-UBA1 complex is responsible for regulating MAP1B levels, a microtubule-associated protein that is critical for neuronal development and has been implicated in neurodegenerative diseases (Bomont et al. 2000; Gordon-Weeks and Fischer 2000). Loss-of-function mutations in *UBA1* result in axon pruning defects in drosophila (Watts et al. 2003) and loss of *UBA1* in adult *C. elegans* leads to late on-set paralysis (Kulkarni and Smith 2008). In addition, reduced UBA1 levels in Schwann cells result in myelin defects in mice (Aghamaleky Sarvestany et al. 2014; Hunter et al. 2014). UBA1 dysfunction has also been linked to multiple neurodegenerative diseases, including other forms of SMA and Huntington's disease (reviewed in (Groen and Gillingwater 2015)). If UBA1 is shown to be important in CMT disease pathogenesis, this would indicate a common pathway is involved in many forms of neurodegeneration.

Both wild-type and mutant GARS appear to interact with members of the heat shock protein pathway, specifically HSP90 and HSPa8, also known as heat shock cognate protein 70 (HSC70). Recent studies in Niemann-Pick type C have demonstrated that HSP90 and HSP70 have a role in stabilizing NPC1 (the protein mutated in Niemann-Pick Type C), while HSPa8/HSC70 destabilizes the NPC1 leading to its degradation (Nakasone et al. 2014). Thus a fine balance of interactions between GARS and these proteins may play a critical role in peripheral nerve health and maintenance, and alterations to these interactions may be involved in the pathogenesis of ARS-mediated CMT disease.

While preliminary data demonstrates altered protein binding between wild-type and 245-248ΔETAQ GARS, follow-up studies are needed to confirm these interactions. Additional mass spectrometry experiments are underway to increase the power of our study. Co-IPs can be used to confirm putative interactions, and immunofluorescence studies will be necessary to determine if interactions are occurring in the puncta themselves or elsewhere in the cell. Once interactions are confirmed, studies of other disease-associated *GARS* mutations are needed to determine if the abnormal interactions of 245-248ΔETAQ GARS are universal to all non-puncta forming mutations and if puncta-forming mutations retain wild-type GARS binding partners.

In conclusion, the data in this chapter not only expand the allelic spectrum of *GARS* mutations by implicating 245-248ΔETAQ *GARS* in disease, but also provide an excellent starting point for ongoing studies into the role of GARS puncta and the mechanism of *GARS*-mediated CMT disease.

CHAPTER 7

Summary and Future Directions

SUMMARY OF DISSERTATION FINDINGS

The data presented in this thesis have contributed to expanding the allelic heterogeneity of ARS mutations in human disease and have added to the understanding of the molecular pathology of ARS-mediated CMT disease. Prior to this thesis work, 21 ARS loci had strong genetic evidence for their role in human disease, including 18 ARS genes in recessive syndromes and three in dominantly inherited CMT disease. We present genetic and functional evidence of pathogenicity for alleles identified at five ARS loci in patients with recessive disease and at four loci for dominantly inherited peripheral neuropathy (Table 7.1).

Recessive mutations in ARS genes have been implicated in a broad spectrum of disease phenotypes that range from multi-organ to tissue specific. The genetic studies presented here mark the first implication of *AARS, YARS, GARS,* and *FARSB* in recessive disease. Recessive mutations in *AARS, YARS, GARS,* and *DARS* lead to a severe central nervous system phenotype and is consistent with previous reports that the majority of ARS mutations cause phenotypes that include a neurologic component. *FARSB,* in contrast, leads to lung and liver disease similar to what has been observed for recessive disease caused by *MARS* mutations. The reason for the tissue specificity of ARS-mediated disease remains unknown. Although characterization is not complete for all mutations, functional studies indicate that impaired function is likely the

Table 7.1. Human diseases associated with ARS mutations

Gene	Locus	Location of Function	Mode of Inheritance	Disease	OMIM no.	Clinical Description	Reference
GARS	7p15	Cytoplasm & Mitochondria	Autosomal dominant	CMT2D or dSMA-V	601472 600794	Upper limb predominant, axonal motor neuropathy with variable sensory involvement	(Antonellis et al. 2006)
GARS	7p15	Cytoplasm & Mitochondria	Autosomal recessive	Unknown	NA	Growth retardation *in utero*, decreased body weight, short stature, atrial septal defect, abnormal facial features, failure to thrive	This study
YARS	1p35.1	Cytoplasm	Autosomal dominant	DI-CMT	608323	Motor and sensory neuropathy with intermediate nerve conduction velocities	(Jordanova et al. 2006)
YARS	1p35.1	Cytoplasm	Autosomal dominant	Unknown	NA	Microcephaly, infantile lethality,	This study
AARS	16q22	Cytoplasm	Autosomal dominant	CMT2N	613287	Motor and sensory neuropathy with variable sensorineural deafness	(Latour et al. 2010)
AARS	16q22	Cytoplasm	Autosomal recessive	EEIE29	616339	Early-onset epileptic encephalopathy, failure to thrive, vertical tali, spasticity, microcephaly	This study
HARS	5q31.3	Cytoplasm	Autosomal dominant	CMT2W	142810	Single individual with sensory > motor axonal neuropathy	(Vester et al. 2012), This study
MARS	12q13.3	Cytoplasm	Autosomal dominant	CMT2U	616280	Motor and sensory neuropathy	(Gonzalez et al. 2013)
MARS	12q13.3	Cytoplasm	Autosomal recessive	ILLD	615486	Interstitial lung and liver disease	(Hadchouel et al. 2015)
KARS	16q23.1	Cytoplasm & Mitochondria	Autosomal recessive	RI-CMTB	613641	Single individual with intermediate CMT, developmental delay, self-abusive behavior, dysmorphic features and vestibular Schwannoma	(McLaughlin et al. 2010)
DARS	2q21.3	Cytoplasm	Autosomal recessive	HBSL	615281	Hypomyelination with brainstem and spinal cord involvement and leg spasticity	(Taft et al. 2013)
FARSB	2q36.1	Cytoplasm	Autosomal recessive	Unknown	NA	Interstitial lung and liver disease	This study

Gene	Locus	Location of Function	Mode of Inheritance	Disease	OMIM no.	Clinical Description	Reference
QARS	3p21.31	Cytoplasmic & Mitochondria	Autosomal recessive	MSCCA	615760	Microcephaly, progressive, seizures, and cerebral and cerebellar atrophy	(Zhang et al. 2014)
RARS	5q34	Cytoplasmic	Autosomal recessive	Leuko-dystrophy	616140	Leukodystrophy, hypomyelination	(Wolf et al. 2014)
AARS2	6p21.1	Mitochondria	Autosomal recessive	Mitochondrial Infantile CMP	614096	Mitochondrial infantile cardiomyopathy	(Götz et al. 2011)
DARS2	1q25.1	Mitochondria	Autosomal recessive	LBSL	611105	Leukoencephalopathy with brainstem and spinal cord involvement and lactate elevation	(Scheper et al. 2007a)
EARS2	16p12.2	Mitochondria	Autosomal recessive	LTBL	614924	Infantile leukoencephalopathy involving thalamus and brainstem, high lactate, myopathy, hypotonia, seizures, retinitis pigmentosa, respiratory failure	(Steenweg et al. 2012)
FARS2	6p25.1	Mitochondria	Autosomal recessive	Alpers encephalo-pathy	614946	Neonatal encephalopathy, developmental delay, refractory seizures, lactic acidosis, liver disease	(Elo et al. 2012)
HARS2	5q31.3	Mitochondria	Autosomal recessive	Perrault Syndrome 2	614926	Sensorineural deafness (both sexes), ovarian dysgenesis (females)	(Pierce et al. 2011)
IARS2	1q41	Mitochondria	Autosomal recessive	CAGSSS; Leigh Syndrome	616007	Cataracts, growth hormone deficiency with short stature, partial sensorineural deafness, peripheral neuropathy	(Schwartzentruber et al. 2014)
LARS2	3p21.31	Mitochondria	Autosomal recessive	Perrault Syndrome 4	615300	Premature ovarian failure (POF) in females and by progressive hearing loss in both females and males	(Pierce et al. 2013)
MARS2	2q33.1	Mitochondria	Autosomal recessive	Spastic Ataxia 3	611390	Spastic ataxia, leukoencephalopathy	(Bayat et al. 2012)
RARS2	6q16.1	Mitochondria	Autosomal Recessive	Spastic Ataxia 3	611390	Spastic ataxia, leukoencephalopathy	(Edvardson et al. 2007)
SARS2	19q13.2	Mitochondria	Autosomal Recessive	HUPRA Syndrome	613845	Hyperuricemia, pulmonary hypertension, infantile renal failure, and alkalosis syndrome	(Belostotsky et al. 2011)

Gene	Locus	Location of Function	Mode of Inheritance	Disease	OMIM no.	Clinical Description	Reference
YARS2	12p11.21	Mitochondria	Autosomal Recessive	MLASA2	613561	Myopathy, lactic acidosis, and sideroblastic anemia, muscle weakness, dysphagia, exercise intolerance	(Riley et al. 2010)
TARS2	1q21.2	Mitochondria	Autosomal Recessive	COXPD21	615918	Axial hypotonia, limb hypertonia, psychomotor delay, and increased serum lactate, thin corpus callosum, cerebral spongiosis, hepatic stenosis	(Diodato et al. 2014)
VARS2	6p21.33	Mitochondria	Autosomal Recessive	COXPD20	615917	Ophthalmoplegia, ptosis, ataxia, seizure	(Diodato et al. 2014; Taylor et al. 2014)

pathomechanism of recessive ARS-mediated disease and suggest that screening of ARS mutations in patients with similar phenotypes should be performed to help build the genetic argument for these loci in recessive disease.

In addition to implicating ARS genes in novel phenotypes, we have identified and characterized alleles in *GARS, YARS, AARS,* and *HARS* in individuals with dominant peripheral neuropathy (Table 7.2). Evaluations of these new alleles led to three significant contributions in the field of ARS-mediate CMT research. First, the identification of multiple *HARS* alleles was critical to implicating *HARS* as a CMT disease locus. Second, we were able to expand the phenotypic spectrum observed for *AARS*, *GARS,* and *HARS* disease. Lastly, combined with functional studies of previously implicated *GARS* mutations, we determined that the majority of CMT-associated ARS alleles demonstrate loss-of-function characteristics, indicating that loss-of-function is an important component of ARS-mediated disease.

Our studies offer valuable insights into the functional consequences of ARS mutations and the pathomechanism of ARS-mediated disease. While loss-of-function is generally accepted as the mechanism in recessive ARS phenotypes, the pathomechanism for dominant ARS-mediated CMT disease is widely debated. Our studies revealed that *AARS* and *YARS* mutations identified in both dominant and recessive disease led to a dramatic reduction in ARS activity. This suggests a direct correlation between levels of ARS activity and the severity of disease (*i.e.* the greater the reduction in enzyme function, the more severe the phenotype). Direct evidence supporting this correlation is observed for the P167T *YARS* allele. P167T *YARS* is a loss of function allele that when in the heterozygous state leads to peripheral neuropathy, while in the homozygous state, leads to recessive microcephaly (Table 7.1). Combined with the studies

Table 7.2. ARS variants identified in dominantly inherited CMT disease

Gene	Variant	Conservation	Enzyme Activity	Yeast Viability	Fly Comp.	Puncta Formation	Dimerization	References
GARS	A57V	Mammals	Normal[1]	NA	NP	Yes	NP	(Rohkamm et al. 2007)
GARS	E71G	Yeast	Reduced	Viable	Reduced	Yes	NP	(Antonellis et al. 2006; Nangle et al. 2007)
GARS	L129P	Bacteria	Reduced	Reduced	No	No	Reduced	(Antonellis et al. 2006; Nangle et al. 2007)
GARS	D146N	Yeast	Reduced	Reduced	NP	NP	NP	(Lee et al. 2012a)
GARS	S211F	Worm	Reduced	NA	NP	NP	NP	(Lee et al. 2012a)
GARS	G240R	Bacteria	Reduced	Viable	NP	No	Reduced	(Antonellis et al. 2006; Nangle et al. 2007)
GARS	P244L	Yeast	Reduced	Lethal	NP	NP	NP	(Abe and Hayasaka 2009)
GARS	245-248 ΔETAQ	Bacteria	Reduced	Lethal	NP	No	NP	This study
GARS	I280F	Yeast	Reduced	Viable	NP	NP	NP	(James et al. 2006)

Gene	Variant	Conservation	Enzyme Activity	Yeast Viability	Fly Comp.	Puncta Formation	Dimerization	References
GARS	H418R	Yeast	Reduced	Lethal	NP	No	Reduced	(Antonellis et al. 2006; Nangle et al. 2007)
GARS	D500N	Zebrafish	Normal	NA	NP	NP	Normal	(Del Bo et al. 2006; Nangle et al. 2007; Stum et al. 2011)
GARS	G526R	Bacteria	Reduced	Lethal	NP	Yes	Normal	(Antonellis et al. 2006; Nangle et al. 2007; Xie et al. 2007)
GARS	S581L	Zebrafish	Normal	NA	NP	NP	Normal	(James et al. 2006; Cader et al. 2007; Nangle et al. 2007)
GARS	G598A	Worm	Reduced	Viable	NP	No	NP	(James et al. 2006; Stum et al. 2011)
YARS	G41R	Bacteria	Reduced	Lethal	NP	No	NP	(Jordanova et al. 2006)
YARS	E196K	Yeast	Reduced	Reduced	NP	No	NP	(Jordanova et al. 2006)
YARS	E196Q	Yeast	Reduced	Reduced	NP	No	NP	This study
YARS	153-156Δ VKQV	Yeast	NP	NP	NP	NP	NP	(Jordanova et al. 2006)
AARS	N71Y	Yeast	Reduced	Lethal	NP	NA	NP	(Lin et al. 2011; McLaughlin et al. 2011)

Gene	Variant	Conservation	Enzyme Activity	Yeast Viability	Fly Comp.	Puncta Formation	Dimerization	References
AARS	G102R	Bacteria	NP	Lethal	NP	NA	NP	This study
AARS	R329H	Zebrafish	Reduced	Lethal	NP	NA	NP	(Latour et al. 2010; McLaughlin et al. 2011)
AARS	D893N	Mouse	NP	NP	NP	NA	NP	(Zhao et al. 2012)
HARS	R137Q	Yeast	NP	Lethal	NP	NA	NP	(Vester et al. 2012)
HARS	T132I	Yeast	NP	Lethal	NP	NA	NP	This study
HARS	P134H	Yeast	NP	Lethal	NP	NA	NP	This study
HARS	V155G	Yeast	NP	Lethal	NP	NA	NP	This study
HARS	D175E	Yeast	NP	Lethal	NP	NA	NP	This study
HARS	Y330C	Bacteria	NP	Reduced	NP	NA	NP	This study
HARS	S356N	Yeast	NP	Reduced	NP	NA	NP	This study

Gene	Variant	Conservation	Enzyme Activity	Yeast Viability	Fly Comp.	Puncta Formation	Dimerization	References
HARS	D364Y	Bacteria	NP	Lethal	NP	NA	NP	This study
MARS	R618C	Bacteria	NP	Lethal	NP	NA	NP	(Gonzalez et al. 2013)

Red text indicates a result consistent with a loss-of-function effect.
[1]This variant was associated with ~50% activity compared with wild-type GARS.
We considered a >90% reduction in activity as dramatically 'Reduced'.
NA – not applicable; mutation could not be modeled in yeast ortholog.
NP- not performed; analysis not included in the literature.
Fly Comp- Fly complementation assays

demonstrating loss-of-function characteristics for dominant ARS mutations, these data support a loss-of-function mechanism for ARS-mediated CMT disease.

We explored two loss-of-function mechanisms for *GARS*-mediated CMT disease: a dominant negative effect and protein mislocalization. We developed a *trans*-complementation assay to assess for a dominant negative effect of *GARS* mutations *in vivo*. Co-expression of a hypomorphic *GARS* allele (L129P) and a null allele (H418R or G526R) revealed that motifs on opposing subunits could complement one and other to improve overall dimer function. This appears to refute the dominant negative hypothesis and instead suggests that a wild-type GARS subunit can compensate for mutations on the opposing subunit. However, not all motifs complement in *trans* as evidenced by a lack of growth observed for yeast co-expressing H418 and G526R alleles. Thus, further studies on the human wild-type/mutant heterodimer are necessary to fully understand the effects of mutations on overall enzyme activity in heterozygous cells. To address the role of GARS puncta in CMT disease, we performed IP/mass spec studies on wild-type and 245-248ΔETAQ GARS and identified multiple putative binding partners for both proteins. These experiments will provide the basis for future studies aimed at understanding the role of GARS puncta in neurons and how disruption of these structures leads to disease.

OUTSTANDING QUESTIONS AND FUTURE DIRECTIONS

Defining the complete spectrum of ARS-mediated disease

While four cytoplasmic ARS loci have been implicated in peripheral neuropathy, it is still unknown if all cytoplasmic ARS genes can lead to dominant disease. To address this, I propose a two-pronged observational and experimental approach. First, individuals who are heterozygous for ARS variants that are implicated in recessive disease should be periodically evaluated for evidence of peripheral neuropathy. Since most ARS-associated CMT cases are late onset and

many of the parents of the affected children are young individuals themselves, longitudinal observation is necessary. To date, only carriers of the P167T *YARS* allele have reported symptoms of a mild neuropathy (discussed above). Currently, evidence indicates that cytoplasmic, not mitochondrial, ARS dysfunction leads to CMT disease (Wallen and Antonellis 2013). Thus of the 21 ARS enzymes implicated in recessive syndromes, only heterozygous individuals in families with mutations in cytoplasmic ARS genes (*AARS, DARS, GARS, FARSB, KARS, MARS, QARS, RARS,* and *YARS* (McLaughlin et al. 2010; Taft et al. 2013; Wolf et al. 2014; Zhang et al. 2014; Hadchouel et al. 2015; Simons et al. 2015)) would be hypothesized to develop symptoms of peripheral neuropathy. Based on mouse studies showing that an expressed, stable mutant GARS protein is necessary to cause dominant disease (Seburn et al. 2006), we hypothesize that parents carrying missense ARS mutations, but not those carrying null (*i.e.* frame shift and nonsense) alleles, would be at risk for developing a peripheral neuropathy. Zebrafish studies on T209R *Gars*, a loss-of function mutant that fails to dimerize (briefly discussed in chapter 5), and the observation that only dimeric ARS enzymes have been implicated in CMT disease indicate that both impaired function and an ability to dimerize are both critical components of ARS-mediated CMT disease pathogenesis. QARS, MARS, and RARS enzymes all function as monomers and have been implicated in recessive disease (Wolf et al. 2014; Zhang et al. 2014; Hadchouel et al. 2015). Thus if missense mutations in these three genes lead to CMT disease in heterozygous individuals, this would refute the hypothesis that oligomerization is an important component of disease.

Since natural history studies are fraught with complications such as the length of the study, attrition rates of participants, and inability to control confounding variables, an experimental approach should be performed in parallel. Our lab is currently employing Programmed Allelic

Series (PALS) mutagenesis (Kitzman et al. 2015) to perform large scale mutagenesis on a subset of ARS loci to catalog all the variants that result in loss of function. In this technique, a single mutation is introduced per DNA template (*i.e.* the ARS open reading frame in yeast expression construct) in a high throughput manner such that every possible allele is generated for that DNA sequence. The pool of ARS alleles can be introduced into a ΔARS haploid yeast strain harboring a *URA3* maintenance vector and grown in either media lacking uracil and leucine or in complete media containing 5-FOA (Boeke et al. 1987). DNA is then extracted from yeast that survive in both conditions, and high throughput sequencing is used to identify the alleles that are present in each condition. Comparison of the two conditions will indicate which variants are able to sustain viability and which are lost during selection, allowing us to catalog all the alleles that are loss of function. Importantly, the functional consequences of variants of unknown significance identified in patients with CMT disease can be mined from the PALS data to rapidly determine the likelihood of pathogenicity for new alleles. Computational analysis and dimerization assays can be used to classify the loss-of function-mutations based on their ability to form dimers. Loss-of-function, dimer forming and loss-of-function, non-dimer forming mutations can expressed be in *C. elegans* to determine if dimerization is a critical component of ARS-mediated dominant neurotoxicity. Finally, a dominantly toxic variant can be modeled in mice to see if the mutation causes a peripheral neuropathy phenotype similar to the C201R or P234KY *Gars* mice. Studies similar to those outlined here are currently underway for threonyl-tRNA synthetase (*TARS*), a dimeric ARS enzyme not previously implicated in disease, to determine if *TARS* variants have the capacity to cause a peripheral neuropathy. To complement the *TARS* work, I propose performing PALS, yeast complementation, and *C. elegans* studies on a monomeric ARS enzyme that has not been previously implicated in disease (*e.g. VARS*) to determine if loss-of-function of

a monomeric ARS can lead to dominantly inherited peripheral neuropathy. These studies will provide valuable insight into the mechanism of ARS-mediated peripheral neuropathy and support the screening of additional ARS loci in idiopathic CMT disease.

Defining the pathomechanism of ARS-mediated dominant disease

As discussed in the introduction, there are five major hypotheses for the mechanism of ARS-mediated CMT disease. Of the five, a loss-of-function dominant negative effect or a toxic gain of function are widely considered the most probable mechanisms. Although this work contributes to the literature that loss of function is an important component of dominant ARS-mediated CMT disease (Wallen and Antonellis 2013; Griffin et al. 2014; Gonzaga-Jauregui et al. 2015; Motley et al. 2015; Safka Brozkova et al. 2015), there is some evidence suggesting a toxic gain of function mechanism may also be involved. In a pure loss-of-function mechanism, improving enzyme activity would be hypothesized to rescue the phenotype. However, two groups have shown that overexpression of wild-type *GARS* is not sufficient to rescue phenotypes observed in fly and mouse models of *GARS*-mediated disease, leading investigators to hypothesize that mutant ARS has a neurotoxic gain of function (Motley et al. 2011; Niehues et al. 2015). A recent study of the mouse P234KY *Gars* mutation demonstrated abnormal interaction of mutant GARS and the neuronal receptor protein neuropilin-1 that may be responsible for the neurotoxicity in this model (He et al. 2015). It is important to note that loss- and gain-of-function mechanisms may not be mutually exclusive. Loss of normal ARS function could free the protein to participate in a toxic interaction with a neuronal protein, or the combination of loss-of-tRNA charging and increased demands on cellular machinery (*e.g.* the ubiquitin protease system) in the presence of a non-functional ARS protein may together be too much for neurons to overcome, resulting in disease. Both of these hypotheses would explain the tissue specificity of ARS-mediated CMT

disease and the discrepancy in the literature between loss and gain of function data. Our studies investigating the binding partners of mutant and wild-type GARS described in Chapter 6 have the potential to identify both the lost and gained interactions of mutant GARS. Follow-up studies outlined in the following section can be used to understand the consequences of novel interactions.

The majority of research into the precise mechanism of ARS-mediated peripheral neuropathy has focused on *GARS* mutations. Given the similarities in phenotypes between *GARS*-, *HARS*-, *YARS*-, and *AARS*-mediated CMT disease, it is hypothesized that all ARS-mediated peripheral neuropathy occurs via a common mechanism. However, this hypothesis needs to be explored. To date, all the data for *YARS*-, *AARS*-, and *HARS*-mediated CMT disease indicate a loss-of-function mechanism. Specifically, all identified mutations at these loci lead to decreased ARS function and studies of yeast co-expressing mutant and wild-type *YARS* showed decreased growth consistent with a dominant negative effect (Jordanova et al. 2006; McLaughlin et al. 2011; Gonzaga-Jauregui et al. 2015; Motley et al. 2015; Safka Brozkova et al. 2015). Carefully assessing *AARS*, *HARS*, and the newly identified *YARS* mutations for a dominant negative effect is critical to addressing the mechanism of pathogenicity. While it is unknown if *HARS* complements *HTS1* in yeast, we can utilize humanized yeast strains to address a dominant negative effect for *AARS* and *YARS* mutations as both of these human genes have been shown to complement in yeast (Stephanie Oprescu, personal communication, (Jordanova et al. 2006; Kachroo et al. 2015)). Understanding the precise pathomechanism is critical to our understanding of the role of ARS in neuronal function and to providing important insights into therapeutic options for CMT disease.

The role of GARS puncta in disease

Mislocalization is an important loss-of-function mechanism that should be considered for *GARS*-mediated CMT disease as the majority of *GARS* alleles fail to form the punctate cellular structures observed for wild-type endogenous GARS. We hypothesize that GARS puncta are critical for axonal health and that disruption of these structures plays a role in pathogenicity. Currently, the function of GARS puncta, how mutations disrupt puncta formation, and what, if any, role loss of these structures play in disease pathogenesis remain open questions in the field. Although we have begun to address these questions using an IP/mass spectrometry approach to identify interacting partners of wild-type and mutant GARS, many follow up studies are needed. The top interacting candidates from the wild-type and 245-248ΔETAQ samples must be confirmed by co-IP and immunofluorescence. Lost or gained protein interactions of mutant GARS may increase or decrease activity of its binding partner, disrupting normal activity, and leading to dominant toxicity in cells. To determine the significance of these interactions in disease, interacting partners could be manipulated in zebrafish. Unpublished work in our lab demonstrates that overexpression of G240R *GARS*, but not wild-type *GARS*, results in gross developmental abnormalities and abnormal neuron projections in zebrafish embryos. We can use zebrafish embryos to assess the effect of knockdown and overexpression of candidate proteins and compare the resulting phenotypes to effects observed for mutant *GARS* expression. Importantly, *UBA1*, whose gene product appeared to preferentially bind mutant GARS, knockdown in zebrafish is known to result in motor neuron axonal defects (Wishart et al. 2014), providing an excellent rationale for the use of zebrafish to address these types of question. We would expect knockdown of proteins essential for puncta formation to result in abnormal motor axon morphology in zebrafish embryos. This would indicate that these components are critical to

axon health and that puncta disruption is an important component of *GARS*-mediated CMT disease. If depletion of proteins that specifically interact with wild-type GARS but not with 245-248ΔETAQ GARS leads to a neuronal phenotype, this would suggest that loss of this interaction is important for disease pathogenesis. If overexpression or depletion of proteins that specifically bind to 245-248ΔETAQ GARS, but do not bind wild-type GARS, leads to an axonal phenotype, this would indicate a potential toxic gain-of-function mechanism for *GARS* mutations. Subsequent studies could then focus on identifying pharmacologic agents that alter these interactions as a next step toward developing a treatment for *GARS*-mediated CMT disease.

Treatment of ARS-mediated disease

Currently, there are no therapeutic interventions for patients with either recessive or dominant ARS-associated disease. Since it appears that loss of protein function is a key component in all ARS-mediated phenotypes, identifying ways to improve ARS function would have the potential to treat both dominant and recessive disease. A quantifiable high throughput system could be used to screen small molecules for effects on ARS function. Yeast assays can be adapted for this purpose. Yeast growth of strains expressing human ARS proteins can be used as a proxy for ARS activity *in vivo*. Specifically, liquid yeast cultures expressing mutant ARS can be treated with a library of compounds and serial OD_{600} readings could be used to assess for effects on yeast growth. Alternatively, a colorimetric assay, such as the pyrophosphate exchange assay, could be used in a similar high throughout manner to screen compounds that improve the first step of the aminoacylation reaction; however, this would only allow for visualization of changes to the K_m, not the k_{cat}, and would not reflect *in vivo* activity of the enzyme or the compound. Another aspect of ARS dysfunction that could be used for drug screening is GARS puncta formation. High throughput cell based florescence assays can be used to screen for compounds

that affect protein localization patterns in cells (Starkuviene and Pepperkok 2007). Cells expressing mutant GARS-EGFP can be treated with compounds and fluorescence intensity can be used to assess for alterations in puncta formation in cells. It should be noted that wild-type overexpression of *GARS* in both a mouse and a fly model of disease did not rescue the neurotoxic phenotype in the presence of the mutant allele (Motley et al. 2011; Niehues et al. 2015). As such, if a toxic interaction is identified in our IP/mass spectrometry and zebrafish studies, therapies targeting this interaction may be needed. If so, a high throughput assay based on a read out of the toxic binding partner's activity should be developed.

CONCLUSIONS

Implicating ARS variants in a variety of disease phenotypes highlights the complexity of ARS biology. In general, our work supports findings in the literature that neurons are exquisitely sensitive to perturbations of protein synthesis. However, understanding why mutations in proteins involved in translation, particularly those within the same protein family (*e.g.* ARS enzymes), lead to dramatically different phenotypes is still an open question. It is our hope that studies presented here will serve as the basis for elucidating the pathogenic mechanism by which ARS mutations lead to recessive and dominant disease and will provide therapeutic targets for improving ARS function in patients.

APPENDIX

Appendix A. Oligonucleotide Primers (5' to 3')

Appendix A1. Cloning primers for yeast complementation constructs

Primer	Sequence
Sc_*DPS1*_locus_GWF	ggggacaagtttgtacaaaaaagcaggctTGACCAAGCCAGTTGTTCC
Sc_ *DPS1*_locus_GWR	ggggaccactttgtacaagaaagctgggtAGGGAGCGTGCTACTTGACC
Sc_*FRS1*_GWF	ggggacaagtttgtacaaaaaagcaggctAGTGCCCAACTTGACAGGAA
Sc_*FRS1*_GWR	ggggaccactttgtacaagaaagctgggtGCTAAACATTTACTCAGTAGCGG
*GRS1*_Sc_GWF	ggggacaagtttgtacaaaaaagcaggctAAGAGAGACAAGAAGTGAATAGTGTTTGC
*GRS1*_Sc_GWR	ggggaccactttgtacaagaaagctgggtGTCCGACCATGCGCAGAGCTGC
Sc_*TYS1*_amp_F	ggggacaagtttgtacaaaaaagcaggctCTTCAACTGCCTTTGTCC
Sc_*TYS1*_amp_R	ggggaccactttgtacaagaaagctgggtTCGAGCATCACTAAGAACG

*Gateway adapter sequences are depicted in lower case and the locus specific primer sequences are depicted in upper case.

Appendix A2. Sequencing primers for pTM3xFLAG_GW constructs

Primer	Sequence
pTM_Upstream_Seq_For	CCAGAAAGGACATCTACACC
pTM_Downstream_Seq_Rev	GAACCTGAAACATAAAATGAATGC

Appendix A3. Mutagenic primers for yeast

Primer	**Sequence**
Sc_*ALA1*_K81T_MutF	CAGAGCTGGTGGTACACACAACGATTTAGAA
Sc_*ALA1*_K81T_MutR	TTCTAAATCGTTGTGTGTACCACCAGCTCTG
Sc_*ALA1*_G102R_MutF	TTTTTTGAAATGCTGCGTAACTGGTCGTTTG
Sc_*ALA1*_G102R_MutR	CAAACGACCAGTTACGCAGCATTTCAAAAAA
Sc_*ALA1*_R329H_MutF	GGTAGAGGATATGTTTTGAGACGCATTCTACATAGAGGTGCCCGTTACG
Sc_*ALA1*_R329H_MutR	CGTAACGGGCACCTCTATGTAGAATGCGTCTCAAAACATATCCTCTACC
Sc_*ALA1*_R751G_MutF	GCAAAGGGTATCAGAGGAATTGTTGCTGTTA
Sc_*ALA1*_R751G_MutR	TAACAGCAACAATTCCTCTGATACCCTTTGC
Sc_*DPS1*_A274V_MutF	GGCCTGTGTTCAGGGTTGAAAACTCCAACAC
Sc_*DPS1*_A274V_MutR	GTGTTGGAGTTTTCAACCCTGAACACAGGCC
Sc_*DPS1*_D367H_MutF	ATTGGTGATTTTGAACACTTGAGTACCGAAA

Primer	Sequence
Sc_*DPS1*_D367H_MutR	TTTCGGTACTCAAGTGTTCAAAATCACCAAT
Sc_*DPS1*_D367Y_MutF	ATTGGTGATTTTGAATACTTGAGTACCGAAA
Sc_*DPS1*_D367Y_MutR	TTTCGGTACTCAAGTATTCAAAATCACCAAT
Sc_*DPS1*_H280L_MutF	AAAACTCCAACACCCTCCGTCACATGACCGA
Sc_*DPS1*_H280L_MutR	TCGGTCATGTGACGGAGGGTGTTGGAGTTTT
Sc_*DPS1*_R460H_MutF	TACTGTGACGGCTTCCACTATGGGTGTCCTCC
Sc_*DPS1*_R460H_MutR	GGAGGACACCCATAGTGGAAGCCGTCACAGTA
Sc_*DPS1*_R494G_MutF	GCTTCATTGTTCCCAGGAGATCCAAAGAGAT
Sc_*DPS1*_R494G_MutR	ATCTCTTTGGATCTCCTGGGAACAATGAAGC
Sc_*FRS1*_T256M_mutF	TTTTGATTGATATAATGGCCACCGATAAGACC
Sc_*FRS1*_T256M_mutR	GGTCTTATCGGTGGCCATTATATCAATCAAAA
Sc_*GRS1*_L129P_MutF	GTTGATTGTACTATGCCAACTCCATATGAGG
Sc_*GRS1*_L129P_MutR	CCTCATATGGAGTTGGCATAGTACAATCAAC
Sc_*GRS1*_D146N_MutF	GTTGACAAATTTTCTAATTGGATGTGTAGAG
Sc_*GRS1*_D146N_MutR	CTCTACACATCCAATTAGAAAATTTGTCAAC
Sc_*GRS1*_P244L_MutF	AGGGTTACTTGAGGCTAGAAACTGCCCAAGG

Primer	Sequence
Sc_*GRS1*_P244L_MutR	CCTTGGGCAGTTTCTAGCCTCAAGTAACCCT
Sc_*GRS1*_245-248ETAQdel_MutF	AAAGGGTTACTTGAGGCCAGGTCAATTCTTGAATTTTAA
Sc_*GRS1*_245-248ETAQdel _MutR	TTAAAATTCAAGAATTGACCTGGCCTCAAGTAACCCTTTA
Sc_*GRS1*_R256Q_MutF	TTCTTGAATTTTAACCAGTTGTTAGAATTTA
Sc_*GRS1*_R256Q_MutR	TAAATTCTAACAACTGGTTAAAATTCAAGAA
Sc_*GRS1*_I280F_MutF	CATTTAGAAATGAGTTTTCCCCAAGGGCCG
Sc_*GRS1*_I280F_MutR	CGGCCCTTGGGGAAAACTCATTTCTAAATGA
Sc_*GRS1*_H418R_MutF	ATGATTTGACCGTTCGTTCCAAAAAGACCA
Sc_*GRS1*_H418R_MutR	TGGTCTTTTTGGAACGAACGGTCAAATCAT
Sc_*GRS1*_G526R_MutF	ATTGAACCATCTTTCCGTATTGGCCGTATCAT
Sc_*GRS1*_G526R_MutR	ATGATACGGCCAATACGGAAAGATGGTTCAAT
Sc_*HTS1*_T132I_MutF	TACGTTACGATTTGATCGTTCCTTTTGCCAG
Sc_*HTS1*_T132I_MutF	CTGGCAAAAGGAACGATCAAATCGTAACGTA
Sc_*HTS1*_T132S_MutF	TACGTTACGATTTGAGCGTTCCTTTTGCCAG
Sc_*HTS1*_T132S_MutF	CTGGCAAAAGGAACGCTCAAATCGTAACGTA
Sc_*HTS1*_P134H_MutF	ACGATTTGACCGTTCATTTTGCCAGGTACGT

Primer	Sequence
Sc_*HTS1*_P134H_MutR	ACGTACCTGGCAAAATGAACGGTCAAATCGT
Sc_*HTS1*_R137Q_MutF	CCGTTCCTTTTGCCCAGTACGTTGCCATGAA
Sc_*HTS1*_R137Q_MutR	TTCATGGCAACGTACTGGGCAAAAGGAACGG
Sc_*HTS1*_V155G_MutF	ATCACATTGCAAAAGGCTACAGAAGAGACCA
Sc_*HTS1*_V155G_MutR	TGGTCTCTTCTGTAGCCTTTTGCAATGTGAT
Sc_*HTS1*_D175E_correctMutF	ATTTTACCAGTGTGAGTTTGACGTTGCAGGC
Sc_*HTS1*_D175E_correctMutR	GCCTGCAACGTCAAACTCACACTGGTAAAAT
Sc_*HTS1*_Y330C_mutF	CAAGAGGTTTGGATTGCTATACAGGTTTGAT
Sc_*HTS1*_Y330C_mutR	ATCAAACCTGTATAGCAATCCAAACCTCTTG
Sc_*HTS1*_S356N_MutF	TTCGTTGGTGTCGGCAATATTGCAGCCG
Sc_*HTS1*_S356N_MutR	CGGCTGCAATATTGCCGACACCAACGAA
Sc_*HTS1*_D364Y_MutF	GCCGGTGGTCGATATTATAATTTGGTCAACA
Sc_*HTS1*_D364Y_MutR	TGTTGACCAAATTATAATATCGACCACCGGC
Sc_*TYS1*_G41R_MutF	ACATTTGAAATTATACTGGCGTACCGCGC
Sc_*TYS1*_G41R_MutF	GCGCGGTACGCCAGTATAATTTCAAATGT
Sc_*TYS1*_P167T_Mut F	GGGCTTATCTATACTTTAATGCAAGCG

Primer	Sequence
Sc_*TYS1*_P167T_Mut R	CGCTTGCATTAAAGTATAGATAAGCCC
Sc_*TYS1*_E196K_MutF	AGAGAAAAATTTTTGTCTTAGCGAAAGAA
Sc_*TYS1*_E196K_MutF	TTCTTTCGCTAAGACAAAAATTTTTCTCT
Sc_*TYS1*_E196Q_MutF	AGAGAAAAATTTTTGTCTTAGCGCAAGAA
Sc_*TYS1*_E196Q_MutF	TTCTTGCGCTAAGACAAAAATTTTTCTCT

Appendix A4. Mutagenic primers for human

Primer	Sequence
Hs_*GARS*_A57V_MutF	GAAGCAAAGGAGCTGGTGTTACAGCCCAAAGAT
Hs_*GARS*_A57V_MutR	ATCTTTGGGCTGTAACACCAGCTCCTTTGCTTC
Hs_*GARS*_D146N_mutF	GTAGACAAATTTGCTAACTTCATGGTGAAAG
Hs_*GARS*_D146N_mutR	CTTTCACCATGAAGTTAGCAAATTTGTCTAC
Hs_*GARS*_S211F_mutF	ACTATAATGTAAAATTTCCCATTACTGGAAA
Hs_*GARS*_S211F_mutR	TTTCCAGTAATGGGAAATTTTACATTATAGT
Hs_*GARS*_P244L_MutF	CTGGGTACTTGAGACTAGAAACTGCACAGGGG
Hs_*GARS*_P244L_MutR	CCCCTGTGCAGTTTCTAGTCTCAAGTACCCAG

Primer	Sequence
Hs_*GARS*_245-248ETAQdel_MutF	ATGCCTGGGTACTTGAGACCGGGGATTTTCTTGAATTTCA
Hs_*GARS*_245-248ETAQdel_MutR	TGAAATTCAAGAAAATCCCCGGTCTCAAGTACCCAGGCAT
Hs_*GARS*_I280F_MutF	TCTTTTAGAAATGAGTTCTCCCCTCGATCTG
Hs_*GARS*_I280F_MutR	CAGATCGAGGGGAGAACTCATTTCTAAAAGA
Hs_*GARS*_D500N_MutF	TTTCAGTTAACAAAAAACATGATCAATGTGA
Hs_*GARS*_D500N_MutR	TCACATTGATCATGTTTTTTGTTAACTGAAA
Hs_*GARS*_S581L_MutF	TTGTCAAGGAATTATTGGAAGCCCTGACCAG
Hs_*GARS*_S581L_MutR	CTGGTCAGGGCTTCCAATAATTCCTTGACAA
Hs_*GARS*_G598A_MutF	GTAGACGATTCCTCTGCGTCAATCGGAAGGCGC
Hs_*GARS*_G598A_MutR	GCGCCTTCCGATTGACGCAGAGGAATCGTCTAC

Appendix A5. Sequencing primers for GARS Exon 16

Primer	Sequence
HS_GARS_16F	AGACAGTAGTTAGATAACACTGG
Hs_GARS_16R	TCCAACACAAATCAGTAAGAAGC

REFERENCES

Abe A, Hayasaka K. 2009. The GARS gene is rarely mutated in Japanese patients with Charcot-Marie-Tooth neuropathy. J. Hum. Genet. 54: 310–312.

Achilli F, Bros-Facer V, Williams HP, Banks GT, AlQatari M, Chia R, Tucci V, Groves M, Nickols CD, Seburn KL, Kendall R, Cader MZ, et al. 2009. 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.

Aghamaleky Sarvestany A, Hunter G, Tavendale A, Lamont DJ, Llavero Hurtado M, Graham LC, Wishart TM, Gillingwater TH. 2014. Label-free quantitative proteomic profiling identifies disruption of ubiquitin homeostasis as a key driver of Schwann cell defects in spinal muscular atrophy. J. Proteome Res. 13: 4546–4557.

Amrani N, Sachs MS, Jacobson A. 2006. Early nonsense: mRNA decay solves a translational problem. Nat. Rev. Mol. Cell Biol. 7: 415–425.

Antonellis A, Ellsworth RE, Sambuughin N, Puls I, Abel A, Lee-Lin S-Q, Jordanova A, Kremensky I, Christodoulou K, Middleton LT, Sivakumar K, Ionasescu V, et al. 2003. 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.

Antonellis A, Green ED. 2008. The Role of Aminoacyl-tRNA Synthetases in Genetic Diseases*. Annu. Rev. Genom. Human Genet. 9: 87–107.

Antonellis A, Lee-Lin SQ, Wasterlain A, Leo P, Quezado M, Goldfarb LG, Myung K, Burgess S, Fischbeck KH, Green ED. 2006. Functional Analyses of Glycyl-tRNA Synthetase Mutations Suggest a Key Role for tRNA-Charging Enzymes in Peripheral Axons. Journal of Neuroscience 26: 10397–10406.

Arnez JG, Moras D. 1997. Structural and functional considerations of the aminoacylation reaction. Trends Biochem. Sci. 22: 211–216.

Babady NE, Pang Y-P, Elpeleg O, Isaya G. 2007. Cryptic proteolytic activity of dihydrolipoamide dehydrogenase. Proc. Natl. Acad. Sci. U.S.A. 104: 6158–6163.

Bandyopadhyay AK, Deutscher MP. 1971. Complex of aminoacyl-transfer RNA synthetases. J. Mol. Biol. 60: 113–122.

Bayat V, Thiffault I, Jaiswal M, Tétreault M, Donti T, Sasarman F, Bernard G, Demers-

Lamarche J, Dicaire M-J, Mathieu J, Vanasse M, Bouchard J-P, et al. 2012. Mutations in the mitochondrial methionyl-tRNA synthetase cause a neurodegenerative phenotype in flies and a recessive ataxia (ARSAL) in humans. PLoS Biol. 10: e1001288.

Belostotsky R, Ben-Shalom E, Rinat C, Becker-Cohen R, Feinstein S, Zeligson S, Segel R, Elpeleg O, Nassar S, Frishberg Y. 2011. Mutations in the mitochondrial seryl-tRNA synthetase cause hyperuricemia, pulmonary hypertension, renal failure in infancy and alkalosis, HUPRA syndrome. Am. J. Hum. Genet. 88: 193–200.

Berthonneau E, Mirande M. 2000. A gene fusion event in the evolution of aminoacyl-tRNA synthetases. FEBS Lett. 470: 300–304.

Biesecker LG, Mullikin JC, Facio FM, Turner C, Cherukuri PF, Blakesley RW, Bouffard GG, Chines PS, Cruz P, Hansen NF, Teer JK, Maskeri B, et al. 2009. The ClinSeq Project: piloting large-scale genome sequencing for research in genomic medicine. Genome Res. 19: 1665–1674.

Boeke JD, Trueheart J, Natsoulis G, Fink GR. 1987. 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. Meth. Enzymol. 154: 164–175.

Bomont P, Cavalier L, Blondeau F, Ben Hamida C, Belal S, Tazir M, Demir E, Topaloglu H, Korinthenberg R, Tüysüz B, Landrieu P, Hentati F, et al. 2000. The gene encoding gigaxonin, a new member of the cytoskeletal BTB/kelch repeat family, is mutated in giant axonal neuropathy. Nat. Genet. 26: 370–374.

Bonnefond L, Fender A, Rudinger-Thirion J, Giegé R, Florentz C, Sissler M. 2005. Toward the full set of human mitochondrial aminoacyl-tRNA synthetases: characterization of AspRS and TyrRS. Biochemistry 44: 4805–4816.

Bouhy D, Timmerman V. 2013. Animal models and therapeutic prospects for Charcot-Marie-Tooth disease. Ann. Neurol. 74: 391–396.

Boyd SD. 2013. Diagnostic applications of high-throughput DNA sequencing. Annu Rev Pathol 8: 381–410.

Brautigam CA, Chuang JL, Tomchick DR, Machius M, Chuang DT. 2005. Crystal structure of human dihydrolipoamide dehydrogenase: NAD+/NADH binding and the structural basis of disease-causing mutations. J. Mol. Biol. 350: 543–552.

Burton PR, Tobin MD, Hopper JL. 2005. Key concepts in genetic epidemiology. Lancet 366: 941–951.

Cader MZ, Ren J, James PA, Bird LE, Talbot K, Stammers DK. 2007. Crystal structure of human wildtype and S581L-mutant glycyl-tRNA synthetase, an enzyme underlying distal spinal muscular atrophy. FEBS Lett. 581: 2959–2964.

Cantalupo G, Alifano P, Roberti V, Bruni CB, Bucci C. 2001. Rab-interacting lysosomal protein (RILP): the Rab7 effector required for transport to lysosomes. EMBO J. 20: 683–693.

Cassandrini D, Cilio MR, Bianchi M, Doimo M, Balestri M, Tessa A, Rizza T, Sartori G, Meschini MC, Nesti C, Tozzi G, Petruzzella V, et al. 2013. Pontocerebellar hypoplasia type 6 caused by mutations in RARS2: definition of the clinical spectrum and molecular findings in five patients. J. Inherit. Metab. Dis. 36: 43–53.

Cen S, Javanbakht H, Kim S, Shiba K, Craven R, Rein A, Ewalt K, Schimmel P, Musier-Forsyth K, Kleiman L. 2002. Retrovirus-specific packaging of aminoacyl-tRNA synthetases with cognate primer tRNAs. J. Virol. 76: 13111–13115.

Chance PF, Alderson MK, Leppig KA, Lensch MW, Matsunami N, Smith B, Swanson PD, Odelberg SJ, Disteche CM, Bird TD. 1993. DNA deletion associated with hereditary neuropathy with liability to pressure palsies. Cell 72: 143–151.

Chien C-I, Chen Y-W, Wu Y-H, Chang C-Y, Wang T-L, Wang C-C. 2014. Functional substitution of a eukaryotic glycyl-tRNA synthetase with an evolutionarily unrelated bacterial cognate enzyme. PLoS ONE 9: e94659.

Chihara T, Luginbuhl D, Luo L. 2007. Cytoplasmic and mitochondrial protein translation in axonal and dendritic terminal arborization. Nat. Neurosci. 10: 828–837.

Christodoulou K, Kyriakides T, Hristova AH, Georgiou DM, Kalaydjieva L, Yshpekova B, Ivanova T, Weber JL, Middleton LT. 1995. Mapping of a distal form of spinal muscular atrophy with upper limb predominance to chromosome 7p. Hum. Mol. Genet. 4: 1629–1632.

Csermely P, Schnaider T, Soti C, Prohászka Z, Nardai G. 1998. The 90-kDa molecular chaperone family: structure, function, and clinical applications. A comprehensive review. Pharmacol. Ther. 79: 129–168.

Dawn Teare M, Barrett JH. 2005. Genetic linkage studies. Lancet 366: 1036–1044.

Deglincerti A, Liu Y, Colak D, Hengst U, Xu G, Jaffrey SR. 2015. Coupled local translation and degradation regulate growth cone collapse. Nat Commun 6: 6888.

Del Bo R, Locatelli F, Corti S, Scarlato M, Ghezzi S, Prelle A, Fagiolari G, Moggio M, Carpo M, Bresolin N, Comi GP. 2006. Coexistence of CMT-2D and distal SMA-V phenotypes in an Italian family with a GARS gene mutation. Neurology 66: 752–754.

Delarue M. 1995. Aminoacyl-tRNA synthetases. Curr. Opin. Struct. Biol. 5: 48–55.

Dewey FE, Grove ME, Pan C, Goldstein BA, Bernstein JA, Chaib H, Merker JD, Goldfeder RL, Enns GM, David SP, Pakdaman N, Ormond KE, et al. 2014. Clinical interpretation and implications of whole-genome sequencing. JAMA 311: 1035–1045.

Dietz HC, Cutting GR, Pyeritz RE, Maslen CL, Sakai LY, Corson GM, Puffenberger EG, Hamosh A, Nanthakumar EJ, Curristin SM. 1991. Marfan syndrome caused by a recurrent de novo missense mutation in the fibrillin gene. Nature 352: 337–339.

Diodato D, Melchionda L, Haack TB, Dallabona C, Baruffini E, Donnini C, Granata T, Ragona

F, Balestri P, Margollicci M, Lamantea E, Nasca A, et al. 2014. VARS2 and TARS2 mutations in patients with mitochondrial encephalomyopathies. Hum. Mutat. 35: 983–989.

Dlamini N, Josifova DJ, Paine SML, Wraige E, Pitt M, Murphy AJ, King A, Buk S, Smith F, Abbs S, Sewry C, Jacques TS, et al. 2013. Clinical and neuropathological features of X-linked spinal muscular atrophy (SMAX2) associated with a novel mutation in the UBA1 gene. Neuromuscul. Disord. 23: 391–398.

Dong J, Qiu H, Garcia-Barrio M, Anderson J, Hinnebusch AG. 2000. Uncharged tRNA activates GCN2 by displacing the protein kinase moiety from a bipartite tRNA-binding domain. Mol. Cell 6: 269–279.

Dyck PJ, Lambert EH. 1968. Lower motor and primary sensory neuron diseases with peroneal muscular atrophy. I. Neurologic, genetic, and electrophysiologic findings in hereditary polyneuropathies. Arch. Neurol. 18: 603–618.

Dyck PJ, Thomas PK. 2005. Peripheral Neuropathy.

Easton DF, Deffenbaugh AM, Pruss D, Frye C, Wenstrup RJ, Allen-Brady K, Tavtigian SV, Monteiro ANA, Iversen ES, Couch FJ, Goldgar DE. 2007. A systematic genetic assessment of 1,433 sequence variants of unknown clinical significance in the BRCA1 and BRCA2 breast cancer-predisposition genes. Am. J. Hum. Genet. 81: 873–883.

Edvardson S, Shaag A, Kolesnikova O, Gomori JM, Tarassov I, Einbinder T, Saada A, Elpeleg O. 2007. Deleterious mutation in the mitochondrial arginyl-transfer RNA synthetase gene is associated with pontocerebellar hypoplasia. Am. J. Hum. Genet. 81: 857–862.

Ellsworth RE, Ionasescu V, Searby C, Sheffield VC, Braden VV, Kucaba TA, McPherson JD, Marra MA, Green ED. 1999. The CMT2D locus: refined genetic position and construction of a bacterial clone-based physical map. Genome Res. 9: 568–574.

Elo JM, Yadavalli SS, Euro L, Isohanni P, Götz A, Carroll CJ, Valanne L, Alkuraya FS, Uusimaa J, Paetau A, Caruso EM, Pihko H, et al. 2012. Mitochondrial phenylalanyl-tRNA synthetase mutations underlie fatal infantile Alpers encephalopathy. Hum. Mol. Genet. 21: 4521–4529.

Eskuri JM, Stanley CM, Moore SA, Mathews KD. 2012. Infantile onset CMT2D/dSMA V in monozygotic twins due to a mutation in the anticodon-binding domain of GARS. J. Peripher. Nerv. Syst. 17: 132–134.

Fallini C, Bassell GJ, Rossoll W. 2012. The ALS disease protein TDP-43 is actively transported in motor neuron axons and regulates axon outgrowth. Hum. Mol. Genet. 21: 3703–3718.

Fanen P, Wohlhuter-Haddad A, Hinzpeter A. 2014. Genetics of cystic fibrosis: CFTR mutation classifications toward genotype-based CF therapies. Int. J. Biochem. Cell Biol. 52: 94–102.

Fermin D, Basrur V, Yocum AK, Nesvizhskii AI. 2011. Abacus: a computational tool for extracting and pre-processing spectral count data for label-free quantitative proteomic analysis.

Proteomics 11: 1340–1345.

Francklyn CS, First EA, Perona JJ, Hou Y-M. 2008. Methods for kinetic and thermodynamic analysis of aminoacyl-tRNA synthetases. Methods 44: 100–118.

Freist W, Logan DT, Gauss DH. 1996. Glycyl-tRNA synthetase. Biol. Chem. Hoppe-Seyler 377: 343–356.

Froelich CA, First EA. 2011. Dominant Intermediate Charcot-Marie-Tooth disorder is not due to a catalytic defect in tyrosyl-tRNA synthetase. Biochemistry 50: 7132–7145.

Garber K, Smith KT, Reines D, Warren ST. 2006. Transcription, translation and fragile X syndrome. Curr. Opin. Genet. Dev. 16: 270–275.

Godinic-Mikulcic V, Jaric J, Greber BJ, Franke V, Hodnik V, Anderluh G, Ban N, Weygand-Durasevic I. 2014. Archaeal aminoacyl-tRNA synthetases interact with the ribosome to recycle tRNAs. Nucleic Acids Res.

Gonzaga-Jauregui C, Harel T, Gambin T, Kousi M, Griffin LB, Francescatto L, Ozes B, Karaca E, Jhangiani SN, Bainbridge MN, Lawson KS, Pehlivan D, et al. 2015. Exome Sequence Analysis Suggests that Genetic Burden Contributes to Phenotypic Variability and Complex Neuropathy. Cell Rep 12: 1169–1183.

Gonzalez M, McLaughlin H, Houlden H, Guo M, Yo-Tsen L, Hadjivassilious M, Speziani F, Yang X-L, Antonellis A, Reilly MM, Züchner S, Inherited Neuropathy Consortium (INC). 2013. Exome sequencing identifies a significant variant in methionyl-tRNA synthetase (MARS) in a family with late-onset CMT2. J. Neurol. Neurosurg. Psychiatr.

Gordon-Weeks PR, Fischer I. 2000. MAP1B expression and microtubule stability in growing and regenerating axons. Microsc. Res. Tech. 48: 63–74.

Götz A, Tyynismaa H, Euro L, Ellonen P, Hyötyläinen T, Ojala T, Hämäläinen RH, Tommiska J, Raivio T, Oresic M, Karikoski R, Tammela O, et al. 2011. Exome sequencing identifies mitochondrial alanyl-tRNA synthetase mutations in infantile mitochondrial cardiomyopathy. Am. J. Hum. Genet. 88: 635–642.

Griffin LB, Sakaguchi R, McGuigan D, Gonzalez MA, Searby C, Züchner S, Hou Y-M, Antonellis A. 2014. Impaired Function is a Common Feature of Neuropathy-Associated Glycyl-tRNA Synthetase Mutations. Hum. Mutat.

Groen EJN, Gillingwater TH. 2015. UBA1: At the Crossroads of Ubiquitin Homeostasis and Neurodegeneration. Trends Mol Med 21: 622–632.

Gumy LF, Yeo GSH, Tung Y-CL, Zivraj KH, Willis D, Coppola G, Lam BYH, Twiss JL, Holt CE, Fawcett JW. 2011. Transcriptome analysis of embryonic and adult sensory axons reveals changes in mRNA repertoire localization. RNA 17: 85–98.

Hadchouel A, Wieland T, Griese M, Baruffini E, Lorenz-Depiereux B, Enaud L, Graf E, Dubus

JC, Halioui-Louhaichi S, Coulomb A, Delacourt C, Eckstein G, et al. 2015. Biallelic Mutations of Methionyl-tRNA Synthetase Cause a Specific Type of Pulmonary Alveolar Proteinosis Prevalent on Réunion Island. Am. J. Hum. Genet. 96: 826–831.

Hanada T, Weitzer S, Mair B, Bernreuther C, Wainger BJ, Ichida J, Hanada R, Orthofer M, Cronin SJ, Komnenovic V, Minis A, Sato F, et al. 2013. CLP1 links tRNA metabolism to progressive motor-neuron loss. Nature 495: 474–480.

Hayasaka K, Himoro M, Sato W, Takada G, Uyemura K, Shimizu N, Bird TD, Conneally PM, Chance PF. 1993. Charcot-Marie-Tooth neuropathy type 1B is associated with mutations of the myelin P0 gene. Nat. Genet. 5: 31–34.

He W, Bai G, Zhou H, Wei N, White NM, Lauer J, Liu H, Shi Y, Dumitru CD, Lettieri K, Shubayev V, Jordanova A, et al. 2015. CMT2D neuropathy is linked to the neomorphic binding activity of glycyl-tRNA synthetase. Nature.

Hirota S, Isozaki K, Moriyama Y, Hashimoto K, Nishida T, Ishiguro S, Kawano K, Hanada M, Kurata A, Takeda M, Muhammad Tunio G, Matsuzawa Y, et al. 1998. Gain-of-function mutations of c-kit in human gastrointestinal stromal tumors. Science 279: 577–580.

Hobert O. 2002. PCR fusion-based approach to create reporter gene constructs for expression analysis in transgenic C. elegans. BioTechniques 32: 728–730.

Hou YM, Westhof E, Giegé R. 1993. An unusual RNA tertiary interaction has a role for the specific aminoacylation of a transfer RNA. Proc. Natl. Acad. Sci. U.S.A. 90: 6776–6780.

Howlett GJ, Minton AP, Rivas G. 2006. Analytical ultracentrifugation for the study of protein association and assembly. Curr Opin Chem Biol 10: 430–436.

Hung M-C, Link W. 2011. Protein localization in disease and therapy. J. Cell. Sci. 124: 3381–3392.

Hunter G, Aghamaleky Sarvestany A, Roche SL, Symes RC, Gillingwater TH. 2014. SMN-dependent intrinsic defects in Schwann cells in mouse models of spinal muscular atrophy. Hum. Mol. Genet. 23: 2235–2250.

Ingoglia NA, Giuditta A, Zanakis MF, Babigian A, Tasaki I, Chakraborty G, Sturman JA. 1983. Incorporation of 3H-amino acids into proteins in a partially purified fraction of axoplasm: evidence for transfer RNA-mediated, post-translational protein modification in squid giant axons. J. Neurosci. 3: 2463–2473.

Ionasescu V, Searby C, Sheffield VC, Roklina T, Nishimura D, Ionasescu R. 1996. Autosomal dominant Charcot-Marie-Tooth axonal neuropathy mapped on chromosome 7p (CMT2D). Hum. Mol. Genet. 5: 1373–1375.

Ivanov PA, Chudinova EM, Nadezhdina ES. 2003. Disruption of microtubules inhibits cytoplasmic ribonucleoprotein stress granule formation. Exp. Cell Res. 290: 227–233.

James PA, Cader MZ, Muntoni F, Childs A-M, Crow YJ, Talbot K. 2006. Severe childhood SMA and axonal CMT due to anticodon binding domain mutations in the GARS gene. Neurology 67: 1710–1712.

Jiang X, Litkowski PE, Taylor AA, Lin Y, Snider BJ, Moulder KL. 2010. A role for the ubiquitin-proteasome system in activity-dependent presynaptic silencing. Journal of Neuroscience 30: 1798–1809.

Jordanova A, Irobi J, Thomas FP, Van Dijck P, Meerschaert K, Dewil M, Dierick I, Jacobs A, De Vriendt E, Guergueltcheva V, Rao CV, Tournev I, et al. 2006. Disrupted function and axonal distribution of mutant tyrosyl-tRNA synthetase in dominant intermediate Charcot-Marie-Tooth neuropathy. Nat. Genet. 38: 197–202.

Jordanova A, Thomas FP, Guergueltcheva V, Tournev I, Gondim FAA, Ishpekova B, De Vriendt E, Jacobs A, Litvinenko I, Ivanova N, Buzhov B, De Jonghe P, et al. 2003. Dominant intermediate Charcot-Marie-Tooth type C maps to chromosome 1p34-p35. Am. J. Hum. Genet. 73: 1423–1430.

Kachroo AH, Laurent JM, Yellman CM, Meyer AG, Wilke CO, Marcotte EM. 2015. Evolution. Systematic humanization of yeast genes reveals conserved functions and genetic modularity. Science 348: 921–925.

Karaca E, Weitzer S, Pehlivan D, Shiraishi H, Gogakos T, Hanada T, Jhangiani SN, Wiszniewski W, Withers M, Campbell IM, Erdin S, Isikay S, et al. 2014. Human CLP1 mutations alter tRNA biogenesis, affecting both peripheral and central nervous system function. Cell 157: 636–650.

Keller A, Nesvizhskii AI, Kolker E, Aebersold R. 2002. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal. Chem. 74: 5383–5392.

Kerem B, Rommens JM, Buchanan JA, Markiewicz D, Cox TK, Chakravarti A, Buchwald M, Tsui LC. 1989. Identification of the cystic fibrosis gene: genetic analysis. Science 245: 1073–1080.

Kerjan P, Cerini C, Sémériva M, Mirande M. 1994. The multienzyme complex containing nine aminoacyl-tRNA synthetases is ubiquitous from Drosophila to mammals. Biochim. Biophys. Acta 1199: 293–297.

Kiga D, Sakamoto K, Kodama K, Kigawa T, Matsuda T, Yabuki T, Shirouzu M, Harada Y, Nakayama H, Takio K, Hasegawa Y, Endo Y, et al. 2002. An engineered Escherichia coli tyrosyl-tRNA synthetase for site-specific incorporation of an unnatural amino acid into proteins in eukaryotic translation and its application in a wheat germ cell-free system. Proc. Natl. Acad. Sci. U.S.A. 99: 9715–9720.

Kitzman JO, Starita LM, Lo RS, Fields S, Shendure J. 2015. Massively parallel single-amino-acid mutagenesis. Nat. Methods 12: 203–6– 4 p following 206.

Kruglyak L, Lander ES. 1995. A nonparametric approach for mapping quantitative trait loci. Genetics 139: 1421–1428.

Kudo M, Sugasawa K, Hori T, Enomoto T, Hanaoka F, Ui M. 1991. Human ubiquitin-activating enzyme (E1): compensation for heat-labile mouse E1 and its gene localization on the X chromosome. Exp. Cell Res. 192: 110–117.

Kulkarni M, Smith HE. 2008. E1 ubiquitin-activating enzyme UBA-1 plays multiple roles throughout C. elegans development. PLoS Genet. 4: e1000131.

Kuo F-T, Bentsi-Barnes IK, Barlow GM, Pisarska MD. 2011. Mutant Forkhead L2 (FOXL2) proteins associated with premature ovarian failure (POF) dimerize with wild-type FOXL2, leading to altered regulation of genes associated with granulosa cell differentiation. Endocrinology 152: 3917–3929.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, et al. 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23: 2947–2948.

Latour P, Thauvin-Robinet C, Baudelet-Méry C, Soichot P, Cusin V, Faivre L, Locatelli M-C, Mayençon M, Sarcey A, Broussolle E, Camu W, David A, et al. 2010. A major determinant for binding and aminoacylation of tRNA(Ala) in cytoplasmic Alanyl-tRNA synthetase is mutated in dominant axonal Charcot-Marie-Tooth disease. Am. J. Hum. Genet. 86: 77–82.

Lee HJ, Park J, Nakhro K, Park JM, Hur Y-M, Choi B-O, Chung KW. 2012a. Two novel mutations of GARS in Korean families with distal hereditary motor neuropathy type V. Journal of the Peripheral Nervous System 17: 418–421.

Lee KS, Zhou W, Scott-McKean JJ, Emmerling KL, Cai G-Y, Krah DL, Costa AC, Freed CR, Levin MJ. 2012b. Human sensory neurons derived from induced pluripotent stem cells support varicella-zoster virus infection. PLoS ONE 7: e53010.

Lee SW, Cho BH, Park SG, Kim S. 2004. Aminoacyl-tRNA synthetase complexes: beyond translation. J. Cell. Sci. 117: 3725–3734.

Levitan IB, Kaczmarek LK. 2002. The Neuron. Oxford University Press, USA.

Lin K-P, Soong B-W, Yang C-C, Huang L-W, Chang M-H, Lee I-H, Antonellis A, Antonellis A, Lee Y-C. 2011. The mutational spectrum in a cohort of Charcot-Marie-Tooth disease type 2 among the Han Chinese in Taiwan. PLoS ONE 6: e29393.

Liu-Yesucevitz L, Bassell GJ, Gitler AD, Hart AC, Klann E, Richter JD, Warren ST, Wolozin B. 2011. Local RNA translation at the synapse and in disease. Journal of Neuroscience 31: 16086–16093.

Lodish H. 2008. Molecular Cell Biology. Macmillan.

Lupski JR, de Oca-Luna RM, Slaugenhaupt S, Pentao L, Guzzetta V, Trask BJ, Saucedo-

Cardenas O, Barker DF, Killian JM, Garcia CA, Chakravarti A, Patel PI. 1991. DNA duplication associated with Charcot-Marie-Tooth disease type 1A. Cell 66: 219–232.

MacArthur DG, Manolio TA, Dimmock DP, Rehm HL, Shendure J, Abecasis GR, Adams DR, Altman RB, Antonarakis SE, Ashley EA, Barrett JC, Biesecker LG, et al. 2014. Guidelines for investigating causality of sequence variants in human disease. Nature 508: 469–476.

Martyn CN, Hughes RA. 1997. Epidemiology of peripheral neuropathy. J. Neurol. Neurosurg. Psychiatr. 62: 310–318.

Mathews M, Sonenberg N, Hershey J. 2007. Translational control in biology and medicine.

Mayer MP, Bukau B. 2005. Hsp70 chaperones: cellular functions and molecular mechanism. Cell. Mol. Life Sci. 62: 670–684.

McLaughlin HM, Sakaguchi R, Giblin W, NIH Intramural Sequencing Center, Wilson TE, Biesecker L, Lupski JR, Talbot K, Vance JM, Züchner S, Lee Y-C, Kennerson M, et al. 2011. A Recurrent loss-of-function alanyl-tRNA synthetase (AARS) mutation in patients with charcot-marie-tooth disease type 2N (CMT2N). Hum. Mutat. 33: 244–253.

McLaughlin HM, Sakaguchi R, Liu C, Igarashi T, Pehlivan D, Chu K, Iyer R, Cruz P, Cherukuri PF, Hansen NF, Mullikin JC, Biesecker LG, et al. 2010. 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.

Michaelevski I, Medzihradszky KF, Lynn A, Burlingame AL, Fainzilber M. 2010. Axonal transport proteomics reveals mobilization of translation machinery to the lesion site in injured sciatic nerve. Mol. Cell Proteomics 9: 976–987.

Mihci E, Türkkahraman D, Ellard S, Akçurin S, Bircan I. 2012. Wolcott-Rallison syndrome due to a novel mutation (R491X) in EIF2AK3 gene. J Clin Res Pediatr Endocrinol 4: 101–103.

Millecamps S, Julien J-P. 2013. Axonal transport deficits and neurodegenerative diseases. Nat. Rev. Neurosci. 14: 161–176.

MORTON NE. 1955. Sequential tests for the detection of linkage. Am. J. Hum. Genet. 7: 277–318.

Motley WW, Griffin LB, Mademan I, Baets J, De Vriendt E, De Jonghe P, Antonellis A, Jordanova A, Scherer SS. 2015. A novel AARS mutation in a family with dominant myeloneuropathy. Neurology.

Motley WW, Seburn KL, Nawaz MH, Miers KE, Cheng J, Antonellis A, Green ED, Talbot K, Yang X-L, Fischbeck KH, Burgess RW. 2011. Charcot-Marie-Tooth-linked mutant GARS is toxic to peripheral neurons independent of wild-type GARS levels. PLoS Genet. 7: e1002399.

Murakami T, Garcia CA, Reiter LT, Lupski JR. 1996. Charcot-Marie-Tooth disease and related inherited neuropathies. Medicine (Baltimore) 75: 233–250.

Nakasone N, Nakamura YS, Higaki K, Oumi N, Ohno K, Ninomiya H. 2014. Endoplasmic reticulum-associated degradation of Niemann-Pick C1: evidence for the role of heat shock proteins and identification of lysine residues that accept ubiquitin. Journal of Biological Chemistry 289: 19714–19725.

Nangle LA, Zhang W, Xie W, Yang X-L, Schimmel P. 2007. Charcot-Marie-Tooth disease-associated mutant tRNA synthetases linked to altered dimer interface and neurite distribution defect. Proc. Natl. Acad. Sci. U.S.A. 104: 11239–11244.

Nesvizhskii AI, Keller A, Kolker E, Aebersold R. 2003. A statistical model for identifying proteins by tandem mass spectrometry. Anal. Chem. 75: 4646–4658.

Ng PC, Kirkness EF. 2010. Whole genome sequencing. Methods Mol. Biol. 628: 215–226.

Ng SB, Buckingham KJ, Lee C, Bigham AW, Tabor HK, Dent KM, Huff CD, Shannon PT, Jabs EW, Nickerson DA, Shendure J, Bamshad MJ. 2010. Exome sequencing identifies the cause of a mendelian disorder. Nat. Genet. 42: 30–35.

Ng SB, Turner EH, Robertson PD, Flygare SD, Bigham AW, Lee C, Shaffer T, Wong M, Bhattacharjee A, Eichler EE, Bamshad M, Nickerson DA, et al. 2009. Targeted capture and massively parallel sequencing of 12 human exomes. Nature 461: 272–276.

Nicholson G, Myers S. 2006. Intermediate forms of Charcot-Marie-Tooth neuropathy: a review. Neuromolecular Med. 8: 123–130.

Niehues S, Bussmann J, Steffes G, Erdmann I, Köhrer C, Sun L, Wagner M, Schäfer K, Wang G, Koerdt SN, Stum M, RajBhandary UL, et al. 2015. Impaired protein translation in Drosophila models for Charcot-Marie-Tooth neuropathy caused by mutant tRNA synthetases. Nat Commun 6: 7520.

Paine H. 2015. Does loss of the normal protein function contribute to the pathogenesis of Huntington's disease? Bioscience Horizons 8:.

Park SG, Schimmel P, Kim S. 2008. Aminoacyl tRNA synthetases and their connections to disease. Proceedings of the National Academy of Sciences 105: 11043–11049.

Patzkó Á, Shy ME. 2010. Update on Charcot-Marie-Tooth Disease. Curr Neurol Neurosci Rep 11: 78–88.

Pericak-Vance MA, Speer MC, Lennon F, West SG, Menold MM, Stajich JM, Wolpert CM, Slotterbeck BD, Saito M, Tim RW, Rozear MP, Middleton LT, et al. 1997. Confirmation of a second locus for CMT2 and evidence for additional genetic heterogeneity. Neurogenetics 1: 89–93.

Pierce SB, Chisholm KM, Lynch ED, Lee MK, Walsh T, Opitz JM, Li W, Klevit RE, King M-C. 2011. Mutations in mitochondrial histidyl tRNA synthetase HARS2 cause ovarian dysgenesis and sensorineural hearing loss of Perrault syndrome. Proceedings of the National Academy of Sciences 108: 6543–6548.

Pierce SB, Gersak K, Michaelson-Cohen R, Walsh T, Lee MK, Malach D, Klevit RE, King M-C, Levy-Lahad E. 2013. Mutations in LARS2, encoding mitochondrial leucyl-tRNA synthetase, lead to premature ovarian failure and hearing loss in Perrault syndrome. Am. J. Hum. Genet. 92: 614–620.

Qin X, Hao Z, Tian Q, Zhang Z, Zhou C, Xie W. 2014. Cocrystal structures of glycyl-tRNA synthetase in complex with tRNA suggest multiple conformational states in glycylation. Journal of Biological Chemistry 289: 20359–20369.

Quevillon S, Robinson JC, Berthonneau E, Siatecka M, Mirande M. 1999. Macromolecular assemblage of aminoacyl-tRNA synthetases: identification of protein-protein interactions and characterization of a core protein. J. Mol. Biol. 285: 183–195.

Raeymaekers P, Timmerman V, Nelis E, De Jonghe P, Hoogendijk JE, Baas F, Barker DF, Martin JJ, De Visser M, Bolhuis PA. 1991. Duplication in chromosome 17p11.2 in Charcot-Marie-Tooth neuropathy type 1a (CMT 1a). The HMSN Collaborative Research Group. Neuromuscul. Disord. 1: 93–97.

Ramser J, Ahearn ME, Lenski C, Yariz KO, Hellebrand H, Rhein von M, Clark RD, Schmutzler RK, Lichtner P, Hoffman EP, Meindl A, Baumbach-Reardon L. 2008. Rare missense and synonymous variants in UBE1 are associated with X-linked infantile spinal muscular atrophy. Am. J. Hum. Genet. 82: 188–193.

Ray PS, Fox PL. 2014. Origin and Evolution of Glutamyl-prolyl tRNA Synthetase WHEP Domains Reveal Evolutionary Relationships within Holozoa. PLoS ONE 9: e98493.

Ray PS, Sullivan JC, Jia J, Francis J, Finnerty JR, Fox PL. 2011. Evolution of function of a fused metazoan tRNA synthetase. Mol. Biol. Evol. 28: 437–447.

Reilly MM, Murphy SM, Laurà M. 2011. Charcot-Marie-Tooth disease. J. Peripher. Nerv. Syst. 16: 1–14.

Riley LG, Cooper S, Hickey P, Rudinger-Thirion J, McKenzie M, Compton A, Lim SC, Thorburn D, Ryan MT, Giegé R, Bahlo M, Christodoulou J. 2010. Mutation of the mitochondrial tyrosyl-tRNA synthetase gene, YARS2, causes myopathy, lactic acidosis, and sideroblastic anemia--MLASA syndrome. Am. J. Hum. Genet. 87: 52–59.

Robinson JC, Kerjan P, Mirande M. 2000. Macromolecular assemblage of aminoacyl-tRNA synthetases: quantitative analysis of protein-protein interactions and mechanism of complex assembly. J. Mol. Biol. 304: 983–994.

Rodova M, Ankilova V, Safro MG. 1999. 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.

Rohkamm B, Reilly MM, Lochmüller H, Schlotter-Weigel B, Barisic N, Schöls L, Nicholson G, Pareyson D, Laurà M, Janecke AR, Miltenberger-Miltenyi G, John E, et al. 2007. Further

evidence for genetic heterogeneity of distal HMN type V, CMT2 with predominant hand involvement and Silver syndrome. J. Neurol. Sci. 263: 100–106.

Rosenberg RN. 2008. The Molecular and Genetic Basis of Neurologic and Psychiatric Disease. Lippincott Williams & Wilkins.

Rual J-F, Venkatesan K, Hao T, Hirozane-Kishikawa T, Dricot A, Li N, Berriz GF, Gibbons FD, Dreze M, Ayivi-Guedehoussou N, Klitgord N, Simon C, et al. 2005. Towards a proteome-scale map of the human protein-protein interaction network. Nature 437: 1173–1178.

Safka Brozkova D, Deconinck T, Beth Griffin L, Ferbert A, Haberlova J, Mazanec R, Lassuthova P, Roth C, Pilunthanakul T, Rautenstrauss B, Janecke AR, Zavadakova P, et al. 2015. Loss of function mutations in HARS cause a spectrum of inherited peripheral neuropathies. Brain 138: 2161–2172.

Salazar-Grueso EF, Kim S, Kim H. 1991. Embryonic mouse spinal cord motor neuron hybrid cells. Neuroreport 2: 505–508.

Sambuughin N, Sivakumar K, Selenge B, Lee HS, Friedlich D, Baasanjav D, Dalakas MC, Goldfarb LG. 1998. 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. J. Neurol. Sci. 161: 23–28.

Sampath P, Mazumder B, Seshadri V, Gerber CA, Chavatte L, Kinter M, Ting SM, Dignam JD, Kim S, Driscoll DM, Fox PL. 2004. Noncanonical function of glutamyl-prolyl-tRNA synthetase: gene-specific silencing of translation. Cell 119: 195–208.

Sarna JR, Hawkes R. 2011. Patterned Purkinje cell loss in the ataxic sticky mouse. Eur. J. Neurosci. 34: 79–86.

Schaffer AE, Eggens VRC, Caglayan AO, Reuter MS, Scott E, Coufal NG, Silhavy JL, Xue Y, Kayserili H, Yasuno K, Rosti RO, Abdellateef M, et al. 2014. CLP1 founder mutation links tRNA splicing and maturation to cerebellar development and neurodegeneration. Cell 157: 651–663.

Scheper GC, van der Klok T, van Andel RJ, van Berkel CGM, Sissler M, Smet J, Muravina TI, Serkov SV, Uziel G, Bugiani M, Schiffmann R, Krägeloh-Mann I, et al. 2007a. Mitochondrial aspartyl-tRNA synthetase deficiency causes leukoencephalopathy with brain stem and spinal cord involvement and lactate elevation. Nat. Genet. 39: 534–539.

Scheper GC, van der Knaap MS, Proud CG. 2007b. Translation matters: protein synthesis defects in inherited disease. Nat. Rev. Genet. 8: 711–723.

Schreier AA, Schimmel PR. 1972. Transfer ribonucleic acid synthetase catalyzed deacylation of aminoacyl transfer ribonucleic acid in the absence of adenosine monophosphate and pyrophosphate. Biochemistry 11: 1582–1589.

Schulman BA, Harper JW. 2009. Ubiquitin-like protein activation by E1 enzymes: the apex for

downstream signalling pathways. Nat. Rev. Mol. Cell Biol. 10: 319–331.

Schuske K, Beg AA, Jorgensen EM. 2004. The GABA nervous system in C. elegans. Trends Neurosci. 27: 407–414.

Schwartzentruber J, Buhas D, Majewski J, Sasarman F, Papillon-Cavanagh S, Thiffault I, Thiffaut I, Sheldon KM, Massicotte C, Patry L, Simon M, Zare AS, et al. 2014. Mutation in the nuclear-encoded mitochondrial isoleucyl-tRNA synthetase IARS2 in patients with cataracts, growth hormone deficiency with short stature, partial sensorineural deafness, and peripheral neuropathy or with Leigh syndrome. Hum. Mutat. 35: 1285–1289.

Seburn KL, Nangle LA, Cox GA, Schimmel P, Burgess RW. 2006. An active dominant mutation of glycyl-tRNA synthetase causes neuropathy in a Charcot-Marie-Tooth 2D mouse model. Neuron 51: 715–726.

Shy ME, Patzkó Á. 2011. Axonal Charcot-Marie-Tooth disease. Curr. Opin. Neurol. 24: 475–483.

Simons C, Griffin LB, Helman G, Golas G, Pizzino A, Bloom M, Murphy JLP, Crawford J, Evans SH, Topper S, Whitehead MT, Schreiber JM, et al. 2015. Loss-of-Function Alanyl-tRNA Synthetase Mutations Cause an Autosomal-Recessive Early-Onset Epileptic Encephalopathy with Persistent Myelination Defect. Am. J. Hum. Genet. 96: 675–681.

Sivakumar K, Kyriakides T, Puls I, Nicholson GA, Funalot B, Antonellis A, Sambuughin N, Christodoulou K, Beggs JL, Zamba-Papanicolaou E, Ionasescu V, Dalakas MC, et al. 2005. Phenotypic spectrum of disorders associated with glycyl-tRNA synthetase mutations. Brain 128: 2304–2314.

Skre H. 1974. Genetic and clinical aspects of Charcot-Marie-Tooth's disease. Clin. Genet. 6: 98–118.

Smirnova EV, Lakunina VA, Tarassov I, Krasheninnikov IA, Kamenski PA. 2012. Noncanonical functions of aminoacyl-tRNA synthetases. Biochemistry Mosc. 77: 15–25.

Smit LS, Strong TV, Wilkinson DJ, Macek M, Mansoura MK, Wood DL, Cole JL, Cutting GR, Cohn JA, Dawson DC. 1995. Missense mutation (G480C) in the CFTR gene associated with protein mislocalization but normal chloride channel activity. Hum. Mol. Genet. 4: 269–273.

Sosnay PR, Siklosi KR, Van Goor F, Kaniecki K, Yu H, Sharma N, Ramalho AS, Amaral MD, Dorfman R, Zielenski J, Masica DL, Karchin R, et al. 2013. Defining the disease liability of variants in the cystic fibrosis transmembrane conductance regulator gene. Nat. Genet.

Spinosa MR, Progida C, De Luca A, Colucci AMR, Alifano P, Bucci C. 2008. Functional characterization of Rab7 mutant proteins associated with Charcot-Marie-Tooth type 2B disease. Journal of Neuroscience 28: 1640–1648.

Starkuviene V, Pepperkok R. 2007. The potential of high-content high-throughput microscopy in drug discovery. Br. J. Pharmacol. 152: 62–71.

Steenweg ME, Ghezzi D, Haack T, Abbink TEM, Martinelli D, van Berkel CGM, Bley A, Diogo L, Grillo E, Water Naudé Te J, Strom TM, Bertini E, et al. 2012. Leukoencephalopathy with thalamus and brainstem involvement and high lactate "LTBL" caused by EARS2 mutations. Brain 135: 1387–1394.

Stum M, McLaughlin HM, Kleinbrink EL, Miers KE, Ackerman SL, Seburn KL, Antonellis A, Burgess RW. 2011. An assessment of mechanisms underlying peripheral axonal degeneration caused by aminoacyl-tRNA synthetase mutations. Mol. Cell. Neurosci. 46: 432–443.

Svitkina TM, Verkhovsky AB, Borisy GG. 1996. Plectin sidearms mediate interaction of intermediate filaments with microtubules and other components of the cytoskeleton. J. Cell Biol. 135: 991–1007.

Taft RJ, Vanderver A, Leventer RJ, Damiani SA, Simons C, Grimmond SM, Miller D, Schmidt J, Lockhart PJ, Pope K, Ru K, Crawford J, et al. 2013. Mutations in DARS cause hypomyelination with brain stem and spinal cord involvement and leg spasticity. Am. J. Hum. Genet. 92: 774–780.

Taylor RW, Pyle A, Griffin H, Blakely EL, Duff J, He L, Smertenko T, Alston CL, Neeve VC, Best A, Yarham JW, Kirschner J, et al. 2014. Use of whole-exome sequencing to determine the genetic basis of multiple mitochondrial respiratory chain complex deficiencies. JAMA 312: 68–77.

Thomas PK, Marques W, Davis MB, Sweeney MG, King RH, Bradley JL, Muddle JR, Tyson J, Malcolm S, Harding AE. 1997. The phenotypic manifestations of chromosome 17p11.2 duplication. Brain 120 (Pt 3): 465–478.

Timmerman V, Clowes VE, Reid E. 2013. Overlapping molecular pathological themes link Charcot-Marie-Tooth neuropathies and hereditary spastic paraplegias. Exp. Neurol. 246: 14–25.

Timmerman V, Strickland AV, Züchner S. 2014. Genetics of Charcot-Marie-Tooth (CMT) Disease within the Frame of the Human Genome Project Success. Genes (Basel) 5: 13–32.

Turner EH, Lee C, Ng SB, Nickerson DA, Shendure J. 2009. Massively parallel exon capture and library-free resequencing across 16 genomes. Nat. Methods 6: 315–316.

Turner RJ, Lovato M, Schimmel P. 2000. One of two genes encoding glycyl-tRNA synthetase in Saccharomyces cerevisiae provides mitochondrial and cytoplasmic functions. J. Biol. Chem. 275: 27681–27688.

van Paassen BW, van der Kooi AJ, van Spaendonck-Zwarts KY, Verhamme C, Baas F, de Visser M. 2014. PMP22 related neuropathies: Charcot-Marie-Tooth disease type 1A and Hereditary Neuropathy with liability to Pressure Palsies. Orphanet J Rare Dis 9: 38.

Vance C, Rogelj B, Hortobágyi T, De Vos KJ, Nishimura AL, Sreedharan J, Hu X, Smith B, Ruddy D, Wright P, Ganesalingam J, Williams KL, et al. 2009. Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. Science 323: 1208–1211.

Vester A, Velez-Ruiz G, McLaughlin HM, NISC Comparative Sequencing Program, Lupski JR, Talbot K, Vance JM, Züchner S, Roda RH, Fischbeck KH, Biesecker LG, Nicholson G, et al. 2012. A Loss-of-Function Variant in the Human Histidyl-tRNA Synthetase (HARS) Gene is Neurotoxic In Vivo. Hum. Mutat. n/a–n/a.

Wallen RC, Antonellis A. 2013. To Charge or Not to Charge: Mechanistic Insights into Neuropathy-Associated tRNA Synthetase Mutations. Curr. Opin. Genet. Dev. 23: 302–309.

Wang W, van Niekerk E, Willis DE, Twiss JL. 2007. RNA transport and localized protein synthesis in neurological disorders and neural repair. Dev Neurobiol 67: 1166–1182.

Watts RJ, Hoopfer ED, Luo L. 2003. Axon pruning during Drosophila metamorphosis: evidence for local degeneration and requirement of the ubiquitin-proteasome system. Neuron 38: 871–885.

Weedon MN, Hastings R, Caswell R, Xie W, Paszkiewicz K, Antoniadi T, Williams M, King C, Greenhalgh L, Newbury-Ecob R, Ellard S. 2011. Exome sequencing identifies a DYNC1H1 mutation in a large pedigree with dominant axonal Charcot-Marie-Tooth disease. Am. J. Hum. Genet. 89: 308–312.

Wilson CJ, Pronk JC, Van der Knaap MS. 2005. Vanishing white matter disease in a child presenting with ataxia. J Paediatr Child Health 41: 65–67.

Winklhofer KF, Tatzelt J, Haass C. 2008. The two faces of protein misfolding: gain- and loss-of-function in neurodegenerative diseases. EMBO J. 27: 336–349.

Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, Bangham R, Benito R, Boeke JD, Bussey H, Chu AM, Connelly C, et al. 1999. Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. Science 285: 901–906.

Wishart TM, Mutsaers CA, Riessland M, Reimer MM, Hunter G, Hannam ML, Eaton SL, Fuller HR, Roche SL, Somers E, Morse R, Young PJ, et al. 2014. Dysregulation of ubiquitin homeostasis and β-catenin signaling promote spinal muscular atrophy. J. Clin. Invest. 124: 1821–1834.

Wolf NI, Salomons GS, Rodenburg RJ, Pouwels PJW, Schieving JH, Derks TGJ, Fock JM, Rump P, van Beek DM, van der Knaap MS, Waisfisz Q. 2014. Mutations in RARS cause hypomyelination. Ann. Neurol. 76: 134–139.

Xiao W. 2006. Yeast Protocols. Springer Science & Business Media.

Xie W, Nangle LA, Zhang W, Schimmel P, Yang X-L. 2007. Long-range structural effects of a Charcot-Marie-Tooth disease-causing mutation in human glycyl-tRNA synthetase. Proc. Natl. Acad. Sci. U.S.A. 104: 9976–9981.

Yook K. 2005. Complementation. WormBook 1–17.

Zhang X, Ling J, Barcia G, Jing L, Wu J, Barry BJ, Mochida GH, Hill RS, Weimer JM, Stein Q, Poduri A, Partlow JN, et al. 2014. Mutations in QARS, encoding glutaminyl-tRNA synthetase,

cause progressive microcephaly, cerebral-cerebellar atrophy, and intractable seizures. Am. J. Hum. Genet. 94: 547–558.

Zhao Z, Hashiguchi A, Hu J, Sakiyama Y, Okamoto Y, Tokunaga S, Zhu L, Shen H, Takashima H. 2012. Alanyl-tRNA synthetase mutation in a family with dominant distal hereditary motor neuropathy. Neurology 78: 1644–1649.