Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Kari A. Mattison Date

Towards a Better Understanding of Genetic Contributions to Epilepsy

By

Kari A. Mattison
Doctor of Philosophy

Graduate Division of Biological and Biomedical Science
Genetics and Molecular Biology

_____
Andrew Escayg, Ph.D.
Advisor

_____
Tamara Caspary, Ph.D.
Committee Member

_____
Victor Faundez, M.D., Ph.D.
Committee Member

_____
Meleah Hickman, Ph.D.
Committee Member

_____
Nigel Pedersen, M.D.
Committee Member

Accepted:

_____
Kimberly Jacob Arriola, Ph.D., MPH
Dean of the James T. Laney School of Graduate Studies

_____
Date

Towards a Better Understanding of Genetic Contributions to Epilepsy

By

Kari A. Mattison
B.S., Iowa State University, 2016

Advisor: Andrew Escayg, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in the Graduate Division of Biological and Biomedical Science
Genetics and Molecular Biology
2022

Abstract

Towards a Better Understanding of Genetic Contributions to Epilepsy

By Kari A. Mattison

Epilepsy, characterized by recurrent and unprovoked seizures, is a common neurological disorder affecting more than 3 million individuals in the United States. Use of next-generation sequencing in genetic testing of epilepsy patients has led to the identification of over 200 known and evidence-based epilepsy genes, representing diverse biological functions. The goal of this dissertation was to gain a better understanding of the genetic contributions to epilepsy. We utilized three approaches for the identification of candidate disease variants. The first approach was to analyze gene panel data consisting of approximately 4,800 evidence-based disease genes from 460 patients with epilepsy. We identified 8 variants in *SLC6A1* from this data set which was a similar to the diagnostic yield of established epilepsy genes, *SCN2A* and *CDKL5*. We showed, for the first time, through functional assays that *SLC6A1* variants identified in epilepsy patients result in decreased GABA re-uptake. In the second approach, we analyzed whole exome sequencing data from 218 patients with epilepsy. From this we identified a single *de novo,* heterozygous variant in *ATP6V0C* (p.A138P). Using GeneMatcher and other publicly available datasets, we subsequently identified 25 additional patients with *ATP6V0C* variants. To confirm pathogenicity of the identified variants, we developed functional assays using *Saccharomyces cerevisiae* which demonstrated that the *ATP6V0C* variants result in loss of vacuolar ATPase function. Our work resulted in the largest cohort to-date of patients with *ATP6V0C* variants and provides strong support for *ATP6V0C* as an epilepsy gene. Lastly, we performed whole genome sequencing in two brothers with epilepsy from a consanguineous family to identify variants within shared regions of homozygosity. We identified a homozygous variant, p.G228R, in *CNTNAP2* and subsequently found co-segregation of the same variant in an unrelated family with overlapping clinical presentations. Taken together, three epilepsy genes were identified, each with a unique disease mechanism and function related to neuronal signaling. The knowledge gained from the identification and functional analysis of variants can provide insight into treatment options and/or development of precision therapies for patients.

Towards a Better Understanding of Genetic Contributions to Epilepsy

By

Kari A. Mattison
B.S., Iowa State University, 2016

Advisor: Andrew Escayg, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in the Graduate Division of Biological and Biomedical Science
Genetics and Molecular Biology
2022

Acknowledgments

I would like to thank Dr. Kameryn McCarty Butler for her guidance when I first came into the Escayg lab as a rotation student and trusting me enough to leave the project she had started in my hands after she graduated. The studies in this dissertation would not exist without her determination to establish and maintain a working relationship with EGL Genetics and her foresight to utilize unevaluated portions of genetic testing data from patients with epilepsy to spur new projects for the lab.

My lab mates, past and present, have been valuable to the completion of my degree from providing training in new techniques, critical review of ideas, and helping a genetics student understand neuroscience! Many thanks to Dr. Jennifer Wong, Dr. George Inglis, Dr. Lindsey Shapiro, Thomas Shiu, Samantha Reed, Jordan Owyoung, and Jackie Thelin. To Lindsey, thank you for letting me decorate your desk with pictures of kittens and puppies in silly costumes while you were studying for quals. I like to think I played a major role in you passing! Lindsey was also an invaluable sounding board for my ideas throughout the years, even after she graduated. To Thomas, no, I don't think I will ever "understand one day", but I do hope that maybe one day I'll get at least one of the many houses you promised me over the years. Now that it's in writing here, I'm expecting it to happen. Samantha, the poor soul who had to deal with having her bench right across from mine, thank you for understanding the ways of the Midwest and putting up with my external monologue while I did experiments/thought things through (especially so when I was complaining about having to do western blots!). Lastly, Jordan thank you for always making us laugh and you are always invited to our grad student hangouts!

I am also indebted to my committee members for believing in my projects and being the needed voices of reason if it was time to let something go. Drs. Meleah Hickman and Tamara Caspary provided valuable guidance on my experiments in yeast and fiercely advocated for me as a person when I was unable to do so myself. Dr. Victor Faundez graciously gave me his time when we thought we might foray into mammalian cell culture and provided critical discussion on aspects related to the cell biology of the V-ATPase. Dr. Nigel Pedersen ensured that my project had a patient focus to it, and I always walked away with new questions to ask our collaborating clinicians. I am grateful that these individuals with distinct expertise in very specific areas came together as a cohesive group to not only support my research, but also my development as a scientist and person.

I owe thanks to many individuals who supported my daily functioning as a graduate student at Emory and within the Department of Human Genetics. Emily Morran and Emily Neutens in GDBBS were great resources and I appreciate their hard work which allowed me to

focus on my training and research. To the front office staff in the DoHG – Christi Bell, Regina Gilbert, Janelle Clark, and Bob Sullivan – I owe thanks for making sure I had access to resources I needed and for their support of the daily operations of DoHG. I enjoyed my chats with Bob about the weather when we saw each other (yes, we are both stereotypical Midwesterners in this sense) and am appreciative of his efforts (along with Drs. Stephen Warren and Peng Jin) to help fund the annual GDBBS Student Research Symposium. Working in a lab that uses mouse models, I also want to recognize the contributions of the DAR staff at Emory with maintenance and care of all the mouse lines in our lab allowing us to focus on our experiments. Lisa Kovack also deserves a great deal of thanks for dealing with me (twice!) in physical therapy. She made PT enjoyable and a bright spot in my week, but I am equally happy that I did not need to see her for a third time.

I made many friends within the GMB program and within GDBBS, but two require special mention. First, Master Stephanie Grewenow (yes, I firmly believe this is the only appropriate way to address someone with a master's degree) where do I begin? Is it with our inexplicable deep hatred towards each other first semester? Or is it being told by Tamara we couldn't sit next to each other in Euk? Or is it basically planning GMB events by ourselves for the better part of two years? No matter where I begin words cannot express how grateful I am for your friendship – even if we have to continually remind people we are not the same person despite the obvious height difference and difference in amount of purple clothing owned. And no, you may never write my name with a heart above the "i". Secondly, Lauren Hodkinson, my other purple loving GMB friend. I'm still not really sure if Stephanie and I forced our friendship on you at your first recruitment as a student, or if it was the other way around. Regardless, I'm glad whatever happened, happened. You literally rescued me each week for a month while I had a boot on my right foot and crutches in my hands. One day I'll forgive you for saying I should be on "What Not to Wear."

To my Iowa State friends, I hope that the therapy you need because you're my friend is going well. Dr. Master Katie (Katherine) Heiderscheit, I'm going to assume by now you have accepted the friendship I forced on you during your first year (partially because you were an AnSci major, but mostly because we wanted our den back). Payton Wise (almost DVM) I will assume has mostly recovered from having me as her Community Adviser her freshman year and even more so from me being "scary" during our first ever meeting. I looked forward (and continue to do so) to the monthly video chats with these two and the craziness that would always ensue. Everyone should be very grateful Katie is not a roller coaster engineer by day as Payton and I have witnessed the atrocities (albeit accidental) that would happen.

Outside of research I am very lucky to have gotten involved with various organizations that helped keep me sane when I was troubleshooting experiments for months on end. To my students at the FirstTee of Metro Atlanta thank you for reminding me of the unbridled joy and curiosity that pushed me into science. I will miss getting to see my students each week, but I was starting to feel old working with kids that are now taller than me but were much shorter when I first got to know them! I am also incredibly fortunate to have gotten involved with the *Journal of Emerging Investigators* while at Emory. I am especially indebted to Dr. Sarah Fankhauser for her mentorship as I took on more and more responsibility with JEI and to my crew of Managing and Head Editors. Special thanks to Claire Otero, Dr. Tori Osinski, Nicole Hensch, M.S., and Dr. Breanna Noffsinger for their commitment to JEI and supporting me and my ambitious projects. Claire played an especially big role by continually being my initial sounding board and always volunteering to help with projects. She is no doubt just as responsible for the current success of JEI as I am.

Lastly, I want to thank my family for their support (and free food when they visited). To my parents, both scientists in their own right, thank you for letting me stand at the kitchen sink as a kid and dump spices into it to create my next “potion” because in doing so you fostered my curiosity in the sciences. I apologize for all the eye rolling when you asked questions about my research that I considered “mundane”. To my brother, Lars, I think I enjoyed our semi-weekly calls when you were driving home from work even if we just sat in silence. But I do want to clarify that even though you got your Ph.D. first, it was my idea to get one first. I am also incredibly appreciative of you asking your own PI if I could work in the lab over summer – I especially enjoyed getting to move you out of your apartment by myself while you were in Germany for a conference. A special shoutout to my amazing sister-in-law, Andrea (the original millennial Dr. Mattison), who I genuinely scared the first time she saw me show true excitement (yes, my stoic nature exists even around family!), and to my almost birthday buddy Tobias, I look forward to our joint celebrations throughout the years. Finally, to my aunts and uncles in the Southeast, thank you for letting me crash your holidays so I could still be with family even though my “nuclear” family was back in the Midwest. I will miss getting to be with you all during the holidays, especially Uncle Philip (Oinkel Fill-up) who puts my level of sarcasm/dry humor to absolute shame.

Table of Contents

Tables and Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Appendix A

Appendix B

Abbreviations

aCGH	Array Comparative Genomic Hybridization
ACMG	American College of Medical Genetics and Genomics
AD	Autosomal Dominant
AED	Anti-Epileptic Drug
AMP	Association for Molecular Pathology
ANOVA	Analysis of Variance
AR	Autosomal Recessive
ASO	Antisense Oligonucleotide
BCE	Before Common Era
CDG	Congenital Disorder of Glycosylation
CL	Cytoplasmic Loop
CNV	Copy Number Variant
DDD study	Deciphering Developmental Disorders Study
DEE	Developmental and Epileptic Encephalopathy
DIC	Differential Interference Contrast
eAUC	Empirical Area Under the Curve
EEG	Electroencephalogram
EL	Extracellular Loop
FCD	Focal Cortical Dysplasia
GABA	γ-aminobutyric acid
GEFS+	Generalized Epilepsy with Febrile Seizures Plus
GFP	Green Fluorescent Protein
gnomAD	Genome Aggregation Database
GSA	Global Screening Array
GUS	Gene of Unknown Significance
HGMD	Human Gene Mutation Database
IBD	Identity by Descent
IBS	Identity by State
LOD	Logarithm of the Odds
MAE	Myoclonic-Atonic Epilepsy
MAF	Minor Allele Frequency
MRI	Magnetic Resonance Imaging
MTR	Missense Tolerance Ratio
NGM	Nematode Growth Media
NGS	Next-Generation Sequencing
NMD	Nonsense Mediated Decay
OMIM	Online Mendelian Inheritance in Man
ORF	Open Reading Frame
PE	Poison Exon
RNAi	RNA Interference
ROH	Regions (or runs) of Homozygosity
SV	Synaptic Vesicle
TM	Transmembrane Domain
UTR	Untranslated Region
V-ATPase	Vacuolar H^+-ATPase
VUS	Variant of Unknown Significance
WES	Whole Exome Sequencing
WGS	Whole Genome Sequencing
WT	Wild-Type
XLD	X-Linked Dominant
XLR	X-Linked Recessive

Chapter 1

Introduction

1.1 Overview

The overall goal of my dissertation research was to gain a better understanding of the genetic contributions to epilepsy. I took three different approaches to identify novel genetic variants in patients with epilepsy. The first two approaches used data from a collaboration with EGL Genetics where we had access to de-identified gene panel and whole exome sequencing (WES) data from patients with epilepsy. Four-hundred and sixty patients were screened using a gene panel containing approximately 4,800 evidenced-based disease genes. My analysis of this data identified eight patients with variants in *SLC6A1*. It was surprising to identify this many patients with predicted pathogenic variants in *SLC6A1* as this gene was just emerging as an epilepsy gene at that time. We collaborated with Dr. Baruch Kanner, an expert on the function of GAT-1, encoded by *SLC6A1*, to provide the first functional evidence that the identified *SLC6A1* variants impacted transport of γ-aminobutyric acid (GABA) by GAT-1 (**Chapter 2**).

In the second approach, I used trio-WES data to identify *de novo* variants in patients with epilepsy. From the analysis of 218 patients, I identified a single heterozygous variant in *ATP6V0C*. My analysis of this variant predicted it was likely pathogenic, but *ATP6V0C* was not yet an established disease gene. To provide further support for the pathogenicity of the identified *ATP6V0C* variant, I looked for additional patients by reviewing the literature, publicly available datasets, and connecting with other researchers or clinicians via GeneMatcher. In addition, I developed a functional assay in yeast for *ATP6V0C* variants using approaches previously used in forward genetic screens. I was able to collect 25 additional patients with overlapping phenotypes and provide functional evidence towards variant pathogenicity (**Chapter 3**).

The last approach I used to identify novel variants associated with epilepsy involved the sequencing and analysis of ten consanguineous families with epilepsy. In each family, we performed whole genome sequencing (WGS) on two individuals with epilepsy and identified shared, homozygous variants. I then performed targeted sequencing in other family members to

look for segregation of these identified variants between affected and unaffected individuals. In Family 201, I found that p.Gly228Arg in *CNTNAP2* segregated with disease. We subsequently identified a second family with this same variant and shared clinical presentation providing further evidence for variant pathogenicity (**Chapter 4**). The results from the remaining nine families sequenced is presented in **Appendix B**.

In this introduction I will provide a general overview of the genetics of epilepsy and the impact that genetic testing has had on the ability to provide diagnoses. Then, I will focus on the criteria used to classify variants identified through genetic testing and, in particular, how functional testing can help with the classification of variants of unknown significance. I will also outline the benefits a genetic diagnosis can provide to patients, their families, and the clinical team. Finally, I will conclude with a description of the specific goals of my research which formed the basis for this dissertation.

1.2 Genetics of Epilepsy

Epilepsy, characterized by recurrent and unprovoked seizures [1], is one of the most common neurological disorders affecting more than 3 million individuals in the United States, and is estimated to affect more than 50 million individuals globally [2, 3]. Seizure onset occurs at the highest rate in young children and the elderly and is often accompanied by a negative change to overall quality of life [4, 5]. Individuals with early-onset and/or uncontrolled (refractory) seizures often have other comorbidities such as developmental delay, intellectual disability, anxiety, and, in severe cases, sudden unexpected death in epilepsy [6, 7]. Identification of the genetic basis of an individual's epilepsy can aid in the treatment and prevention of severe comorbidities post-seizure onset.

To date over 200 known and evidence-based epilepsy genes have been identified. I compiled a list of epilepsy genes by first searching publicly available epilepsy and/or seizure gene panel lists and by searching "epilepsy" and "seizure" in Online Mendelian Inheritance in

Man (OMIM) and the Human Gene Mutation Database (HGMD; v.2021.4). The resulting list was refined by excluding genes where epilepsy or seizures was not a primary phenotype reported in patients from the literature (i.e., patients with spinal muscular atrophy that also report seizures). This resulted in the list of known and evidence-based epilepsy genes shown in **Table 1.1**. Genes within **Table 1.1** were placed into the functional category that best fit the known canonical function of the gene/protein. Although some genes could be placed into multiple categories, for simplicity each gene was only assigned to one category.

Ion Channels							
CACNA1A	*CACNA1B*	*CACNA1E*	*CACNA2D2*	*CACNB4*	*CHRNA2*	*CHRNA4*	*CHRNB2*
CLCN4	*FGF12*	*FGF13*	*GABBR2*	*GABRA1*	*GABRA2*	*GABRA5*	*GABRB1*
GABRB2	*GABRB3*	*GABRE*	*GABRG2*	*GRIN1*	*GRIN2A*	*GRIN2B*	*GRIN2D*
HCN1	*HCN2*	*KCNA2*	*KCNB1*	*KCNC1*	*KCNC2*	*KCND2*	*KCNH1*
KCNJ10	*KCNK4*	*KCNMA1*	*KCNQ2*	*KCNQ3*	*KCNT1*	*KCNT2*	*KCTD7*
SCN1A	*SCN1B*	*SCN2A*	*SCN3A*	*SCN8A*	*TRPM3*		
Enzymes							
ADPRS	*ADSL*	*ALDH5A1*	*ALDH7A1*	*ALG1*	*ALG13*	*ALG14*	*ARFGEF2*
ARHGEF9	*ASNS*	*BCKDK*	*CARS2*	*CASK*	*CDKL5*	*CERS1*	*CILK1*
CPA6	*CSNK2B*	*CSTB*	*DENND5A*	*DHDDS*	*DNM1*	*DNM1L*	*DOCK7*
DOLK	*DYRK1A*	*EPM2A*	*EXT2*	*FAR1*	*FARS2*	*GAD1*	*GAMT*
GATM	*GNAO1*	*GOT2*	*HECW2*	*HUWE1*	*ITPA*	*LGI1*	*LIAS*
MDH1	*MDH2*	*MTOR*	*NDUFA8*	*NEDD4L*	*NGLY1*	*NHLRC1*	*NPRL2*
NPRL3	*NUS1*	*OTUD7A*	*PAFAH1B1*	*PAK1*	*PARS2*	*PHGDH*	*PIGA*
PIGC	*PIGH*	*PIGN*	*PIGO*	*PIGS*	*PIGT*	*PIGW*	*PLCB1*
PNKP	*PNPO*	*POLG*	*PPP3CA*	*PPT1*	*PTPN23*	*QARS1*	*RAPGEF2*
RHOBTB2	*SIK1*	*SPATA5*	*ST3GAL3*	*ST3GAL5*	*SYNGAP1*	*SYNJ1*	*TBC1D24*
TIMM50	*TMTC3*	*TPP1*	*TRIM8*	*TRIT1*	*TSC1*	*TSC2*	*TSEN15*
UBA5	*UBE3A*	*UBR7*	*VARS1*	*WARS2*	*WWOX*		
Transporters, Receptors, and Membrane Trafficking							
ATP1A2	*ATP1A3*	*ATP6AP2*	*ATP6V0A1*	*ATP6V0C*	*ATP6V1A*	*ATP6V1B2*	*CLN3*
CLN5	*CLN6*	*CLN8*	*CLTC*	*CPLX1*	*DCX*	*DMXL2*	*DNAJC5*
DYNC1H1	*FOLR1*	*GNAI1*	*GNB1*	*GOSR2*	*GPAA1*	*GPR37L1*	*GRIK2*
GRM7	*IER3IP1*	*IQSEC2*	*KIF2A*	*KIF5C*	*KPNA7*	*MFSD8*	*NRXN1*
PACS1	*PACS2*	*SCAMP5*	*SCARB2*	*SLC12A5*	*SLC13A5*	*SLC1A2*	*SLC25A12*
SLC25A22	*SLC2A1*	*SLC35A2*	*SLC45A1*	*SLC6A1*	*SLC6A8*	*SLC7A6OS*	*SLC9A6*
SNAP25	*STX1B*	*STXBP1*	*SYN1*	*TRAPPC6B*			
Nucleotide Binding and Chromatin Remodeling							
ANKRD11	*ARX*	*CHD2*	*CUX2*	*DEAF1*	*DMBX1*	*EEF1A2*	*EHMT1*
FOXG1	*HCFC1*	*HNRNPU*	*KANSL1*	*MBD5*	*MECP2*	*MEF2C*	*NACC1*
NEUROD2	*PURA*	*RORA*	*RORB*	*SETD1A*	*SETD1B*	*SMARCA2*	*SMC1A*
SNIP1	*STRADA*	*TBL1XR1*	*TCF4*	*YEATS2*	*ZEB2*		
Other							
ADAM22	*APC2*	*ATN1*	*BRAT1*	*CNKSR2*	*CNTN2*	*CNTNAP2*	*COL4A2*
CTNNA2	*CYFIP2*	*DCLK2*	*DEPDC5*	*FBXO28*	*FLNA*	*GAL*	*GPHN*
LAMC3	*LMNB2*	*NBEA*	*NEXMIF*	*NRROS*	*OPHN1*	*PCDH19*	*PLPBP*
PRICKLE1	*PRRT2*	*RELN*	*SEMA6B*	*SPTAN1*	*SZT2*	*TBCD*	*TMX2*
TUBB2A	*TUBG1*	*UFM1*	*WDR26*	*WDR37*	*WDR45*	*WDR45B*	*WDR62*
YWHAG							

Table 1.1. Known and Evidence-Based Epilepsy Genes. Inheritance pattern: Autosomal Dominant (AD; black), Autosomal Recessive (AR; blue), X-Linked Dominant (XLD; green), X-Linked Recessive (XLR; pink), AD/AR (purple), XLD/R (orange). Genes are grouped by known function, including genes that are known to modulate the function of other genes in that group. For example, *FGF12* which modulates the function of voltage-gated sodium channels was placed in the "Ion Channels" category.

1.2.1 Heritability

Seizures were first documented circa 2500 BCE, but it was not until the early 1900s that the first evidence for a genetic component to seizures and epilepsy was demonstrated when Herman Lundborg traced the presence of progressive myoclonic epilepsy in one family back to the 1700s [8, 9]. William Lennox showed further evidence of a genetic basis for epilepsy through the study of 143 twin pairs from 1934-1958. He observed an 86% concordance rate in monozygotic twins and 60% in dizygotic twins when looking at whether both twins were diagnosed with the same epilepsy syndrome [10, 11].

Subsequent studies of twin pairs with epilepsy of unknown cause in the late 1990s and early 2000s also showed a strong genetic contribution. The estimates from these studies suggest that between 70-80% of epilepsy is genetic in nature and heritable [12, 13]. This is further supported by studies that showed acquired epilepsy (i.e., epilepsy developed following an insult to the brain such as stroke, traumatic brain injury, or infection) only accounts for approximately 20% of all epilepsy cases [14, 15].

1.2.2 Modes of Inheritance

As with other genetic diseases, epilepsy can follow many different inheritance patterns. This is primarily determined by the gene, and in some cases the type of variant, that is responsible. The three main Mendelian modes of inheritance, autosomal dominant, autosomal recessive, and X-linked, are discussed in more detail below; however, it is important to note that other factors may play a role in the “assumed” inheritance of epilepsy within a family.

1.2.2.1 Autosomal Dominant

One of the most common modes of inheritance identified in patients with genetic epilepsy is autosomal dominant (AD). With AD epilepsy only one copy of a pathogenic variant is

needed to cause disease. This copy can be inherited from an affected biological parent or arise *de novo* during embryonic development. The first AD epilepsy gene, *CHRNA4,* was identified in 1995 from a large family with multiple affected individuals [16]. Since this original report, over 100 epilepsy genes that follow an AD inheritance pattern have been identified (**Table 1.1**).

1.2.2.2 Autosomal Recessive

Autosomal recessive (AR) variants have the highest rate of prevalence in consanguineous families. In AR epilepsy, two copies of a pathogenic variant in a single gene are needed. This can occur by the affected individual inheriting a variant from each biological parent, or by inheriting a variant from one biological parent and having the other variant arise *de novo* during embryonic development. These variants can either be identical (homozygous) or different (compound heterozygous or biallelic). One of the first identified AR epilepsy genes was *PPT1* in which a homozygous p.R122W variant was identified in 40 Finnish patients with infantile neuronal ceroid lipofuscinosis [17]. The prevalence of this variant in the Finnish population supported the presence of a founder effect which is common in AR disease [17, 18]. Similarly to AD epilepsy genes, over 100 epilepsy genes with AR inheritance are known (**Table 1.1**).

1.2.2.3 X-Linked

The last major mode of inheritance in epilepsy is X-linked where the disease gene is located on the X chromosome. Both X-linked dominant (XLD) and X-linked recessive (XLR) variants are known to cause epilepsy (**Table 1.1**). X-linked epilepsies can be difficult to identify as being genetic given that the phenotype is likely to skip generations and in simplex families it may be more difficult to identify a biological sex-linked inheritance pattern. Additionally, female

carriers may exhibit mild phenotypes for XLR epilepsies due to skewed X-inactivation that can vary by tissue type and further complicate the diagnostic journey for individuals and families.

1.2.2.4 Other Factors

In addition to the three main modes of inheritance discussed above, there are several factors that can impact inheritance and phenotypic presentation in patients with epilepsy, making it harder to determine whether the disorder is genetic in nature and/or if the same genetic variant is responsible for the phenotypes observed in all affected members of a family.

Modifier genes/alleles can affect the expression of other genes and/or the phenotypic presentation of a specific trait or disease. For example, in families with genetic epilepsy with febrile seizures plus (GEFS+) variable disease severity and presentation can be seen between affected family members, suggestive of the presence of modifier genes/alleles [19].

Reduced penetrance can also make it more difficult to identify the genetic cause of epilepsy. Penetrance is the percentage of individuals with a (likely) pathogenic variant that show the expected phenotype, so with reduced penetrance an individual with the likely pathogenic variant might be asymptomatic. In families with AD partial epilepsy with auditory features resulting from variants in *LIG1,* penetrance has been shown to be 67% [20]. Therefore, approximately one-third of individuals in these families will carry the pathogenic variant, but not have epilepsy.

In addition to modifiers and reduced penetrance, somatic mosaicism (variants that arise post-fertilization) can also contribute to inherited epilepsies. With somatic mosaicism, an affected individual may appear to have a *de novo* variant when in fact they have inherited it from an unaffected parent. The timing of post-zygotic mutational events, as well as the tissues impacted, have a large influence on phenotypic presentation in individuals with somatic mosaicism [21].

1.3 Diversifying out of the "Channelopathy" era

The use of advanced technologies, such as next-generation sequencing (NGS), has not only increased the rate of epilepsy gene discovery, but has also increased our knowledge regarding the types of genes responsible for epilepsy. The majority of genes identified between 1995 and 2004 were ion channels, known to be directly involved in the modulation of excitation and inhibition in the central nervous system [4, 22]. This initially led to the idea that epilepsy was a "channelopathy"; however, NGS has resulted in the identification of epilepsy genes with diverse biological functions. **Table 1.1** shows the variety of functions associated with epilepsy genes and further discussion on these different categories is provided below.

1.3.1 Enzymes

Over 90 genes that encode enzymes are known to cause epilepsy (**Table 1.1**). Over half of AR epilepsy genes are enzymes (61/110, 55.5%) and within this category 65% (61/94) of the genes follow AR inheritance patterns. This is unsurprising as 50% of normal enzyme activity is sufficient in most cases to maintain normal enzymatic function [23].

Within the enzyme category a wide range of enzymatic functions are represented. There are E3 ubiquitin ligases, such as *HECW2* and *HUWE1,* that play a key role in the ubiquitination of proteins to maintain cellular homeostasis. Additionally, pathogenic variants in enzymes involved in tRNA synthesis, such as *TRIT1,* which plays a role in modifying tRNAs that bind codons starting with uridine, have been identified in recent years [24].

Interestingly, there are several genes within glycan modification pathways where pathogenic variants lead to congenital disorders of glycosylation (CDG). Over 130 genes are known to lead to CDG, including several that cause CDG with epilepsy [25]. In CDG with epilepsy, pathogenic variants are seen in enzymes involved in asparagine (N)-glycosylation

(e.g., *ALG13, ALG14*) and synthesis of the glycosylphosphatidylinositol anchor (e.g., *PIGA, PIGC, PIGH, etc.*).

1.3.2 Transporters

Unsurprisingly, variants in several genes involved in the transport of neurotransmitters have been identified in patients with epilepsy. GAT-1 (encoded by *SLC6A1*), which recycles GABA, one of the major inhibitory neurotransmitters, is one such transporter. Patients with *SLC6A1* mutations present with myoclonic-atonic epilepsy, intellectual disability, and behavioral problems. In **Chapter 2**, we provide functional evidence to demonstrate that *SLC6A1* variants from patients with epilepsy result in decreased GABA reuptake.

Perhaps less obvious causes of epilepsy are transporters that do not directly transport neurotransmitters. For example, *SLC13A5* is a sodium-dependent citrate transporter that mediates citrate entry into cells [26]. Pathogenic variants in this gene lead to developmental and epileptic encephalopathy (DEE) characterized by onset of refractory seizures in infancy with global developmental delay and intellectual disability [27].

1.3.3 Chromatin Remodelers and Nucleic Acid Binding Proteins

Since the use of NGS has become standard in genetic testing, genes that function as chromatin remodelers and transcription factors have also been identified as epilepsy genes. In 2013, the first report of pathogenic variants in in *CHD2*, a chromatin remodeler, was published. Patients with pathogenic variants in *CHD2* have DEE, and despite ubiquitous expression of *CHD2* throughout the body only the brain appears to be affected in these patients [28].

1.3.4 Other

There are also epilepsy genes that do not fit into these previously mentioned categories. This includes genes that play a role in the structural organization of the brain (e.g., *ARX*), genes that have active roles in cytoskeletal organization (e.g., *FLNA*), and genes that help facilitate cell-to-cell adhesion [29, 30]. One gene involved in cell-to-cell adhesion is *CNTNAP2* which facilitates axo-glial contacts and the clustering of voltage-gated potassium channels on myelinated axons [31, 32]. In **Chapter 4** I describe the identification of a novel homozygous missense variant in *CNTNAP2* from a consanguineous Pakistani family with epilepsy, intellectual disability, and autism spectrum disorder.

There are also genes in which pathogenic variants cause epilepsy, but the function of the gene is not fully understood. One example of this is *NRROS* in which pathogenic variants lead to onset of seizures within the first 18 months of life, followed by developmental regression and cerebral atrophy with brain calcification [33, 34]. Potential functions of *NRROS* include response to Toll-like receptor signaling, regulation of reactive oxygen species production, and processing or activation of TGF-β; however, the exact role and/or the main mechanism behind epilepsy in patients with pathogenic *NRROS* variants is still unknown [34].

1.4 Identifying Pathogenic Variants

There are many tools available to clinicians and patients to identify the genetic cause behind the patient's epilepsy. The type and number of tests needed will vary on family history and if the phenotypic presentation is consistent with a well-known epilepsy subtype with a primary genetic cause. The major tools used in genetic testing for epilepsy are described below.

1.4.1. Genetic Linkage and Targeted Sequencing

Prior to the development of NGS, linkage mapping followed by targeted sequencing were used to identify the likely region harboring the causal gene and the subsequent causal variant. The first pathogenic epilepsy variant, identified in *CHRNA4* (p.S252F), was found in a large Australian family with nocturnal frontal lobe epilepsy after previous linkage analysis had identified 20q13.2-q13.3 as the likely region harboring the causal gene [16, 35, 36].

1.4.2 Copy Number Variants and Gene Panels

Array comparative genomic hybridization (aCGH) was the next method used for the identification of epilepsy genes [37]. aCGH detects copy number variation (CNV) across the genome or can be designed to target a specific region. Across 43 studies, aCGH showed an average diagnostic yield of 7% (95% CI = 7%-11%) [38]. However, with the rise of long-read sequencing and programs that can accurately map CNVs, the use of aCGH has been decreasing.

The introduction of gene panels replaced the slow, gene-by-gene approaches to identify pathogenic variants in patients with epilepsy. Gene panels allow for the sequencing of the coding regions of a large set of predefined target genes. A recent meta-analysis showed that the diagnostic yield from gene panels is 19% (95% CI = 16%-24%) and that the diagnostic yield was higher in cohorts with DEE compared to patients with focal epilepsy [38]. Additionally, the diagnostic yield of a gene panel tends to increase when more genes are included in the analysis; however, the inclusion of more genes also results in the identification of a larger number of variants of unknown significance (VUSs). This can unintentionally lengthen the diagnostic process as each variant must be evaluated to determine its likely contribution to the phenotype.

1.4.3 Whole Exome and Whole Genome Sequencing

WES and WGS provide more unbiased approaches to identifying genetic variants compared to targeted sequencing or a gene panel. Like a gene panel, WES covers the coding regions of the genome but does so for the entire genome instead of over a selected set of genes. WGS is more comprehensive than WES, capturing all non-coding and coding variants in the genome. Both WES and WGS allow the identification of variants in genes that might not otherwise be screened for epilepsy. The diagnostic yield of WES and WGS is generally higher compared to gene panels with 20-40% for WES and an average of 48% (95% CI = 28%-70%) for WGS [38-40]. Sequencing trios (proband + biological parents) can increase the diagnostic yield in both WES and WGS as shared variants can be ruled out when *de novo* inheritance and full penetrance are assumed.

It has also been shown that not only can WES lead to an increased diagnostic yield, but that if it is performed as the first line of genetic testing, it can reduce the overall diagnostic costs by one third [41, 42].

1.5 Follow-up on VUS's and GUS's

As WES and WGS become more commonly used in genetic testing, more VUSs and genes of unknown significance (GUSs) will be identified in patients. This may complicate the diagnostic process as these VUSs and GUSs will need to be thoroughly evaluated to determine whether they are contributing to the observed phenotypes. A VUS is a variant in a known disease gene without enough information on the specific variant or region of the gene/protein to definitively determine whether it is pathogenic or benign. A GUS is a gene whose exact function and/or localization is not known making it difficult to determine what effect, if any, the identified variant may have on gene/protein function. The identification of VUSs and/or GUSs can have unintended consequences for patients and their families. It can be confusing and lead to

psychological impacts given the uncertainty surrounding the results. In addition, the identification of VUSs and GUSs may not be useful in diagnoses or treatment guidance [43].

Given the challenges associated with classifying variants identified through genetic testing, the American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP) have developed criteria (described in more detail below) that are to be used when analyzing variants for pathogenicity [43].

1.5.1 Classification of Variants

The ACMG/AMP guidelines outline two main categories of evidence: one for classifying pathogenic variants and the other for classifying benign variants. Each of these main categories are further divided into several sub-categories (very strong, strong, moderate, and supporting) that indicate the strength of evidence towards classification of a variant (**Table 1.2**). Ultimately it is the weighted combination of all the pieces of evidence that are taken together to assign a classification of pathogenic, likely pathogenic, VUS, likely benign, or benign (**Figure 1.1**).

When classifying variants, all variants start off as a VUS and move towards pathogenic or benign as evidence is applied. The burden to call a variant pathogenic is much higher than the burden to call a variant benign, and some variants will not have enough evidence to be classified as anything other than a VUS (**Figure 1.1**). Using these classification guidelines, I was able to identify likely pathogenic variants in *SLC6A1* (**Chapter 2**) and *ATP6V0C* (**Chapter 3**) from an unselected cohort of patients with epilepsy.

Evidence		Definition
Very Strong	**PVS1**	Null variant (nonsense, frameshift, canonical ±1 or 2 splice site, start codon, exon deletion) when LOF is known mechanism of disease in gene
Strong	**PS1**	Same missense change (amino acid level) as previously established pathogenic variant
	PS2	*de novo* variant with confirmed maternity and paternity
	PS3	Well-established *in vitro* or *in vivo* functional studies demonstrate damaging effect
	PS4	Significant increase of prevalence of variant in cases compared to controls
Moderate	**PM1**	Variant is in a mutation hot-spot and/or in a well-established functional domain
	PM2	Not present in controls (if recessive, at extremely low frequency)
	PM3	Detected in trans with a pathogenic variant (for recessive disorders)
	PM4	Length of mature protein changes due to in-frame indel (in non-repetitive region) or stop-loss
	PM5	Different missense change (amino acid level) at same residue a pathogenic variant has previously been identified
	PM6	*de novo* assumed, but no confirmation of paternity and maternity
Supporting	**PP1**	Segregation with other affected family members in a gene previously linked to the disease
	PP2	Missense variant in a gene intolerant to variation and where missense variants are predicted to be a disease mechanism
	PP3	*in silico* evidence supporting a deleterious effect on the gene or protein
	PP4	Family history and/or patient phenotype specific to a disorder with a known, single-gene etiology
	PP5	Recent report from reputable source classifies variant as pathogenic, but evidence not readily available to testing site
Stand-alone	**BA1**	ExAC/gnomAD, 1000 Genomes, or Exome Sequencing Project allele frequency is >5%
Strong	**BS1**	Greater than expected allele frequency for disorder
	BS2	If full penetrance expected at early age, variant is observed in a healthy adult at the same zygosity expected to cause disease
	BS3	Well-established *in vitro* or *in vivo* functional studies demonstrate no damaging effect
	BS4	Variant does not segregate with affected status in a family
Supporting	**BP1**	Missense variant in gene where truncating variants are primary known cause of disease
	BP2	Detected in trans with pathogenic variant for a dominant disorder with full penetrance or in cis with a pathogenic variant of any inheritance pattern
	BP3	In-frame indels located in a repetitive region with no known function
	BP4	*in silico* evidence does not support a deleterious effect on the gene or protein
	BP5	Identified in case with alternative molecular basis for disease
	BP6	Recent report from reputable source classifies variant as benign, but evidence not readily available to testing site
	BP7	Synonymous variant with no predicted effect on splicing and nucleotide is not highly conserved across species

Table 1.2. ACMG/AMP Evidence for Variant Classification. Each variant is assigned evidence codes for the criteria met. The combination of evidence is used to determine the pathogenicity of the variant (see **Figure 1.1**). For further discussion and caveats to assigning evidence to variants for classification, see Richards *et al.* 2015 [43].

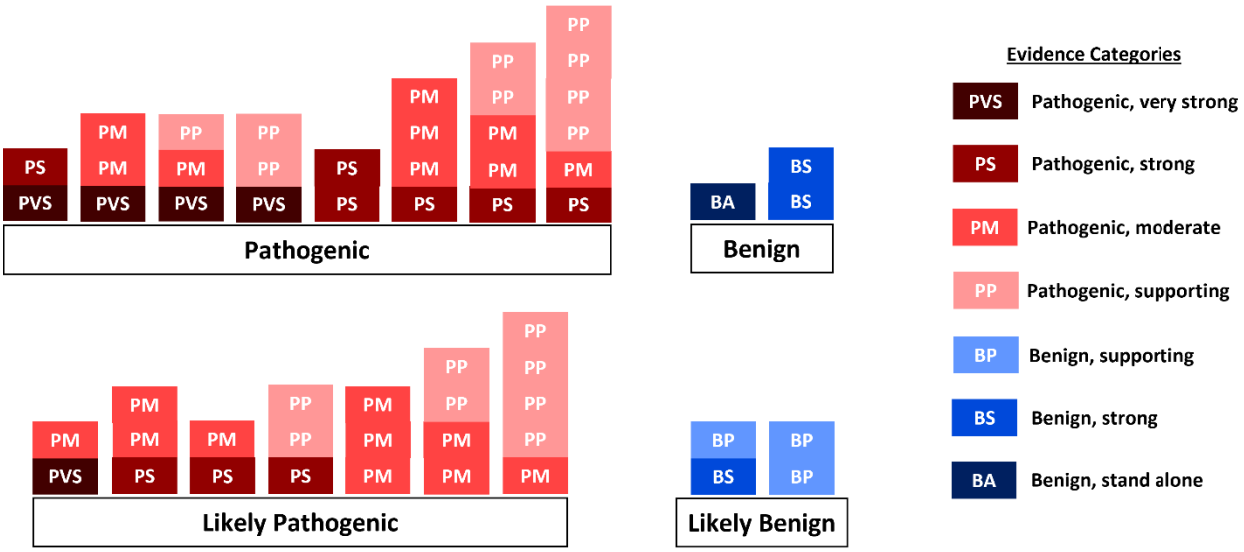


Figure 1.1. Overview of ACMG/AMP Criteria for Variant Classification. Columns represent possible combinations of evidence for variant classification as based on ACMG/AMP guidelines [43]. Evidence categories are represented by two letters, where the first indicates support for pathogenic (P) or benign (B) classification and the second indicates evidence strength. Darker colors represent stronger evidence. Variants with any combination of evidence not shown here are classified as VUSs.

1.5.2 Functional Testing

The initial classification of variants is usually done without the use of *in vitro* or *in vivo* functional experiments. This is because much of the evidence used to classify variants can be determined through sequencing of the proband's biological relatives, publicly available databases (such as gnomAD or HGMD), and *in silico* prediction tools (**Table 1.2**). Thus, if a variant is classified as a VUS after the initial round of classification, performing functional testing can be a powerful tool to provide definitive variant classification. A previous study showed that 97% of VUSs could be reclassified through the addition of functional evidence, and interestingly only 3% of VUSs reclassified were classified as benign [44].

However, there are many caveats that one must consider when using functional testing to provide evidence for pathogenicity. In 2019, new recommendations were made for the application of functional evidence when classifying variants [45]. These recommendations include accounting for the physiological context of the system being used to perform the functional tests, that molecular consequences are accurately modeled particularly for splice variants or those predicted to result in nonsense mediated decay (NMD), and ensuring that appropriate controls, which can demonstrate the extreme outputs for the assay, are used [45].

In **Chapter 2**, functional testing of *SLC6A1* missense variants demonstrated a reduction in GABA uptake in cells, and a splicing assay showed that a splice site variant leads to exon skipping. In **Chapter 3**, functional testing of multiple *ATP6V0C* missense variants in yeast demonstrated damaging effects which would result in all variants being classified as pathogenic if the PS3 criteria was applied.

1.5.3 Problem of Singleton Cases

The other major challenge with establishing the pathogenicity of VUSs and GUSs is that they are often identified in singleton cases in which no other individual in the cohort has been

reported with the same variant (and the gene might also be a GUS). Some of the ACMG/AMP criteria rely on the classification of previous variants identified in a gene or the proportion of cases vs controls that have the variant (**Table 1.2**). These criteria can be difficult to meet unless the variant has been previously published and this further highlights the weight that the total number of likely pathogenic or pathogenic variants identified in a single gene provides to move variant classification away from a VUS.

The use of online variant repositories (e.g., ClinVar) as well as “matchmaking” services has greatly helped to reduce the number of singleton variants. One tool that is particularly useful for both clinicians and researchers is GeneMatcher [46]. GeneMatcher is an online service where researchers and clinicians can provide information about an identified variant and the general phenotype of the patient. They are then connected with other individuals who match selected criteria, such as gene, variant type, or phenotype, and information can be shared between the two parties. I demonstrate the power of these services in **Chapter 3**. I had initially identified one patient with a variant in *ATP6V0C*; however, I was able to identify an additional 25 patients with *ATP6V0C* variants using GeneMatcher and ClinVar.

1.6 Benefits of a Genetic Diagnosis

A genetic diagnosis provides many benefits to the patient, their family, and the clinical team caring for them. Knowing and understanding the genetic basis of a patient’s epilepsy can provide the opportunity to look at the natural progression of symptoms within previously published patients. From this knowledge, preventative measures can be started in order to stop or lessen the onset of some symptoms. Understanding the mechanism behind a patient’s seizures also allows the clinical team to prescribe more efficacious drugs (discussed more below). Importantly, there can be both positive and negative psychological effects of receiving a genetic diagnosis. One positive impact is the ability to join a support group of individuals and families with the same genetic diagnosis which is also discussed below. However, clinicians and

genetic counselors must also be aware that a genetic diagnosis may lead patients and family members to believe that the epilepsy is a direct result of something they did [47].

1.6.1 Therapeutics

There are many different classes of anti-epileptic drugs (AEDs), and they can be broadly grouped based on their mechanism of action. AEDs such as phenytoin and carbamazepine act as sodium channel blockers, others are known to enhance $GABA_A$ receptor activity (benzodiazepines), and some have multiple mechanisms of action, acting on more than one type of ion channel and/or receptor, such as lamotrigine and topiramate [48]. Knowing the genetic cause of a patient's epilepsy can allow the clinical team to make an informed decision on the AED(s) to prescribe in order to most effectively control seizures. For example, patients with Dravet syndrome typically have *SCN1A* variants that cause a loss of function in $Na_v1.1$, a major voltage-gated sodium channel in the central nervous system [49]. Prescribing a sodium channel blocker in these patients would exacerbate the seizure phenotype by further reducing the already low $Na_v1.1$ function and should be avoided.

In recent years a substantial amount of work has gone into developing and testing precision therapies for patients with epilepsy. The research that has yielded the most promising results to date is the use of antisense oligonucleotides (ASOs) to modulate mRNA expression. ASOs can decrease mRNA levels by binding to RNA creating a RNA-DNA hybrid structure that will be degraded by RNase H1 or by blocking the binding of RNA binding proteins required for translation [50]. Additionally, ASOs can be used to modulate splicing of mRNA transcripts by causing targeted exon skipping or inclusion [50].

An ASO (STK-001) that upregulates wild-type *SCN1A* is currently in clinical trials. STK-001 increases wild-type *SCN1A* mRNA levels and therefore wild-type $Na_v1.1$ protein levels by preventing the inclusion of a naturally occurring "poison" exon that leads to the degradation of mRNA transcripts [51]. In preclinical mouse studies, STK-001 administration decreased seizure

frequency and increased survival in pre-and post-weaning *Scn1a*$^{+/-}$ knock-out mice [51]. Studies where ASOs have been used to decrease mRNA levels also show promising preclinical results. Pathogenic variants in *SCN8A,* which encodes $Na_V1.6$, have been shown to have a gain of function mechanism [52]. ASO treatment targeting the 3' untranslated region (UTR) of *Scn8a* resulted in delayed seizure onset and increased survival in mice with a gain of function *Scn8a* allele (*Scn8a*$^{R1872W/+}$) [53].

1.6.2 Support Groups

In addition to a better understanding of appropriate treatment options, a genetic diagnosis can also provide a patient and their family a chance to join a support and/or family advocacy organization. These organizations are usually nonprofits that advocate for patients, provide support for the families, and may even fund research related to their cause. Many gene and syndrome specific organizations exist already, such as the International *SCN8A* Alliance, the Dravet Foundation (*SCN1A*-derived epilepsy), and the Lennox-Gastaut Syndrome Foundation. In **Chapter 2**, I describe my work on *SLC6A1* which was an emerging epilepsy gene at the time of publication of our work (2018); however, thanks to the formation of SLC6A1Connect, a patient advocacy group, our understanding of the disease mechanism and development of specific treatments has rapidly progressed in the last four years.

1.7 Summary and Goals of this Dissertation

In summary, genetic testing in the epilepsies has proved to be a powerful tool both for clinicians and patients with epilepsy. However, one must be cognizant of the likelihood that many VUS will likely be identified, and further testing will need to be done in order to provide more definitive variant classification. Thus, the overall goal of my dissertation was to analyze genetic testing data for VUSs that were likely to explain the epilepsy seen in patients and

develop functional tests to establish whether identified variants were likely to have a deleterious effect on protein function.

In **Chapter 2**, I show that *SLC6A1* variants identified in patients with epilepsy and intellectual disability decrease uptake of GABA. In **Chapter 3**, I show that *ATP6V0C* variants in patients with severe neurodevelopmental disorders, including epilepsy, result in decreased function of the vacuolar ATPase (V-ATPase). This also represents the first report of a large cohort of patients with *ATP6V0C* variants and neurological disease. In **Chapter 4**, I report the findings from WGS and segregation analysis in a consanguineous Pakistani family with two brothers affected with epilepsy, intellectual disability, and autism spectrum disorder. Finally, in **Chapter 5**, I provide a discussion of these findings, future directions, and a general discussion on the future of genetic testing in the epilepsies.

Chapter 2

SLC6A1 Variants Identified In Epilepsy Patients Reduce γ-Aminobutyric Acid Transport

Reproduced with Permission From:

Mattison KA,* Butler KM,* Inglis GAS, Dayan O, Boussidan H, Bhambhani V, Philbrook B, da Silva, C, Alexander, JJ, Kanner BI, and Escayg A (2018) *SCL6A1* variants identified in epilepsy patients reduce γ-aminobutyric acid transport. *Epilepsia* **59**(9):e145-e41. doi:10.1111/epi.14531

*__Co-first authors__.

2.1 Summary

In this chapter, I describe the identification and functional testing of eight *SLC6A1* variants identified from the analysis of sequence data from 460 patients with epilepsy screened at EGL Genetics using the Epilepsy and Seizure Disorders (ESD) panel. The identification of such a large number of likely causative variants in *SLC6A1* was unexpected since the contribution of *SLC6A1* to epilepsy was not clear at that time. The almost 2% diagnostic yield for *SLC6A1* variants from the ESD panel was similar to what was seen with *SCN2A* and *CDKL5* in a previous analysis of 339 patients screened using the ESD panel [54], suggesting that *SLC6A1* was a major epilepsy gene. In addition, the diagnostic yield and the top identified genes from the ESD panel was similar to other gene panel analyses with similar patient cohorts [55, 56].

SLC6A1 encodes GAT-1 which is responsible for the reuptake/recycling of GABA at the synapse. Through a collaboration with Dr. Baruch Kanner, we were able to show for the first time, that *SLC6A1* missense variants identified in patients with epilepsy decrease GABA reuptake. This assay was performed in HeLa cells which do not endogenously express *SCL6A1*. Therefore, we assumed that any GABA uptake into HeLa cells was due to the activity of the *SLC6A1* constructs expressed and not from any endogenous transporters. In addition, we were able to confirm that an identified splice site variant resulted in exon skipping, leading to nonsense mediated decay of the mutant transcript.

This work established a likely mechanism underlying seizure generation in patients with *SLC6A1*-derived epilepsy. Since publication of our study, numerous other studies have been published on *SLC6A1* variants and their involvement in epilepsy and other neurodevelopmental disorders. Overall, our study in combination with subsequent reports has led to a better understanding of the mechanism behind *SLC6A1*-derived epilepsy and more efficacious treatments for these patients are being developed and tested using this knowledge.

The published manuscript that resulted from this work is presented below with minor modifications.

2.2 Introduction

SLC6A1 encodes the electrogenic sodium and chloride-coupled GABA transporter, GAT-1, which is responsible for the re-uptake of the inhibitory neurotransmitter GABA from the synapse [57]. Previous reports identified heterozygous *SLC6A1* variants in patients with myoclonic-atonic epilepsy (MAE; also called myoclonic-astatic epilepsy or Doose syndrome) and other generalized epilepsies [58-62]. While it has been hypothesized that *SLC6A1* epilepsy mutations are likely to be loss-of-function, the functional effect of these reported variants has not yet been experimentally determined.

In the present study, we report eight *SLC6A1* variants identified from an unselected cohort of 460 epilepsy patients referred for genetic testing. Using splicing and GABA transport assays, we demonstrated that these variants reduce or abolish the function of the GAT-1 GABA transporter.

2.3 Subjects and Methods

Next-Generation Sequencing and Sanger Confirmation

Next-generation sequencing and Sanger confirmation were performed as previously published [54]. Briefly, DNA samples from 460 epilepsy patients underwent targeted resequencing of approximately 4,800 genes associated with human disease. Variants were called within the coding exons and ±10bp into the introns using NextGENe® (SoftGenetics, State College, PA). Variants were filtered for population frequency using the genome Aggregation Database (gnomAD). The Institutional Review Board of Emory University approved this study.

GABA Transport Assay

Each patient variant was generated by site-directed mutagenesis of the rat GAT-1 cDNA (which shares 98% amino acid identity with human GAT-1) in the vector pBluescript SK$^-$ (Stratagene) as described previously [63, 64].

HeLa cells were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, 200 units/ml penicillin, 200 µg/ml streptomycin, and 2 mM glutamine. Infection with recombinant vaccinia/T7 virus vTF7-3 and subsequent transfection with plasmid DNA in the expression vector pBluescript SK^- were done as previously described [65].

GABA transport assays were performed as previously published.[66] Briefly, using the sub-saturating [^{3}H]-GABA concentration of 22.3 nM, transport of radioactive GABA was performed for 10 minutes. The uptake of GABA by each mutant was normalized to that of wild-type (WT) GAT-1. Statistical evaluation of GAT-1 transporter activity used a one-way analysis of variance (ANOVA) with a post-hoc Dunnett's multiple comparison test, where $p<0.05$ was significant.

Minigene Splicing Assay

A 1450bp fragment containing the last 276 bases of *SLC6A1* intron 7 through the first 263 bases of intron 10 (NM_003042.3) was PCR amplified from human genomic DNA using Phusion HotStart II Polymerase (Invitrogen) (5'-CACCTCCTGTCACCACATGCAATAC-3', 5'-CTGCCATCTTTCTAGCTCCATAC-3'). The fragment was cloned into a pENTR/D-TOPO vector (Invitrogen) and verified by restriction digestion using *Asc*I and *Sac*II. The *SLC6A1* fragment was then gateway-cloned into the pDESTsplice minigene splicing vector and verified by restriction digestion using *Hind*III, *Xho*I, and *Sac*II [67]. The c.850-2A>G variant was subsequently introduced into the *SLC6A1* fragment by site-directed mutagenesis using the QuikChange XL Site-Directed Mutagenesis Kit (Agilent Technologies). Sanger sequencing confirmed the presence of the c.850-2A>G variant as well as the absence of any unwanted substitutions.

Human embryonic kidney 293 (HEK293) cells were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum. HEK293 cells were transfected with 1.25 µg of pDESTsplice vector using Lipofectamine 3000 (Invitrogen). The pDESTsplice vector contained one of two inserts: 1) wild-type *SLC6A1* (WT), or 2) *SLC6A1* c.850-2A>G

(Splice Mutant). Transfections were performed in triplicate. RNA was extracted using the PureLink RNA Mini Kit (Invitrogen) 24 hours post-transfection. Before first strand cDNA synthesis, RNA was treated with DNase at 37°C for 30 min. First strand cDNA synthesis was performed using SuperScript III Reverse Transcriptase with oligo(dT) primers (Invitrogen), and subsequent cDNA was amplified using primers to exons 8 and 10 of *SLC6A1* (5'-GATCATCCTGTTCTTCCGTGG-3', 5'-GAGAAGATGACGAATCCTGCG-3').

PCR products were visualized by gel electrophoresis. Bands were extracted using the PureLink Quick Gel Extraction Kit (Invitrogen), and Sanger sequenced to verify the splice products.

2.4 Results

2.4.1 *SLC6A1* Variants Identified in Individuals with Epilepsy

To identify disease-associated *SLC6A1* variants, we examined all coding exons of *SLC6A1* (NM_003042.3) from 460 unselected epilepsy patients referred for gene panel analysis at EGL Genetics. We identified eight heterozygous variants: five missense, one nonsense, one splice-site, and one in-frame deletion (**Figure 2.1, Table 2.1**). All substitutions affected highly conserved residues (**Figure 2.1**) and were predicted to be damaging by *in silico* tools (**Table 2.1**).

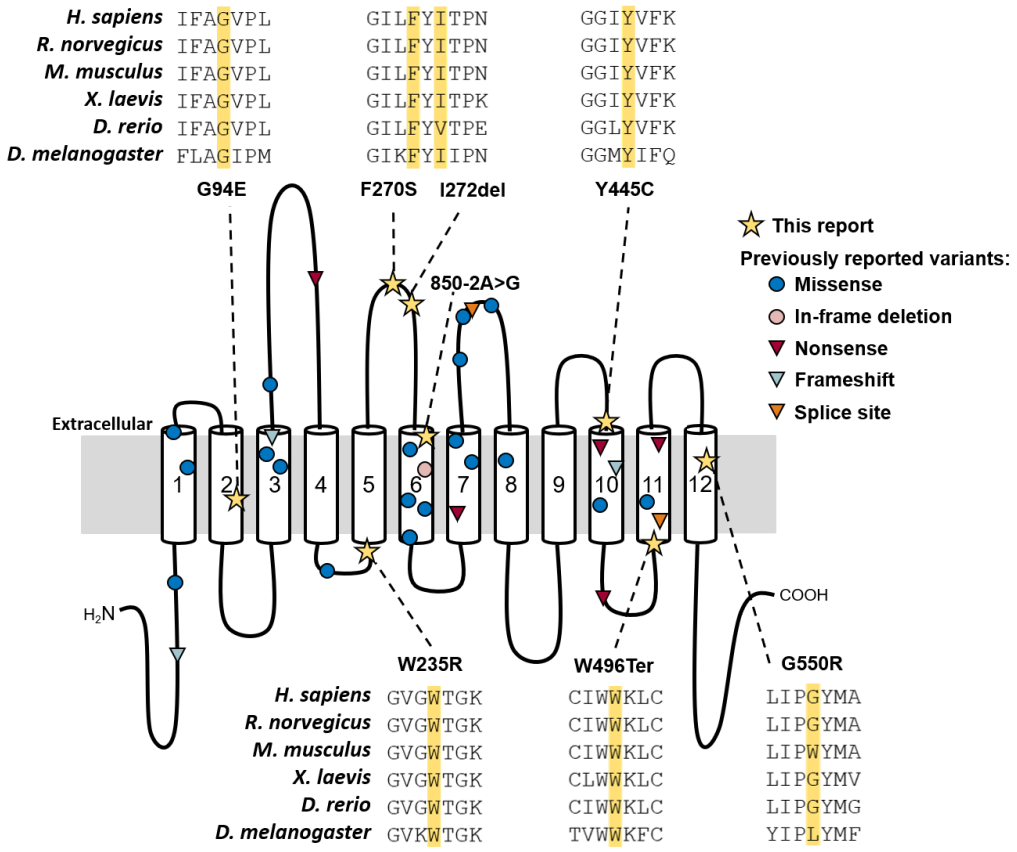


Figure 2.1. Location and Conservation of GAT-1 Variants Identified in Epilepsy Patients. Schematic representation of the GAT-1 protein. Numbered cylinders represent the twelve transmembrane domains. Stars indicate approximate locations of the eight variants identified in the present study. Circles indicate previously published missense (dark blue) and in-frame deletion (pink) variants, and truncating variants are indicated by triangles: nonsense (red), frameshift (light blue) and splice-site (orange) [58-62, 68-75]. Variant positions are based on the previously published LeuT crystal structure [76]. The affected amino acid residues are highlighted in yellow in the protein alignments. The following GAT-1 protein sequences were used in the alignments: *H. sapiens*, NP_003033.3; *R. norvegicus*, NP_0077347.1; *M. musculus*, AAH59080.1; *X. laevis*, AAI70214.1; *D. rerio*, NP_001007363.1; *D. melanogaster*, NP_651930.2.

	Patient 1	Patient 2	Patient 3[†]	Patient 4
Sex	Female	Female	Male	Male
Variant[a]	c.218G>A (p.G94E)	c.703T>C (p.W235R)	c.809T>C (p.F270S)	c.815_817delTCA (p.I272del)
Inheritance	unknown	unknown (adopted)	*de novo*	*de novo*
gnomAD[b]	0	0	0	0
CADD	29	24.5	24.5	-
PolyPhen-2[c]	Probably Damaging	Probably Damaging	Benign	-
SIFT[c]	Deleterious	Deleterious	Deleterious	Deleterious
Epilepsy Syndrome	NA	Intractable absence epilepsy	MAE; Intractable primary generalized epilepsy	MAE: Intractable primary generalized epilepsy
Seizure Types	NA	Generalized tonic-clonic, absence, atypical absence, tonic	Absence, eyelid myoclonias, atonic head nodding	Intractable absence, history of focal, generalized tonic and atonic
Intellectual Disability	NA	Moderate	Moderate	Moderate
Behavioral	NA	ASD	ADHA, ODD/aggressive behavior	Normal
Other Features	NA	Hypogammaglobulinemia, precocious puberty, insomnia, *in utero* drug and HIV exposure	Hypotonia, insomnia, hypoxia at birth	Bilateral upper extremity tremor, mild tandem gait ataxia
[^{3}H]-GABA uptake[d]	0%	27%	2%	13%

	Patient 5[‡]	Patient 6	Patient 7	Patient 8
Sex	Female	Female	Male	Male
Variant[a]	c.850-2A>G	c.1334A>G (p.Y445C)	c.1487G>A (p.W496Ter)	c.1648G>A (p.G550R)
Inheritance	*de novo*	unknown	unknown	unknown
gnomAD[b]	0	0	0	0
CADD	23.5	26.6	43	25
PolyPhen-2[c]	-	Probably Damaging	-	Probably Damaging
SIFT[c]	-	Deleterious	-	Deleterious
Epilepsy Syndrome	Generalized epilepsy	Generalized epilepsy	Generalized epilepsy	Generalized epilepsy
Seizure Types	Focal and generalized	NA	Absence	Absence, tonic-clonic
Intellectual Disability	NA	NA	NA	NA
Behavioral	NA	NA	ASD, behavioral problems	NA
Other Features	Mild speech delay, episodes of falling and eye deviation	NA	Macrocephaly, mild hypotonia	NA
[^{3}H]-GABA uptake[d]	-	6%	0%	0%

Table 2.1. *SLC6A1* Variants Identified from 460 Individuals with Epilepsy

[a]Based on reference sequence NM_003042.3.

[b]Number of times observed in the gnomAD database out of approximately 277,200 alleles.

[c]PolyPhen-2 and SIFT score missense variants.

[d][^{3}H]-GABA uptake relative to wild-type.

[†]Previously published as Patient 10 in Johannessen *et al.* [60]

[‡]Previously published as Patient 32 in Johannessen *et al.* [60]

Abbreviations: ADHD = attention deficit hyperactivity disorder, ASD = autism spectrum disorder, ID = intellectual disability, MAE = myoclonic-atonic epilepsy, NA = not available, ODD = oppositional defiant disorder.

Additionally, all identified variants were absent from gnomAD, a population database that includes approximately 277,200 alleles from whole-exome and genome sequencing but excludes individuals with severe pediatric diseases. Five of the identified variants were novel, while three (p.G550R, p.F270S, c.850-2A>G) were previously reported (**Table 2.1**). The p.G550R variant carried by Patient 8 was previously identified in an autism patient by Wang and colleagues [74]. Coincidentally, two of the patients in the current study were included in a recent publication by Johannesen and colleagues. Specifically, Patient 3 (p.F270S) and Patient 5 (c.850-2A>G) correspond to Patient 10 and Patient 32, respectively, in the study by Johannesen and colleagues [60].

2.4.2 GABA Transport is Reduced by Variants Identified in Epilepsy Patients

Variant location in **Figure 2.1** is based on the published crystal structure of LeuT, the bacterial homolog of Na^+/Cl^- -dependent neurotransmitter transporters [76]. Six of the variants affect the transmembrane helices (TM) of GAT-1 or their extracellular boundaries, and two variants are in the third extracellular loop (EL) 3 connecting TMs 5 and 6 (see **Figure 2.1** and Discussion). Introduction of the identified *SLC6A1* variants into the rat GAT-1 sequence resulted in a reduction or complete loss of [^{3}H]-GABA transport activity (**Figure 2.2**). Small levels of residual activity were seen for transporters containing p.F270S, p.Y445C and p.I272del variants (2%, 6%, and 13% of wild-type activity, respectively), while p.G94E, p.W496Ter, and p.G550R abolished GABA transport activity. p.W235R had the smallest impact; however, it still reduced GABA transport to 27% of wild-type activity.

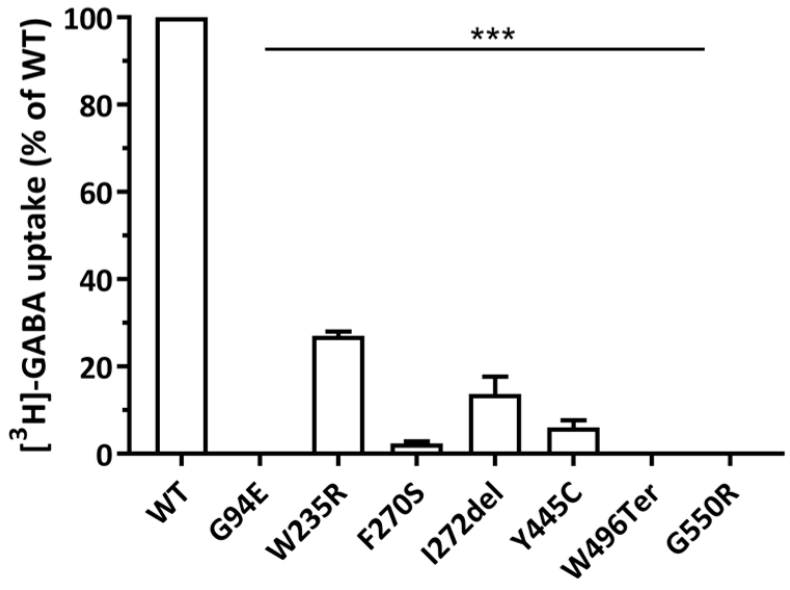


Figure 2.2. *SLC6A1* Epilepsy Variants Reduce GABA Transport. GAT-1 wild-type (WT) and missense variants identified in epilepsy patients were transiently expressed in HeLa cells. Sodium-dependent [^{3}H]-GABA transport was measured at room temperature for 10 minutes. Results represent mean ± S.D. of three separate transfections performed in quadruplicate. Means of variant transport were compared to wild-type (WT) using a one-way ANOVA with a post-hoc Dunnett's multiple comparison test (***$p < 0.001$).

2.4.3 c.850-2A>G Affects the Splicing of *SLC6A1*

The *SLC6A1* c.850-2A>G variant alters the canonical splice acceptor site of intron 8. *In silico* analysis using Human Splicing Finder v3.0 predicted that this variant would eliminate the acceptor site. In HEK293 cells, expression of the c.850-2A>G splice site variant caused exon 9 to be incorrectly spliced out of the RNA transcript (**Figure 2.3A-C**). Sanger sequencing of PCR bands confirmed that the c.850-2A>G variant caused exons 8 and 10 to be spliced together, excluding exon 9 (**Figure 2.3D**). Loss of exon 9 would be predicted to decrease GAT-1 expression through the introduction of a premature stop codon and subsequent NMD of the mutant transcript.

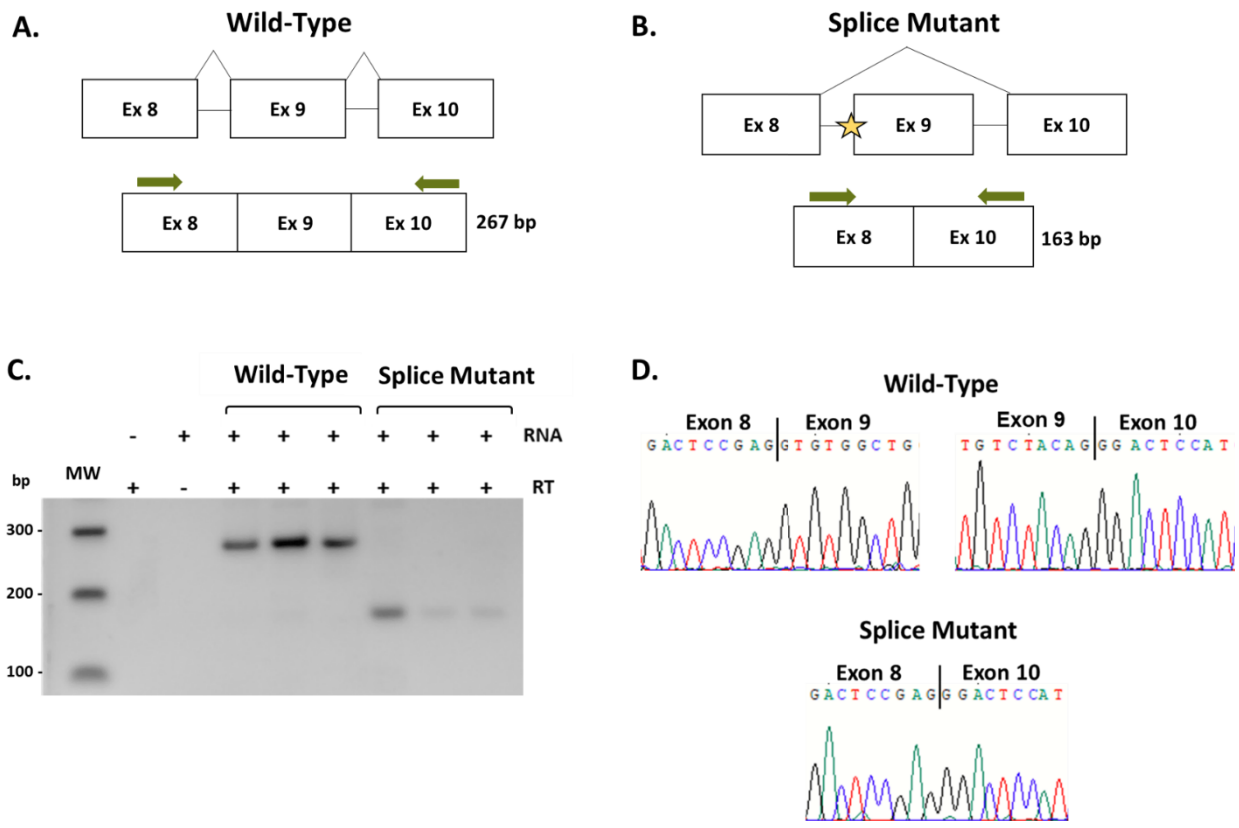


Figure 2.3. c.850-2G>A Results in Exon 9 Skipping. A-B) Schematics showing expected splicing pattern of vectors transfected into HEK293 cells. The expected mRNA structure is indicated below the gene structure. Green arrows indicate the location of primers used to amplify cDNA. **A**) Splicing pattern for wild-type *SLC6A1*. **B**) Predicted splicing pattern for the *SLC6A1* c.850-2A>G splice mutant. The star indicates the location of the c.850-2A>G variant. **C**) Gel image showing splice products. RT indicates the presence (+) or absence (-) of reverse transcriptase during cDNA synthesis. RNA indicates the presence (+) or absence (-) of RNA during cDNA synthesis. **D**) Sanger sequencing traces of gel extracted bands showing exon boundaries.

2.5 Discussion

Variants in *SLC6A1* were first identified in patients presenting with MAE, which is characterized by a range of seizure types including myoclonic, myoclonic-atonic, atonic, and absence seizures. Additionally, patients with MAE also have variable degrees of intellectual disability, developmental delay, and in some cases, autism spectrum disorder, and other behavioral disorders [58, 60-62, 72]. Recently, Johannesen and colleagues showed that variants in *SLC6A1* are more broadly associated with generalized epilepsies, with absence seizures and intellectual disability being common phenotypes [60]. In the present study, we identified eight *SLC6A1* variants in individuals with epilepsy and evaluated their functional effects.

Using a radioactive GABA transport assay, the five missense, one in-frame deletion, and one nonsense variant identified in this study were found to reduce GABA transport (**Figure 2.2**). Three variants (p.G94E, p.W496Ter, and p.G550R) completely abolished GABA transport. The lack of transport activity associated with p.G94E is unsurprising given the role that this glycine residue plays in the bending of TM2 during substrate translocation [77]. Similarly, p.G550R likely leads to defective transport due to destabilization of the TM12 region by the introduction of the positively charged arginine residue. Residual transporter activity ranging from 2% to 27% compared to WT was observed for the p.F270S, p.I272del, p.Y445C, and p.W235R variants. p.F270S and p.I272del are located in the EL3, which connects TM5 and TM6 that are part of the scaffold and core domains, respectively [76]. These domains move relative to each other during substrate translocation, and perturbations of the EL3 linker likely affects this movement [77]. Finally, both p.Y445C and p.W235R involve the replacement of aromatic residues near TM boundaries. Aromatic residues typically stabilize TM domains; therefore, replacement of these residues may affect protein stability [78].

Using a minigene splicing assay, we also confirmed that the c.850-2A>G variant disrupts the canonical splicing of *SLC6A1*, resulting in the exclusion of exon 9 from the mRNA transcript (**Figure 2.3A-B**). This result is predicted to introduce a premature stop codon, which would likely result in nonsense mediated decay and possibly explain the fainter mutant PCR product observed in **Figure 2.3C**.

As an increasing number of *SLC6A1* variants are identified, the ability to quantify the impact of variants on GABA transport will provide the opportunity to explore genotype-phenotype correlations. Based on the clinical information available to us, there were no clear differences in clinical presentation between patients with variants causing complete loss of GABA transport activity and those with residual activity (**Table 2.1**). For example, Patient 2 with the p.W235R variant, which retained 27% of WT activity, presented with intractable absence epilepsy, moderate intellectual disability, and autism spectrum disorder. This presentation was similar to Patient 7, carrying the p.W496Ter variant that completely abolished transport activity.

We hypothesize that reduced GAT-1 function could influence neuronal excitability via multiple mechanisms. GAT-1 dysfunction is expected to reduce GABA clearance, leading to increased GABA levels, both at the synapse and extrasynaptically [79]. Increased GABA levels could lead to the overstimulation of extrasynaptic $GABA_A$ and $GABA_B$ receptors, which are responsible for slow and sustained (tonic) inhibitory responses [80, 81]. Cope and colleagues showed that $GABA_A$-mediated tonic inhibition is increased in thalamocortical neurons in GAT-1 knock-out mice [81]. Increased $GABA_A$-mediated tonic inhibition can lead to neuronal hyperpolarization and burst pattern firing in thalamocortical neurons, which can promote the generation of spike-wave discharges [82]. Similarly, prolonged activation of $GABA_B$ receptors is known to stimulate low voltage-activated (T-type) Ca^{2+} channels, which can cause recurrent excitation within the thalamocortical system through successive Na^+ spikes [80, 83]. Previous studies show that $GABA_B$ receptor activation causes absence seizures in mice and rats and that pretreatment with a $GABA_B$ antagonist can decrease the duration of chemically induced

absence seizures [84, 85]. Reduced GAT-1 function could also decrease the amount of intracellular GABA available for release to activate $GABA_A$-mediated synaptic (phasic) signaling. Decreased $GABA_A$-mediated synaptic signaling is already associated with variants in several $GABA_A$ receptor subunits in genetic epilepsies [86]. Together, these results suggest that reduced GAT-1 function might lead to epilepsy through overactivation of extrasynaptic $GABA_A$ and $GABA_B$ receptors, and reduction in $GABA_A$ synaptic signaling.

In summary, we identified eight *SLC6A1* variants, five of which are novel, in an unselected cohort of 460 epilepsy patients, representing a 1.7% diagnostic yield. Functional analyses of these variants identified reduced GABA transport as a common underlying disease mechanism.

2.6 Acknowledgements

We would like to thank the patients and their families for their participation in this study. We would also like to acknowledge Elizabeth Duffy for her assistance in contacting patients, Lindsey Shapiro for her discussions involving GABA receptors and absence epilepsy, and Deborah Cook for her assistance with editing this manuscript. This study was supported by the following training grant appointments: 5T32GM008490 to K.A.M, K.M.B., and G.S.I. and 5T32NS007480 to G.S.I. The work of the Kanner group was supported by grant 531/15 from the Israel Science Foundation and a grant from the Rosetrees Trust.

Chapter 3

Variants in *ATP6V0C* Associated with Epilepsy Decrease the Function of the Vacuolar ATPase

3.1 Summary

This work started with the identification of a *de novo*, heterozygous missense variant (p.A138P) in *ATP6V0C* in a patient with epilepsy who had trio-WES performed at EGL Genetics. p.A138P is immediately adjacent to p.E139, a necessary residue for the function of ATP6V0C and the larger vacuolar H^+-ATPase (V-ATPase) complex. This single variant was interesting by itself; however, identification of additional patients with variants in *ATP6V0C* would provide further support towards *ATP6V0C* being a disease gene. I used GeneMatcher, an online clinician and researcher "matchmaking" service to identify additional patients with presumed pathogenic *ATP6V0C* variants. Through the combination of GeneMatcher, large sequencing studies of patients with epilepsy and/or neurodevelopmental disorders, and a review of previously published literature, I was able to identify 25 additional patients with neurodevelopmental disorders that had variants in *ATP6V0C.*

Based on previous forward genetic screens in *Saccharomyces cerevisiae*, I was able to develop assays to functionally test a subset of variants located across different regions of *ATP6V0C* in *S. cerevisiae.* These assays utilized the selective requirement for V-ATPase function when yeast is grown in different concentrations of Ca^{2+}*.*

Through collaborations we established over the course of this project, we were able to add additional functional data using *Drosophila* (to answer whether loss of *ATP6V0C* is sufficient to cause seizures) and *C. elegans* (to look at the effect of variants in a more neurologically relevant model system), as well as *in silico* modelling of patient variants using a recently generating cryo-EM structure of the V-ATPase in humans.

This study is the first report of a large cohort of patients with *ATP6V0C* variants and the first report in which patient variants were tested for impact on V-ATPase function. We were able to determine that patient variants reduce the function of the V-ATPase and that missense variants likely act through a dominant-negative mechanism in which they are predicted to interfere with the rotary mechanism of the V-ATPase required for proton translocation. Further

studies are needed to 1) follow this cohort of patients longitudinally to determine the natural course of disease and 2) better evaluate the exact mechanism behind the neurodevelopmental phenotypes seen in patients. Discussion of these areas is provided in **Chapter 5**.

3.2 Introduction

The V-ATPase is a highly conserved enzymatic complex that functions in an ATP-dependent manner to pump protons across membranes and acidify organelles. The V-ATPase is comprised of a peripheral V_1 domain and an integral V_0 domain (**Figure 3.1**). The V_1 domain is responsible for the hydrolysis of ATP creating the necessary energy to translocate protons through the V_0 domain via a rotational mechanism [87]. The V-ATPase plays a crucial role in many cellular processes involving membrane trafficking by creating a proton/pH gradient used by several different types of transporters [87, 88].

The human V-ATPase, comprised of 13 different subunits, is encoded by 23 genes (**Table A1**). This genetic redundancy allows for the formation of tissue-specific V-ATPase complexes, including in synaptic vesicles (SVs) where it creates the necessary proton/pH gradient to load various neurotransmitters [87, 88]. To date, twelve genes corresponding to eight different subunits of the V-ATPase have been associated with human disease (**Table A1**) [89-99]. Early-onset epilepsy has been observed in patients with variants in either *ATP6V1A* or *ATP6V0A1*, with heterozygous, *de novo* variants leading to less severe presentations when compared to patients with biallelic variants [100-102]. Pathogenic variants in *ATP6V1B2* can cause epileptic conditions such as Zimmerman-Laband (MIM#616455) and DOORS syndromes, or deafness and nail dysplasia without epilepsy (DDOD, MIM#124480) [97, 98, 103, 104]. An accessory protein to the V-ATPase, encoded by *ATP6AP2*, is associated with X-linked syndromic intellectual disability that can present with or without epilepsy (MIM#300423) [105, 106].

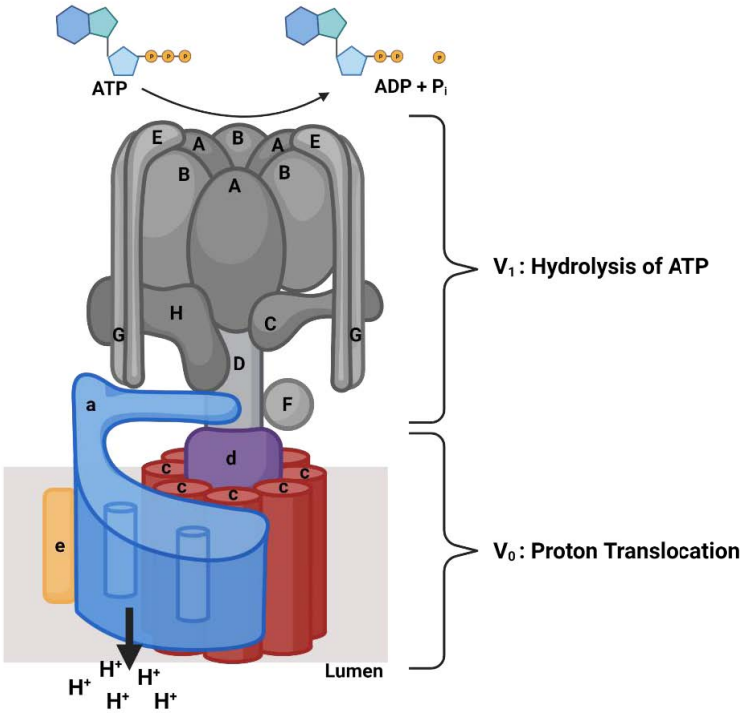


Figure 3.1. V-ATPase Structure. The peripheral domain (V_1, uppercase letters, in grey) is the site of ATP binding and hydrolysis. The integral domain (V_0, lowercase letters, in purple, red, blue, and yellow) transports protons across membranes. The c-ring (red) is composed of 9 c-subunits (encoded by *ATP6V0C*) and 1 c"-subunit (encoded by *ATP6V0B*, not shown) and rotates after ATP hydrolysis to bring protons to ATP6V0A (blue). ATP6V0A possess two hemi-channels and a buried arginine residue (p.R735) which are required along with p.E139 in ATP6V0C for proton translocation [87].

ATP6V0C, a three-exon gene (**Figure 3.2A**) located on human chromosome 16p13.3, encodes the 155 amino acid c-subunit of the V_0 domain which along with the c'' subunit (encoded by *ATP6V0B*) forms the intramembrane c-ring that facilitates the movement of protons across the membrane (**Figure 3.1**) [107]. The process of proton translocation is reliant on a glutamate residue at position 139 (p.E139) in ATP6V0C as well as an arginine residue (p.R735) in ATP6V0A [87].

We previously described patients with developmental delay, intellectual disability, microcephaly, and seizures with 16p13.3 microdeletions encompassing a minimal overlapping region that included *TBC1D24, ATP6V0C,* and *PDPK1* [108]. By reviewing the known function(s) and expression patterns of genes in the minimal overlapping region, we proposed haploinsufficiency of *ATP6V0C* as the primary contributor to the clinical features of 16p13.3 microdeletion syndrome [109]. However, we did not provide any functional evidence to support our claim.

Most recently, Ittiwut *et al.* reported a *de novo* stop-loss variant in *ATP6V0C* in an individual with epilepsy and intellectual disability [110]. Analysis of RNA derived from the patient's leukocytes revealed that, as expected, the mutant transcript escaped nonsense mediated decay (NMD). The authors proposed haploinsufficiency as the likely pathomechanism given the observed decrease in mRNA levels; however, a dominant negative effect is also possible given the transcript escape from NMD. Hence, thus far it is unknown if *ATP6V0C* missense variants are a cause of human disease, and the mechanistic basis of *ATP6V0C*-associated human disease is unclear.

In this study, we report the identification of heterozygous *ATP6V0C* missense variants in patients with a novel syndrome of developmental delay, epilepsy, and intellectual disability. We present multiple lines of computational and functional analyses to demonstrate that these variants disrupt V-ATPase activity.

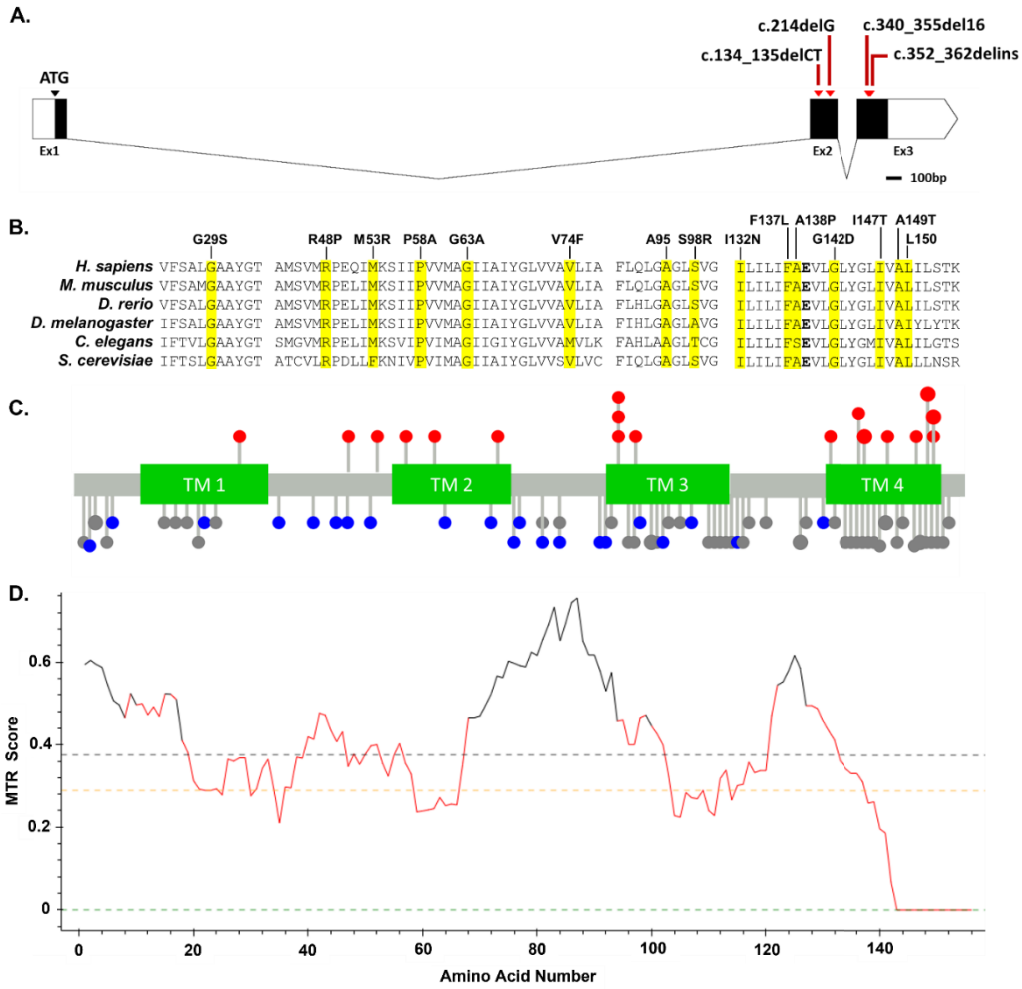


Figure 3.2. Location and Conservation of ATP6V0C Variants. A) Exon/intron structure of *ATP6V0C*. Boxes represent exons with black denoting coding regions. **B**) Protein alignment showing conservation of affected residues (yellow). Glutamate residue (p.E139) required for proton transport is bolded. The following protein sequences were used in the alignments: *H. sapiens*, NP_001685.1; *M. musculus*, NP_001348461.1; *D. rerio*, NP_991117.7; *D. melanogaster*, NP_476801.1; *C. elegans,* NP_499166.1; *S. cerevisiae,* NP_010887.3. **C**) Lollipop plot showing the transmembrane structure (green) and location of variants throughout ATP6V0C. Patient missense variants are indicated with red. Missense (blue) and synonymous (grey) variants observed in gnomAD are shown. Based on UniProt accession P27449. There is a significant enrichment of patient variants in TM4 (p = 0.006, Fisher's Exact test). **D**) The missense tolerance ratio (MTR) for ATP6V0C was calculated using 21 codon window sizes. An MTR score of < 1 indicates intolerance to missense variation. Dashed lines on the plot denote ATP6V0C-specific MTRs: green = 5th percentile, yellow = 25th percentile, and black = 50th percentile.

3.3 Materials and Methods

Identification of Individuals with ATP6V0C Variants

Informed consent was obtained through protocols approved by the institutional review boards at each site of patient recruitment. Patients with *ATP6V0C* variants were identified via GeneMatcher, and by interrogating the 100,000 Genomes database, the Deciphering Developmental Disorders (DDD) study, and ClinVar [46, 111, 112]. We also screened whole-exome sequencing (WES) data from epilepsy patients referred for genetic testing at EGL Genetics. Three patients (Patients 2, 13 and 26) were previously reported in other publications [110, 111, 113]. For Patients 2 and 26, clinical information was obtained from the previous publications, whereas the referring clinician provided clinical information for Patient 13. Patient 5 was reported in ClinVar, and clinical information was provided by the depositing organization (**Table A2**). Deidentified clinical information for all other patients was collected using a custom form provided to each site.

Sequencing and Analysis of Sequence Data

Whole exome or whole genome sequencing was performed on patient DNA extracted through standard protocols. All libraries were sequenced on Illumina HiSeq systems. Sequence alignment and variant calling were performed at each site and further details are provided in **Table A2**. When possible, *ATP6V0C* variant segregation was confirmed with Sanger sequencing using standard protocols.

Lollipop and Missense Tolerance Ratio Diagram

The Lollipop diagram was created as previously described using the UniProt accession number, P27449 [114, 115]. Non-synonymous and synonymous population variants were downloaded from gnomAD (v.2.1.1) [116]. Resulting diagrams of gnomAD and patient variants were merged into a single image for ease of visualization, and the locations of the transmembrane domains were superimposed over the resulting image. An MTR plot was

generated using MTR-Viewer v2, with a codon window size of 21, on the ENST00000330398 transcript [117].

Drosophila Studies

The Drosophila ortholog, Dmel\Vha16-3 (CG32090), was knocked-down using pan-neuronal (elaV-GAL4) expression of a gene-specific RNAi construct (102067), provided by the Vienna Drosophila Resource Center. As a control, an RNAi to GFP was used. Seizure activity was elicited using electroshock of wall-climbing third instar larvae as described in Marley and Baines 2011 [118]. Drugs, solubilized in DMSO, were fed to larvae by mixing (at 2mM) in molten fly food which was then allowed to cool and set before being seeded with first instar larvae.

In silico Variant Modeling

Patient and gnomAD variants were displayed in the context of a structure for the V_0 domain of human V-ATPase (PDB: 6wlw) [107, 119]. PyMOL and Swiss PDB Viewer were used for visualization of protein structure [120].

Saccharomyces cerevisiae Strains and Plasmids

E. coli and yeast manipulations were performed following standard molecular biology protocols [121]. The *vma3::kanMX* yeast strain (*vma3Δ,* Cat No. YSC6273-201929081) was obtained from GE Dharmacon and is isogenic with BY4741 (*MATa his3Δ leu2Δ met15Δ ura3Δ*). Plasmids are listed in **Table A3**. A plasmid (pKM16) containing the promoter and wild type open reading frame for *VMA3* (the yeast ortholog of *ATP6V0C*) was generated by amplifying a 924bp fragment from *S. cerevisiae* gDNA and cloning it into pRS316 using *BamHI* and *SacI* (Forward cloning primer: 5' taagcaggatccagcaatgaaataggccgtctac, Reverse cloning primer: 5' taagcagagctccttgaaatgaggtagtttgg) [122]. pKM16 was used as the backbone to generate all variants via conventional cloning techniques or the QuikChange XL Site-Directed Mutagenesis Kit (Agilent Technologies). Sanger sequencing was performed to confirm the presence of each variant as well as the absence of unwanted substitutions.

Plasmids were transformed into the *vma3Δ* strain and selected on plates containing synthetic minimal media plus dextrose supplemented to select for a *URA3* plasmid (SD-ura). Selected transformants were maintained on SD-ura throughout the course of experiments.

Serial Dilution Spotting Assay

Liquid cultures of transformants were grown at 30°C overnight in SD-ura adjusted to a pH of 5.5. 1 OD_{600} per ml of cells was collect and suspended in dH_20. Serial 10-fold dilutions were spotted onto SD-ura plates. Plates were imaged after incubation at 30°C for 48 hours.

LysoSensor Uptake and Confocal Imaging

Cells were collected at 0.6-1.0 OD_{600} and incubated with LysoSensor Green DND-189 (Invitrogen, Cat# L7535) as previously described [123]. Cells were resuspended in 1X PBS (pH 7.6) to an OD_{600} of 0.6, deposited on 1.5% agarose pads, and visualized immediately. Images were taken at room temperature using a confocal laser scanning microscope (A1R HD25, Nikon) with an Apo TIRF 60x 1.49 NA oil immersion lens, WD 0.12mm (Nikon). Images were acquired using NIS-Elements (AR 5.21.02, Nikon) and processed using FIJI. All cells visible in the DIC channel were selected, the selection was copied to the FITC channel, and the mean grey area was calculated for each cell. Measurements were corrected for background signal by subtracting the mean grey area of a background only selection from each image. Two biological replicates totaling 71-132 cells for each variant were analyzed.

Generation of Growth Curves

Liquid cultures of transformants were grown at 30°C overnight in SD-ura adjusted to a pH of 5.5. Overnight cultures were diluted 1:20 in dH_20 and further diluted 1:10 in YPD, pH 5.5, with 5mM, 100mM or 200mM $CaCl_2$. OD_{600} measurements were taken every 15 minutes for 30 hours at 30°C with shaking using an ELx808 plate reader (BioTek). Three independent transformants were assayed in triplicate for each variant. The R package, GrowthCurver, was used to calculate the empirical area under the curve (eAUC) from the OD_{600} measurements for each replicate [124].

Caenorhabditis elegans Studies

All *C. elegans* strains were cultured and handled as per standard methods. All experiments were carried out at 20°C. Mutations were knocked in via CRISPR/Cas9, and homozygous worms were studied. Strains used in this project are described in **Table A4**.

Paralysis assays

In three separate experiments, 30-40 L4 worms (in triplicate) were picked to standard NGM plates either with 51mM NaCl (physiological conditions), or with 200mM or 300mM NaCl, and scored daily for paralysis starting the day after they were picked. A total of 270-360 worms were tested per condition. Worms were counted as paralyzed if they failed to move their body upon prodding with a platinum wire and were considered dead if they failed to move their head and showed no pharyngeal pumping when prodded. Dead or lost animals were censored from statistical analyses. Worms were transferred every 2 days to avoid progeny.

Lifespan assays

Lifespan experiments were conducted similarly to paralysis assays. Two separate experiments were performed using 35-40 worms in triplicate, leading to a total 315-360 worms tested per condition. Worms were counted every second day from day 1 of adulthood until death. Lost animals were censored from statistical analyses, paralyzed worms were not censored and were kept until death.

Liquid culture motility assays

Synchronized day 1 adult worms were transferred into 200uL of M9 buffer with or without 350mM NaCl in each well of a 96 well plate, to a density of 30 worms per well. Motility was automatically analyzed for 4 hours using a WMicrotracker-One plate reader (PhylumTech).

WormLab analysis

Synchronized day 1 adult worms were video recorded for 30-33 seconds using a Leica Stereomicroscope S9i. Automated movement and worm size analyses were conducted using WormLab software (MBF Biosciences) which tracks individual worms from the recorded videos.

Activity index is defined by the brush stroke (area "painted" by the animal's body in a single complete stroke) normalized by the time taken to perform two strokes. Wave initiation rate is defined as the number of body waves initiated from either the head or the tail per minute. Swimming speed was measured over a two-stroke interval [125]. For body size analysis, worms were recorded on bacteria free NGM plates. For swimming parameters, worms were placed in M9 with 500mM NaCl and recorded 30 minutes later. At least 50 worms were recorded in two independent experiments.

Aldicarb sensitivity assays

Worms were grown on standard NGM plates until day 1 of adulthood and then transferred to NGM plates containing 1mM aldicarb. Paralysis was assayed every 30 minutes for 2 hours. Worms were counted as paralyzed if they failed to move upon prodding with a platinum wire.

Statistical Analyses

Drosophila recovery time was analyzed using one-way ANOVA with Dunnett's post-hoc test for multiple comparisons. Recovery time after drug treatment was normalized to the vehicle only control. A Fisher's Exact test was used to demonstrate the presence of a variant hot-spot in the fourth transmembrane domain of ATP6V0C. For LysoSensor, fluorescence values for each cell measured were normalized to the mean fluorescence of the wild type rescue. For the growth rate assay, eAUC values were normalized to the mean eAUC of the wild type rescue within each plate. A one-sample t-test was used to compare the mean fluorescence and eAUC for each variant to a hypothetical mean of 100 (representing the wild type rescue). Significance levels were corrected for multiple comparisons using a Bonferroni correction (LysoSensor and 5mM, $\alpha = 0.0003125$; 100mM, $\alpha = 0.00714$; 200mM, $\alpha = 0.00833$). A two sample t-test (two-tailed) was used to compared p.G103S to p.F137L at 200mM. Paralysis curves and lifespan assays were compared using the log-rank (Mantel–Cox) test. The liquid culture motility assays (WormTracker) results were analyzed using two-way ANOVA to compare each variant to the

wild type N2 strain. WormLab results were analyzed using a one-way ANOVA with a Dunnett's post-hoc test for multiple comparisons to compare each variant to the wild type N2 strain. The WormLab data is presented as box and whisker plots indicating minimal and maximal data points. Normalization and statistical analyses were carried out using Prism 9.0 (GraphPad Software Inc.). All data is presented as mean ± SEM and $\alpha = 0.05$ was used unless otherwise noted.

3.4 Results

3.4.1 Identification of *ATP6V0C* Variants in Patients

We identified 26 patients with heterozygous *ATP6V0C* variants through the use of GeneMatcher, 100,000 Genomes database, the DDD study, ClinVar, published literature, and EGL Genetics. Of the 26 patients, 21 had missense substitutions (18 variants were unique), 4 had frameshifting variants, and 1 had a stop-loss variant (**Table 3.1**). The variants p.A138P, p.A149T, and p.L150F were recurrent, each being seen in two unrelated individuals. Different substitutions at p.A95 and p.L150 were also observed. Patients 4, 15 and 19 were each found to be mosaic for their identified *ATP6V0C* variant.

Multiple lines of evidence support the pathogenicity of the identified patient variants. Firstly, all patient variants were absent from gnomAD (v.2.1.1). Secondly, in 23 patients where biologic parent DNA was available, the variants were shown to have occurred *de novo*. Thirdly, all missense substitutions affected highly conserved residues (**Figure 3.2B**), with 17 out of the 18 unique missense variants having CADD scores that placed them in the top 1% of predicted deleterious variants. Lastly, *ATP6V0C* has a predicted intolerance to missense and loss of function variation with only 21 population missense variants observed in gnomAD compared to the expectation of 108 (observed/expected = 0.19) and zero loss of function variants observed compared to the expectation of 4.5 (**Figure 3.2C-D)**. In addition, all *ATP6V0C* gnomAD variants

are observed at low frequencies (three times or less) [126]. Taken together, when the ACMG/AMP variant classification guidelines are applied, all *ATP6V0C* variants identified in the patients are predicted to be likely pathogenic or pathogenic (**Table A2**) [43].

Table 3.1 (next page) Clinical Presentation of Patients with *ATP6V0C* Variants
[a]Based on reference sequence NM_001694.4.
[b]Scores obtained using CADD GRCh37-v1.6.
[c]Intellectual Disability can usually be first assessed at 5 years of age
[d]Patients 1, 12, and 23 have severe neurodevelopmental diseases but detailed clinical information was unavailable.
[e]Previously published as Patient T1911 in Carvill *et al.* [113]
[f]ClinVar Variant: VCV000870676
[g]Previously published as DDD4K.04123 in DDD study [111]
[h]Previously published in Ittiwut *et al.* [110]
Abbreviations: Ab = absence, At = atonic, ASD = autism spectrum disorder, FDS = focal dyscognitive seizures, FOA = Focal onset aware (Partial), FIA = Focal impaired aware, Myo = Myoclonic, NA = not available, T = Tonic, TCS = Tonic-clonic seizures

Patient	Variant[a]	CADD Score[b]	Inheritance	Seizures (Age at Onset)	Seizure Types	Developmental Delay	Intellectual Disability[c]	Behavior
1[d]	c.85G>A; p.G29S	26.2	*de novo*	NA	NA	NA	NA	NA
2[e]	c.134_135delCT; p.(S45CfsTer37)	NA	*de novo*	Yes (7mo)	GTCS, At, FDS, Myo, T	NA	Severe, with regression	NA
3	c.143G>C; p.R48P	27.2	*de novo*	Yes (18mo)	NA	Yes	NA	ASD
4	c.158T>G; p.M53R	25.9	*de novo*, mosaic	No	NA	Yes	Too Young for Evaluation	Tantrums, communication related
5[f]	c.172C>G; p.P58A	23.1	NA	Yes	Cryptogenic focal	Yes, psychomotor	NA	NA
6	c.188G>C; p.G63A	23.9	*de novo*	Yes (8mo)	Infantile spasms, GTCS, At Myo	Yes, non-verbal	Severe	Head banging, hair pulling
7	c.214delG; p.(V72WfsTer9)	NA	*de novo*	Yes	Infantile spasms	Yes	Too Young for Evaluation	Too Young for Evaluation
8	c.220G>T; p.V74F	26	*de novo*	Yes (12mo)	GTCS, Ab, FOA	Yes, motor and speech	Severe	ASD, ADHD, food seeking
9	c.283G>A; p.A95T	26.2	NA	Yes (10mo)	GTCS, FOA, Febrile	Yes, regression to non-verbal	Profound	ASD
10	c.283G>C; p.A95P	24.6	*de novo*	Yes (10mo)	GTCS, staring	Yes, motor and non-verbal	Too Young for Evaluation	Picky eater
11	c.284C>T; p.A95V	23.7	*de novo*	Yes (5mo)	Febrile, Ab, Myo, T (nocturnal)	No	Moderate	ASD, ADHD, aggression
12[d]	c.294C>A; p.S98R	13.2	*de novo*	NA	NA	NA	NA	NA
13[g]	c.340_355del16; p.(D115AfsTer12)	NA	*de novo*	Yes (16mo)	Focal with secondary generalization	Yes, speaks only in short sentences	Yes	ASD
14	c.352_362delins; p.(V118HfsTer19)	NA	*de novo*	Yes (30mo)	GTCS, focal to bilateral TCS, Ab, FIA	Yes, motor and speech	Mild, regression in adulthood	ASD, psychoses (>30yrs)
15	c.395T>A; p.I132N	25.3	*de novo*, mosaic	Yes (12yr)	NA	Yes	Yes	Tantrums
16	c.409T>C; p.F137L	25.2	*de novo*	Yes (13mo)	GTCS, Myo, At, FOA	Yes, motor and non-verbal	Profound	Stereotypies
17	c.412G>C; p.A138P	25.5	*de novo*	NA	TCS	NA	NA	NA
18	c.412G>C; p.A138P	25.5	*de novo*	Yes (6mo)	GTCS, multifocal	Yes, motor and speech	Too Young for Evaluation	None
19	c.425G>A; p.G142D	24.7	*de novo*, mosaic	No	NA	Yes, motor and non-verbal	Too Young for Evaluation	Screams to indicate wants
20	c.440T>C; p.I147T	23.3	*de novo*	Yes	NA	Yes, speech	Profound	Autistic traits
21	c.445G>A; p.A149T	24.3	*de novo*	Yes (38mo)	Febrile, TCS, Myo, Ab	Yes, motor and speech	Mild	None
22	c.445G>A; p.A149T	24.3	*de novo*	Yes (6mo)	GTCS, At, TCS	Yes, motor and speech	Mild	ADHD
23[d]	c.448C>T; p.L150F	25.1	NA	NA	NA	NA	NA	NA
24	c.448C>T; p.L150F	25.1	*de novo*	Yes (6mo)	Infantile flexor spasms, T (w/ asymmetrical limb stiffening)	Yes, motor and speech	Profound	ASD-like sensory, self-biting and scratching
25	c.449T>C; p.L150P	24.7	*de novo*	Yes (12mo)	Febrile, GTCS, T, At, Ab	Yes, fine motor	Mild	Autistic traits
26[g]	c.467A>T; p.(Ter156LeuextTer35)	NA	*de novo*	Yes (24mo)	GTCS, TCS, At, afebrile	No	Yes, with regression	NA

3.4.2 *ATP6V0C* Variants Cause a Human Syndrome of Developmental Delay, Epilepsy, and Intellectual Disability

The primary clinical presentation of the identified patients was developmental delay with early-onset epilepsy and intellectual disability (**Table 3.1**). The mean age of seizure onset was 20.72 ± 7.56 months, with 14 of 17 patients for whom this information was available having onset prior to 24 months. Based on clinical information from 18 patients, the most common seizure types observed were generalized tonic-clonic (11/18), focal (7/18), atonic (6/18), and myoclonic (5/18). Intellectual disability, ranging from mild to severe, was seen in 15/15 patients who were old enough for a formal diagnosis and for whom this information was available. Development delay was seen in 19/21 patients. Nineteen patients had MRIs with thirteen showing abnormalities (**Table A2**). Common findings in MRIs included agenesis/hypoplasia of the corpus callosum (**Figure A1**) or cerebellar vermis (Patients 5, 6, 18, 20 and 25), and delayed myelination (Patients 6, 18, 22 and 24). Four patients were reported to have cardiac abnormalities: Patient 3 had pulmonary valve stenosis, Patient 6 had a thickened left ventricular wall, Patient 7 had a heart murmur, and Patient 13 exhibited several cardiac defects including hypertrophic cardiomyopathy, mitral valve prolapse, and mild to moderate mitral valve regurgitation. Patients 5 and 24 both showed dental enamel defects, with Patient 24 lacking dental enamel. It should be noted that Patient 7 also has a *de novo* 2.3Mb deletion in 20q11.22-11.23 which has been associated with developmental delay and dysmorphism, and Patient 13 has biallelic variants in *LZTR1* which has been associated with Noonan-like syndrome (MIM#605375) [127].

Collectively, these data show that *ATP6V0C* variants cause a human syndrome of developmental delay, intellectual disability, and epilepsy. Furthermore, as most individuals were ascertained on the basis of genotype (i.e., having a variant in *ATP6V0C*), their phenotypic

convergence on reverse phenotyping further supports the pathogenicity of the variants described in this study [128].

3.4.3 *ATP6V0C* Knockdown in Drosophila Results in Seizure Phenotype

Based on our hypothesis that haploinsufficiency of *ATP6V0C* drives the neurological phenotype of 16p13.3 microdeletion syndrome and identification of patients with frameshift variants, we first tested the consequences of ATP6V0C knockdown in Drosophila [109]. The orthologous protein in Drosophila, Dmel\Vha16-3, (CG32090), shows 78% amino acid identity to ATP6V0C. CG32090 was knocked-down via pan-neuronal expression of a gene-specific RNAi construct (VDRC-102067). Controls of the same pan-neuronal driver line (elaV-GAL4) driving expression of GFP RNAi (elaV>GFP RNAi), and the homozygous 102067 RNAi line without the elaV-GAL4 driver (102067 RNAi) were used. Following knock-down of CG32090 (elaV>102067 RNAi), wall climbing third instar larvae showed a significant increase in recovery time (i.e., longer seizure duration) following electroshock-induced seizure ($p < 0.0001$, one-way ANOVA, **Figure 3.3A**). Pre-treatment of larvae in which CG32090 was knocked down (elaV>102067 RNAi) with a variety of established antiepileptic drugs show that significant reduction of recovery time (i.e., reduction in seizure duration) was achieved with levetiracetam and topiramate ($p < 0.001$), and to a lesser extent with lamotrigine and valproate ($p < 0.01$ and < 0.05, respectively, **Figure 3.3B**). Phenytoin, at the same concentration (2mM in fly food) did not significantly alter recovery time. These results are consistent with the hypothesis that haploinsufficiency of *ATP6V0C* contributes to seizures.

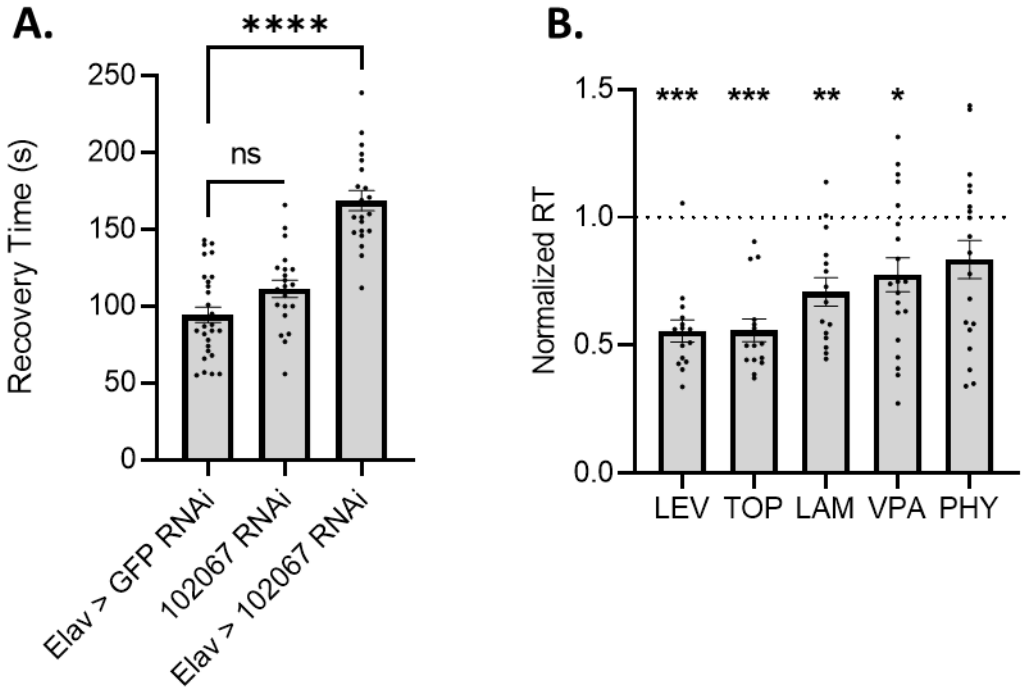


Figure 3.3. Knockdown of the *Drosophila* Ortholog of *ATP6V0C* Increases Seizure Duration. A) Pan-neuronal (elaV-GAL4) RNAi-mediated knockdown of Dmel\Vha16-3 (CG32090) using RNAi (elaV>102067 RNAi) is sufficient to increase the recovery time (RT) of third instar larvae to electroshock-induced seizure. Controls expressed GFP RNAi via elaV-GAL4 (elaV>GFP RNAi) or the RNAi (102067) without a driver (102067 RNAi). **B**) Seizure-induction due to expression of 102067 RNAi is preferentially rescued by pre-treatment of larvae with levetiracetam (LEV) or topiramate (TOP). Lamotrigine (LAM) and valproate (VAL) were also effective, but not phenytoin (PHY). RT was normalized to a vehicle (DMSO) only control. Data shown as mean ± SEM. *p* = ****<0.0001; ***<0.001, **<0.01; *<0.05, One-way ANOVA with post-hoc comparison (Dunnett's).

3.4.4 *ATP6V0C* Variants Are Predicted to Interfere with V-ATPase Rotary Mechanism

Higher conservation across ATP6V0C orthologs was seen at sites of patient variants compared to gnomAD variants (**Figure A2**). To understand the basis of pathogenicity of the *ATP6V0C* missense variants, we first turned to *in silico* modelling. Upon hydrolysis of ATP, the c-ring (comprised of nine copies of ATP6V0C and one copy of ATP6V0B) rotates within the membrane delivering protons to the ATP6V0A subunit (encoded by *ATP6V0A*) for transport across the membrane (**Figure 3.1**) [107]. Transmembrane (TM) domains 2 and 4 of ATP6V0C are outward facing and interact with ATP6V0A during this rotational mechanism [129, 130]. The location of patient variants shows an enrichment within TM domains and the presence of a 'hot-spot' in the fourth TM of ATP6V0C ($p = 0.006$, Fisher's Exact test, **Figure 3.2C**). When viewed structurally, some patient and gnomAD variants are located at sites of packing between c-ring subunits; however, many more patient variants are located outward-facing from the c-ring so as to potentially interfere with interactions between mutant ATP6V0C subunits and the ATP6V0A subunit (**Figure 3.4A-B**). These data suggest a dominant negative effect as the likely underlying mechanism for most missense variants in *ATP6V0C*.

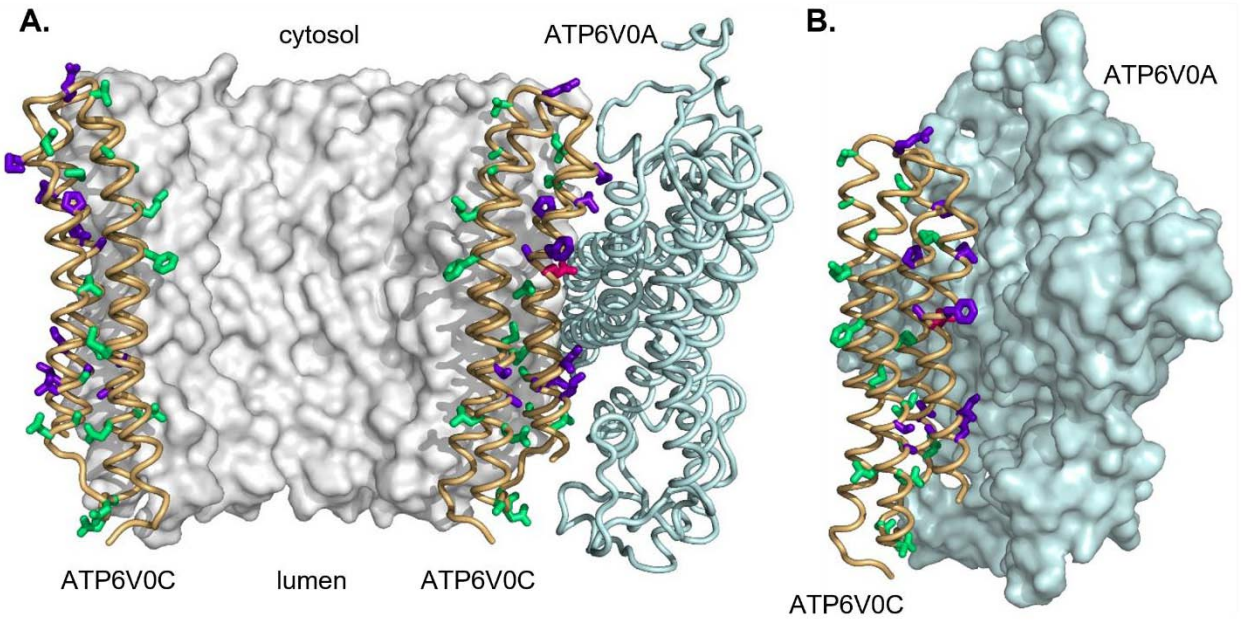


Figure 3.4. Molecular Modeling of Patient and gnomAD Variants. A) Structure of part of the V_0 region of human V-ATPase (PDB: 6wlw) [107]. Sites of patient (purple) and gnomAD (green) variants are shown superposed on ribbon backbone for two ATP6V0C subunits (gold), one next to the ATP6V0A subunit (cyan) and one on the opposite side of the c-ring. The back part of the c-ring is filled with grey, and the front part has been omitted for clarity. **B**) Isolated view of the interaction between ATP6V0C variants and ATP6V0A. The functional amino acid p.E139 is also displayed (pink).

3.4.5 *ATP6V0C* Patient Variants are Deleterious in Yeast

Budding yeast, *S. cerevisiae,* possesses an *ATP6V0C* ortholog, *VMA3,* which shares a 72% amino acid identity and a conserved four transmembrane protein structure. *S. cerevisiae* has been previously used to study the functional effects of variants in other V-ATPase subunits [94, 131-134]. Given that all identified patient missense variants affected residues that are conserved between human and yeast, we expressed 12 of the patient variants in *VMA3* using a centromeric plasmid in a *vma3Δ* yeast strain (**Figure A2, Table A3**). Six additional variants were identified after completion of these experiments and, therefore, were not modeled in yeast.

We also examined the functional effects of three population variants in *ATP6V0C* from gnomAD (p.R48W, p.G103S, p.M131I; **Table 3.2, Figure A2**) on V-ATPase function. p.R48W was chosen as it effects the same residue as the p.R48P patient variant. p.G103S and p.M131I were chosen as they have the highest CADD scores for variants seen twice and once, respectively, in gnomAD. The altered residues are also conserved between human and yeast (**Figure A2**). All *ATP6V0C* variants in gnomAD (21 total) are rare (**Figure A2**), being seen no more than three times out of approximately 250,000 alleles. In addition, we also generated and tested p.E139A, which removes the glutamate residue necessary for V-ATPase function [87].

gnomAD Variant[a]	c.142C>T; p.R48W	c.307G>A; p.G103S	c.393G>C; p.M131I
gnomAD count[b]	1	2	1
CADD Score[c]	21.2	24	23
Poly-Phen2[c]	Benign (0.235)	Probably Damaging (0.957)	Benign (0.168)
SIFT[c]	Deleterious (0.04)	Deleterious (0.05)	Deleterious (0.04)
Location	CL1	TM3	TM4

Table 3.2. gnomAD Variants Assessed in Yeast.
[a]Based on reference sequence NM_001694.4.
[b]Number of times observed in gnomAD (v.2.1.1) out of approximately 250,000 alleles.
[c]CADD, Poly-Phen2 and SIFT scores obtained from CADD GRCh37-v1.6
Abbreviations: CL = cytoplasmic loop, TM = Transmembrane domain

To first establish the ability of the *vma3Δ* strain to grow when transformed with plasmids containing patient or population variants, a serial dilution spot assay was performed on SD-ura plates. We confirmed the ability of all transformants to grow under no selective V-ATPase pressure, thereby allowing us to examine V-ATPase function in our yeast model (**Figure 3.5**).

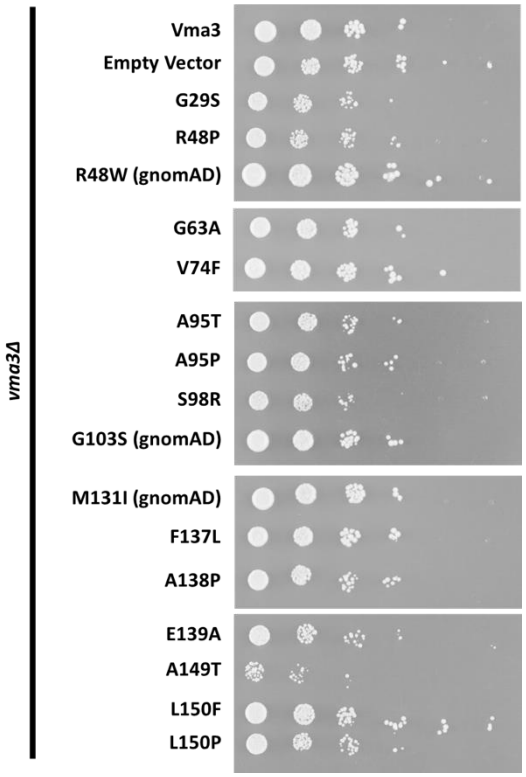


Figure 3.5. Transformants can Grow in Permissive Conditions. Liquid cultures of the *vma3Δ* strain transformed with a wild-type plasmid or a plasmid expressing a patient variant were serially diluted and spotted onto SD-ura, pH 5.5 and grown at 30°C for 48hrs.

LysoSensor Green DND-189 is an acidotropic dye that accumulates in the membranes of vacuoles. Upon protonation, quenching of the fluorescent probe is relieved and fluorescent intensity increases in a pH dependent manner. It has been previously demonstrated that V-ATPase activity in yeast and fluorescent intensity of LysoSensor Green are correlated [123]. We looked for rescue of V-ATPase function by transformation of each patient or population variant

into the *vma3Δ* strain. Nine patient variants and p.E139A resulted in little to no fluorescence (**Figure 3.6**). Two patient variants (p.G63A and p.L150F) and two gnomAD variants (p.R48W and p.G103S) showed intermediate levels of fluorescence intensity compared to the wild type rescue. One gnomAD variant (p.M131I) and one patient variant (p.F137L) show levels of fluorescence intensity that were comparable to wild type rescue. Overall, 10 of 12 patient variants and 2 of 3 gnomAD variants elicited significant decreases in fluorescence intensity when compared to the wild type rescue ($p < 0.0001$, one sample t-test, **Figure 3.6**).

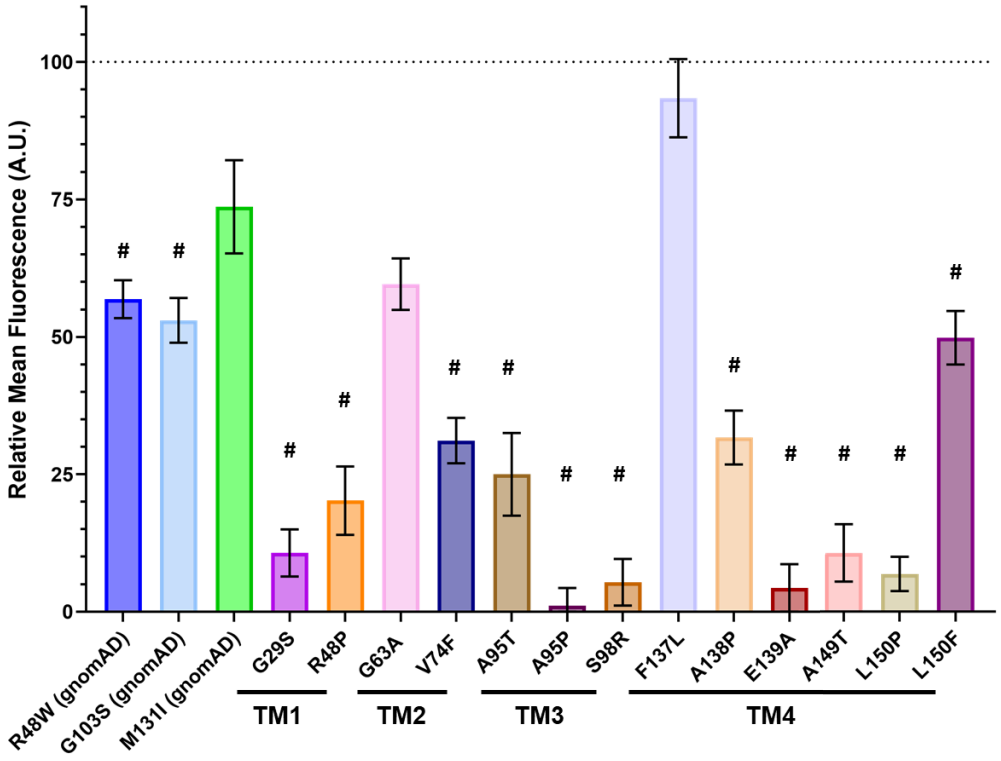


Figure 3.6. Patient Variants Show Decreased LysoSensor Fluorescence. Quantification of average fluorescent intensity for each variant in the LysoSensor assay. Variants are grouped based on their location within or proximity to the nearest TM domain. Data was normalized with mean of wild type as 100% (denoted by dotted line) and mean of empty vector as 0%. Data shown as mean ± SEM (n = 71-132 cells/variant). A one sample t-test to a hypothetical mean of 100 was conducted with a Bonferroni correction (adjusted α level = 0.0003125). $^{\#}p < 0.0001$.

V-ATPase, and therefore ATP6V0C, function is required for yeast to grow at increased $CaCl_2$ concentrations [131, 135]. To further examine the impact of patient *ATP6V0C* variants on V-ATPase function, transformants with patient or population variants were used to inoculate YPD with 5mM $CaCl_2$ and growth curves were generated for each variant (**Figure 3.7A-D**). Eight patient variants and p.E139A showed negligible growth at 5mM $CaCl_2$. The remaining four patient variants and the three gnomAD variants showed varying degrees of growth at 5mM $CaCl_2$. Overall, significantly less growth compared to wild type rescue was seen for 11 of 12 patient and the three gnomAD variants tested ($p < 0.0001$, one sample t-test, **Figure 3.7E**). The results seen at 5mM $CaCl_2$ mirrored those seen with LysoSensor uptake.

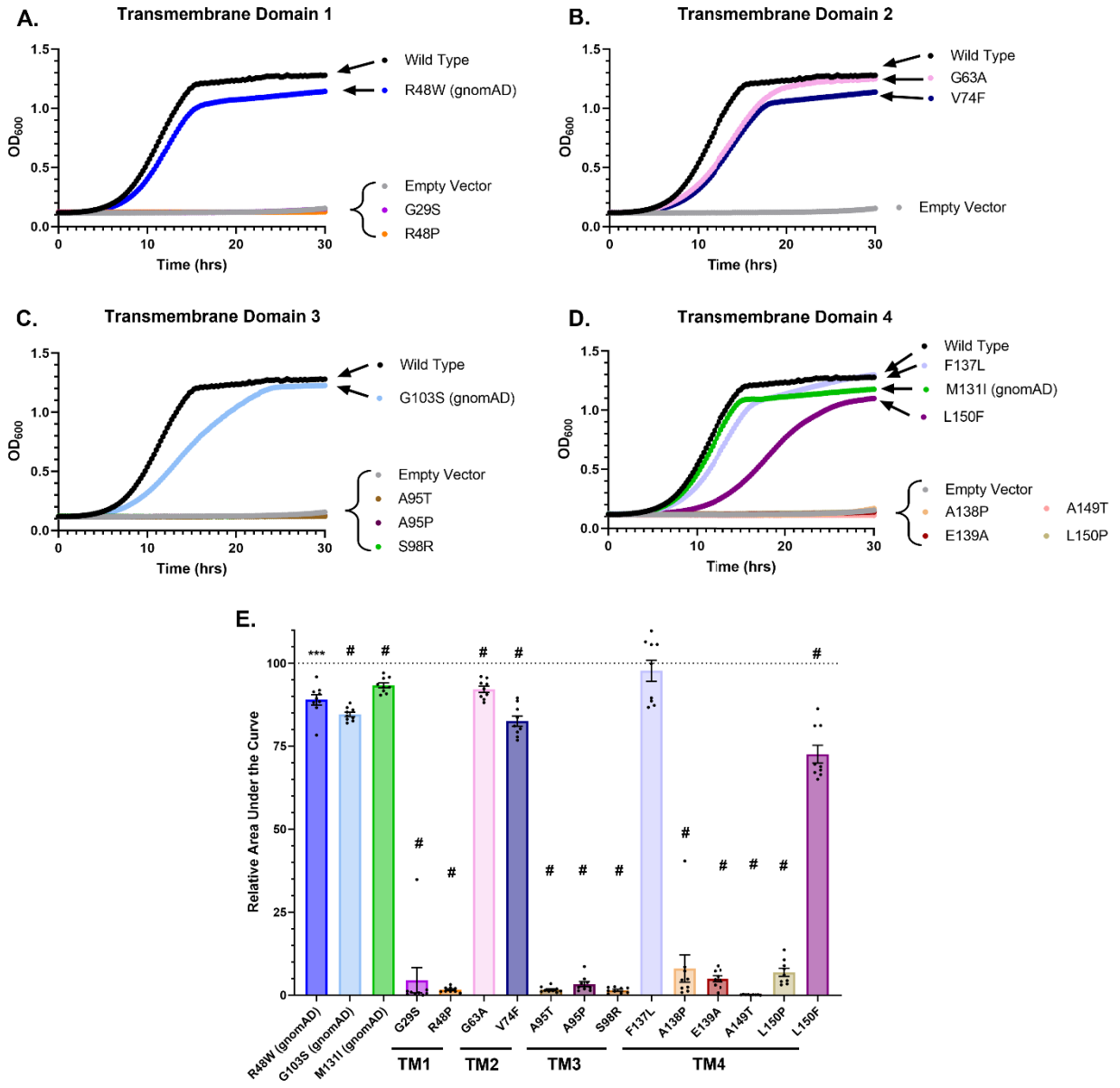


Figure 3.7. Patient Variants Result in Decreased Growth at 5mM $CaCl_2$. A-D) Growth curves of *vma3Δ S. cerevisiae* expressing patient or gnomAD variants when grown in YPD, pH 5.5 with 5mM $CaCl_2$. In all panels, wild type is shown in black and the empty vector in grey. Mean of 9 replicates per construct is shown with error bars omitted for clarity. Variants are grouped based on their location within or proximity to the nearest TM domain. **E**) eAUC was calculated using GrowthCurver [124]. Data was normalized within each plate with wild type as 100% (denoted by dotted line) and empty vector as 0% and is shown as mean ± SEM. A one sample t-test to a hypothetical mean of 100 was conducted with a Bonferroni correction (adjusted α level = 0.0003125). ***p = 0.0001, $^{\#}p$ < 0.0001.

Next, patient variants that grew at 5mM $CaCl_2$ along with the three gnomAD variants were tested at 100mM and 200mM $CaCl_2$ to determine whether a higher concentration of calcium would provide further separation of variants relative to the wild type rescue. At 100mM $CaCl_2$, p.L150F showed almost no growth compared to wild type ($p < 0.0001$, one sample t-test, **Figure 3.8A-B)** while the three other patient variants (p.G63A, p.V74F, p.F137L) showed intermediate growth relative to the wild type rescue. The growth of one gnomAD variant (p.M131I) was similar to the wild type rescue, while p.R48W and p.G103S both showed less growth relative to the wild type rescue. The three patient variants with intermediate growth at 100mM $CaCl_2$ and the three gnomAD variants were then tested at 200mM $CaCl_2$ (**Figure 3.8C-D**). Significantly less growth was seen with the three patient variants compared to the wild type rescue ($p < 0.0001$), while two gnomAD variants (p.R48W and p.M131I) were comparable to the wild type rescue. p.G103S showed less growth relative to the other gnomAD variants (p.R48W and p.M131I), but still yielded a significantly larger eAUC (**Figure 3.8D**) compared to the best growing patient variant at 200mM $CaCl_2$, p.F137L (59.22 ± 2.278 vs. 45.63 ± 2.243, $p = 0.0006$, two sample t-test).

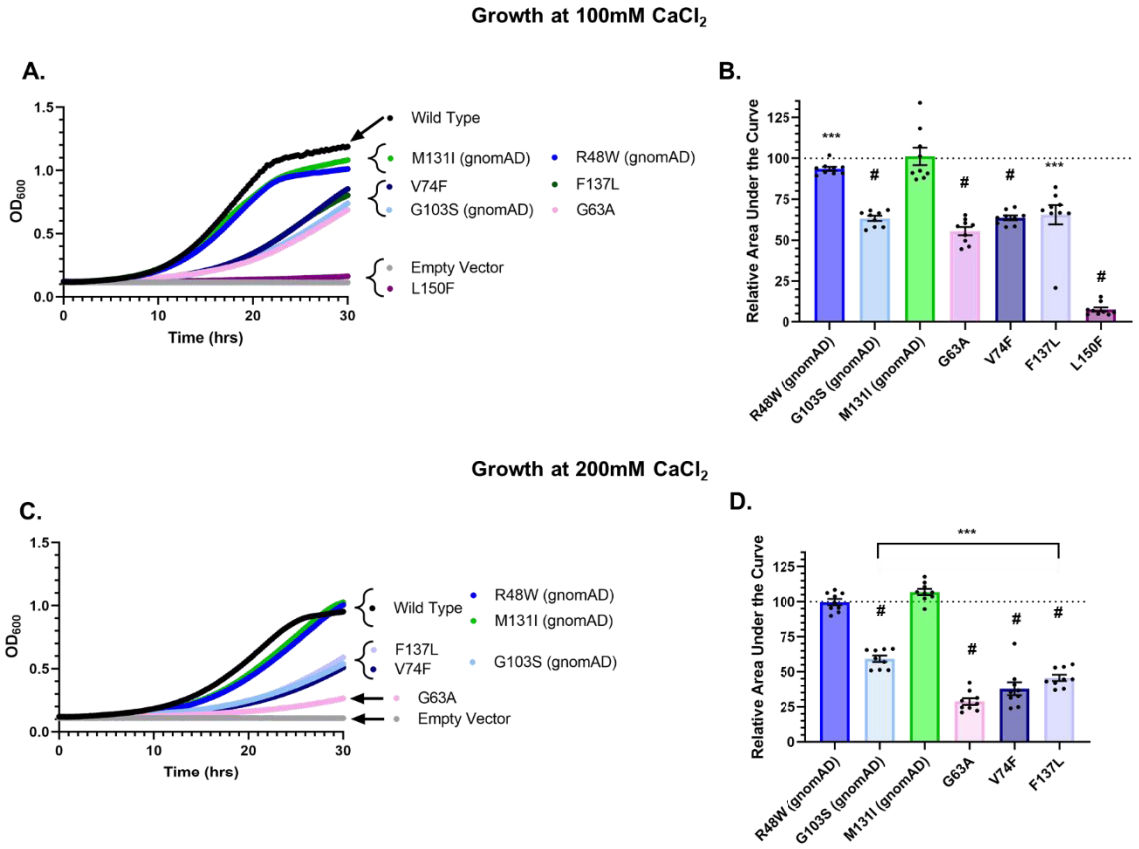


Figure 3.8. Growth of Selected Patient Variants at 100mM and 200mM $CaCl_2$. Growth curves of *vma3Δ S. cerevisiae* expressing patient variants at **A**) 100mM $CaCl_2$ or **C**) 200mM $CaCl_2$. In all panels, wild type is shown in black and the empty vector in grey. Mean of 9 replicates per construct is shown with error bars omitted for clarity. **B** and **D**) Empirical area under the curve was calculated using GrowthCurver [124]. Data was normalized within each plate with wild type as 100% (denoted by dotted line) and empty vector as 0% and is shown as Mean ± SEM. A one sample t-test to a hypothetical mean of 100 was conducted with a Bonferroni correction (adjusted α level = 0.00714 for 100mM and 0.00833 for 200mM). A two sample t-test between F137L and G103S at 200mM was conducted. ***p = 0.0001, $^{\#}p$ < 0.0001.

3.4.6 Assessment of Three Patient Variants in *C. elegans*

Next, we assessed the effects of a subset of patient *ATP6V0C* variants on neurological function using *C. elegans.* Worms express three orthologous genes to *ATP6V0C* in neurons, *vha-1, vha-2* and *vha-3.* VHA-2 and VHA-3 have identical amino acid sequences and share 66.7% amino acid identify with ATP6V0C, while VHA-1 has slightly less homology to ATP6V0C at 63% amino acid identity. The variants that were selected for further analysis were distributed throughout the protein, showed tolerance to 5mM $CaCl_2$ in the yeast growth assay, and were identified in patients with severe neurocognitive deficits and poly-medicated epilepsy. Specifically, we studied p.F137L (corresponding to p.F143L in VHA-2) and p.G63A and p.L150F (corresponding to p.G69A and p.L156F, respectively, in VHA-3*)*. A fourth strain carrying the p.A95T variant (corresponding to p.A101T in VHA-2) was generated but caused sterility in homozygous worms and was not studied further.

Worms expressing each variant were shorter and smaller than N2 (wild type) controls at day 1 of adulthood, indicating a morphological delay even in ideal physiological conditions ($p < 0.05$, one-way ANOVA, **Figure 3.9A-B**). When tested in liquid physiological M9 culture over a 4-hour period, movement of day 1 young adult worms with each mutation was comparable to N2 worms (**Figure 3.9C**)**.** However, when motor function was examined under osmotic stress conditions (350mM NaCl), mutants expressing p.G63A and p.F137L exhibited significantly reduced movement scores when compared to N2 worms ($p < 0.01$, two-way ANOVA **Figure 3.9D**). Although mutants expressing p.L150F also exhibited less movement than N2 worms, this difference was not statistically significant ($p = 0.0869$).

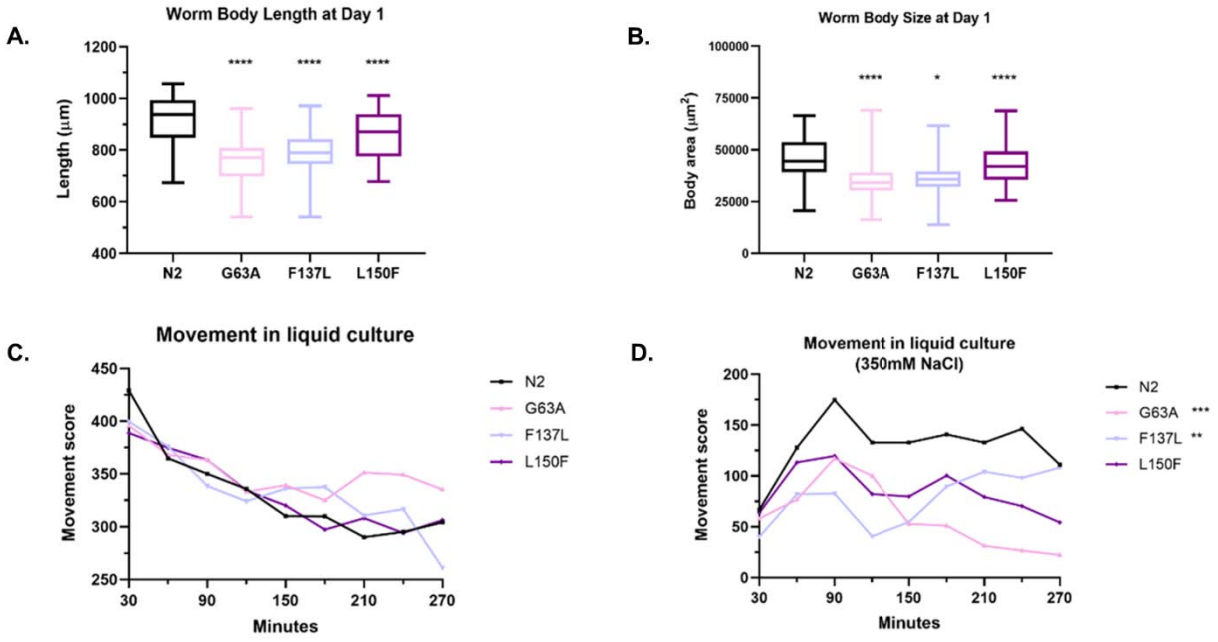


Figure 3.9. Expression of Patient Variants in *C. elegans* Result in Morphological Delay and Exacerbate Motor Dysfunction. A-B) WormLab analysis of body length and size at day 1 of adulthood. All mutants are shorter and smaller than N2 controls. **C-D**) Automated analysis of worm movement in liquid culture by WormTracker software. **C**) In physiological M9 solution, all mutants show no motor deficits. **D**) In presence of 350mM NaCl concentration the p.G63A ($p < 0.0001$) and p.F137L ($p < 0.0062$) mutants show reduced movement scores in liquid culture over 270 minutes. Reduced movement was also observed with the p.L150F variant, but this difference was not statistically significant ($p = 0.0869$). *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$ compared to N2 controls.

We next compared paralysis and lifespan of each mutant strain with N2 worms when maintained on NGM plates under physiological conditions or exposure to osmotic stress (200mM and 300mM NaCl). All mutants showed greater levels of paralysis when compared to N2 worms over a 14-day period, and these differences were strongly exacerbated in the presence of osmotic stress ($p < 0.0001$, log-rank (Mantel–Cox) test, **Figure 3.10A-C**). Lifespans of the mutant strains were also reduced when compared to N2 worms when maintained under physiological conditions and osmotic stress ($p < 0.0001$, log-rank (Mantel–Cox) test, **Figure 3.10D-F**).

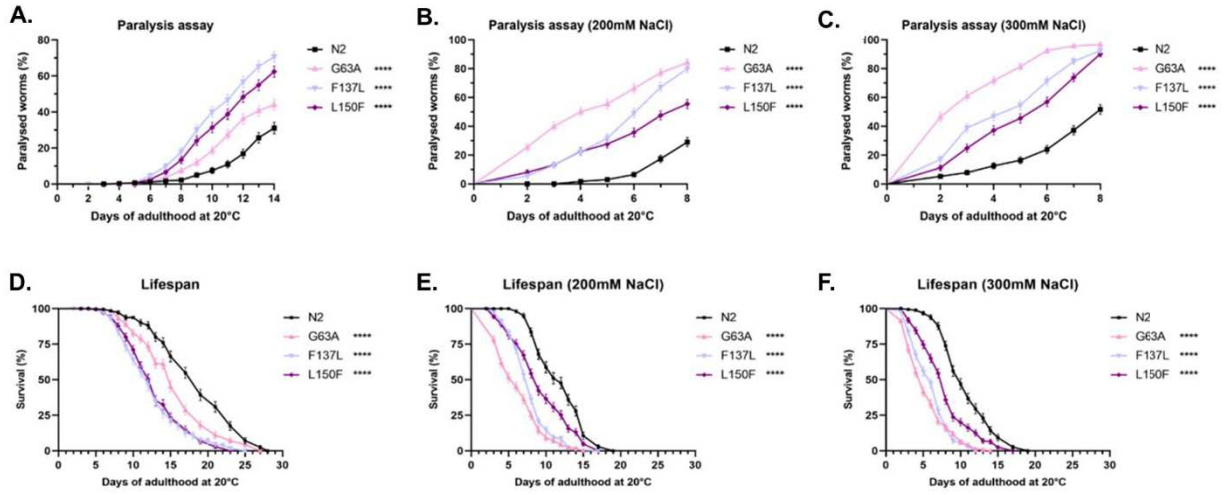


Figure 3.10. Expression of Patient Variants in *C. elegans* Results in Paralysis and Decreased Lifespan. A) All mutant strains showed increased paralysis over 14 days compared to N2 controls (n = 313-317 / strain, $p < 0.0001$). **B-C**) In presence of osmotic stress (200 or 300mM NaCl) the paralysis phenotype is exacerbated, leading to almost 100% paralysis after 8 days for the p.G63A strain (n = 246 – 260 / strain, $p < 0.0001$). **D**) All mutant strains exhibited reduced lifespan compared to N2 controls (n = 219 – 233 / strain, $p < 0.0001$). **E-F**) All mutant strains have reduced lifespans in presence of osmotic stress compared to N2 controls. (200mM NaCl: n = 182 – 228 / strain. 300mM, $p < 0.0001$) (300mM NaCl: n = 200 – 244 / strain, $p < 0.0001$). *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$ compared to N2 controls.

To compare fine motor phenotypes, we exposed the mutant strains and N2 worms to liquid M9 with 500mM NaCl for 30 minutes and analyzed movement using WormLab. Activity index and wave initiation were significantly increased in the mutant strains when compared to N2 worms ($p < 0.05$), but swimming speeds were not significantly altered, suggesting increased but uncoordinated movements ($p > 0.05$, one-way ANOVA, **Figure 3.11A-C**).

To test if the mutant strains have an impairment in nervous system signaling, we added an acetylcholinesterase inhibitor, aldicarb, to NGM plates and scored the number of paralyzed worms over a 2-hour period. Aldicarb causes an accumulation of acetylcholine in neuromuscular junctions resulting in muscle hypercontraction and acute paralysis and can be used to evaluate if there is dysfunction of either GABA or acetylcholine signalling [136, 137]. To confirm proper aldicarb effect, we included the unc-47(e307) and unc-64(e246) mutants. Unc-47(e307) is hypersensitive to aldicarb due to the lack of a vesicular GABA transporter gene (orthologous to *SLC32A1* in humans) required for GABA transmission [138]. Unc-64(e246) (orthologous to *STX1A* in humans) mutants have reduced cholinergic neurotransmission, making them resistant to aldicarb-induced paralysis [139, 140]. Worms expressing each patient variant showed greater paralysis in presence of aldicarb, compared to N2 worms ($p < 0.0001$, log-rank (Mantel–Cox) test, **Figure 3.11D**).

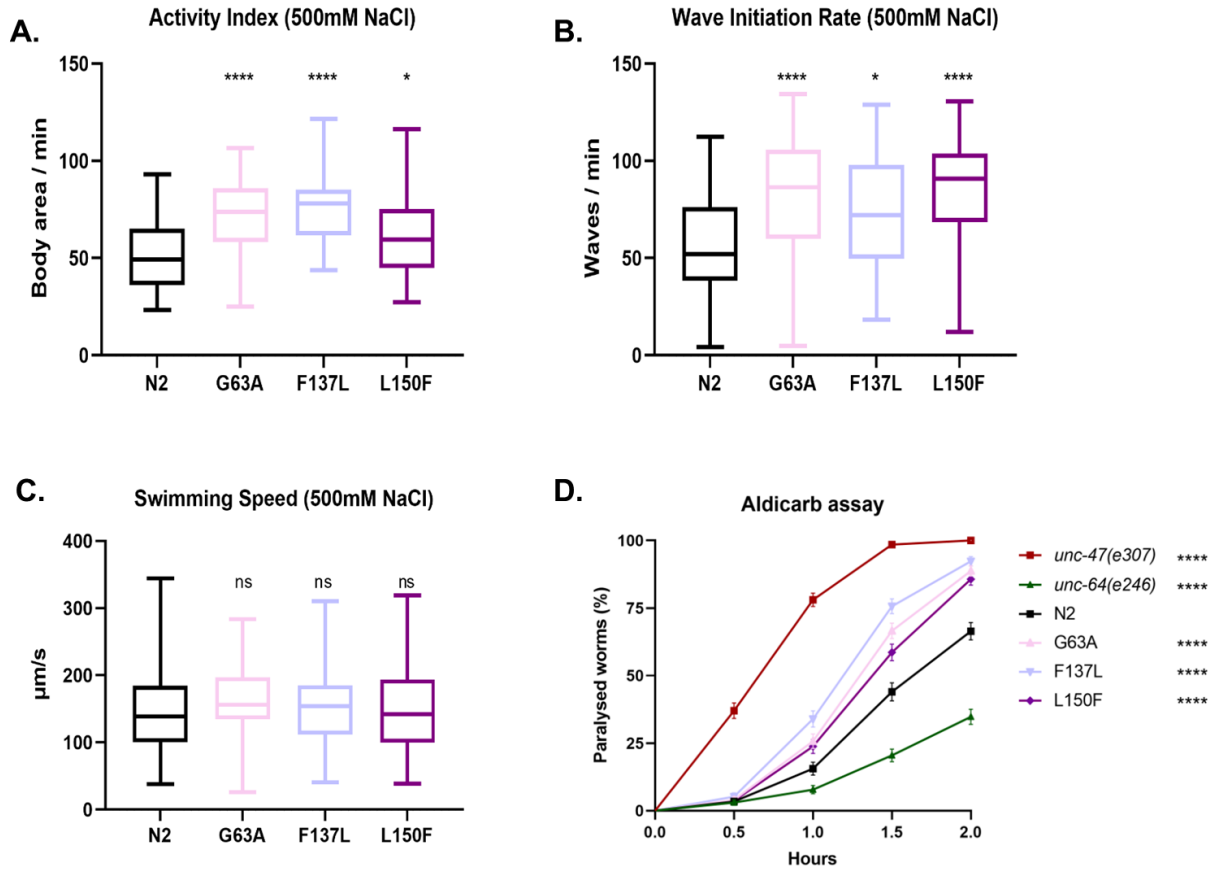


Figure 3.11. Patient variants cause increased uncoordinated movement and neuronal signaling dysfunction in *C. elegans*. A-C) Analysis of fine motor movement of worms after 30 minutes in 500mM NaCl liquid culture. Mutants show increased activity index and wave initiation (**A&B**), but swimming speed was not significantly altered (**C**). **D**) Synaptic transmission was evaluated by exposing day 1 adult worms to aldicarb. Worms were scored over a 2-hour period for paralysis. All mutants were hypersensitive to aldicarb treatment compared to N2 worms (n = 236 – 296 / strain, $p < 0.0001$). *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$ compared to N2 controls.

3.5 Discussion

In this study, we report the identification of heterozygous *ATP6V0C* variants in 26 patients with neurodevelopmental phenotypes. In general, this cohort of patients presented with development delay, early-onset epilepsy (mean age of onset: 20.72 ± 7.56 months), and varying severities of intellectual disability. Five patients with MRIs show hypoplasia or agenesis of the corpus callosum. Congenital cardiac abnormalities were also observed in four patients.

ATP6V0C is an evolutionary constrained gene as reflected by the high degree of amino acid homology between the human and yeast orthologs (72%), human and worm orthologs (63-67%), and human and Drosophila orthologs (78%), and the low number of missense variants seen in gnomAD (n=21 compared to an expected 108.5) [116]. Of the 18 unique patient missense variants, 16 are located in TM domains, with nine in TM4 which encompasses the p.E139 residue that is required for proton transport by the V-ATPase [87]. Consistent with evolutionary constraints on TM4, only one variant in gnomAD is located in this region of the protein (**Figure 3.2C**).

Interestingly, three patients (Patients 4, 15 and 19) were found to be mosaic for their identified *ATP6V0C* variant. Based on the available clinical information, the seizure phenotype of these patients may be less severe than for those with a germline variant (**Table 3.1**). Patient 4 has seizure-like episodes that started at 14 months, but are not supported electrographically, Patient 15 had seizure onset at 12 years of age, and Patient 19 has not reported any seizure or seizure-like episodes. It is well established that the timing of the post-zygotic mutation event, and the affected tissues, can have a large influence on phenotypic presentation and severity in patients with somatic mosaicism, underlying differences in clinical presentation between patients with germline and mosaic variants [21].

Nine copies of ATP6V0C and one copy of ATP6V0B assemble to form the intramembrane c-ring of the V-ATPase which uses a rotary mechanism to translocate protons across the membrane (**Figure 3.1**) [107]. Normally, frameshifting variants are predicted to

cause NMD of the mutant mRNA, which would result in reduced protein levels. However, two of the four frameshift variants (c.340_355del16 and c.352_362delins) are located in last exon of *ATP6V0C* and are expected to escape NMD. The two frameshift variants in exon 2 may also escape NMD given the proximity of c.134_135delCT to the start codon and c.214delG being within 50bp of the last exon-exon junction (**Figure 3.2A**) [141]. Additionally, modeling showed that the majority of missense variants are likely to act as a 'stone in the gear' between ATP6V0C and ATP6V0A inhibiting the rotatory mechanism, consistent with a dominant negative mechanism (**Figure 3.4A-B)**. Therefore, we believe missense variants and those predicted to escape NMD are likely to act via a dominant negative mechanism, but nonsense variants and microdeletions containing *ATP6V0C* act via haploinsufficiency as demonstrated by our Drosophila data (**Figure 3.3**).

Twelve disease-associated missense variants were examined in yeast. When the uptake of LysoSensor was measured, we saw that nine variants were associated with little to no fluorescence, indicating significant reduction or loss of V-ATPase activity. Higher levels of fluorescence were seen when the p.G63A, p.F137L, and p.L150F patient variants were expressed (**Figure 3.6**). Growth curves generated by yeast expressing the 12 patient variants mirrored observations from the LysoSensor assay, with most variants resulting in little or no growth, and intermediate levels of growth observed with the p.G63A, p.V74F, p.F137L, and p.L150F variants (**Figure 3.7E**). To further examine the effect of these intermediate variants on developmental and neurological function, we modeled p.G63A, p.F137L, and p.L150F in worms. Expression of all three variants resulted in morphological delay as indicated by reduced body size and length at day 1 of adulthood (**Figure 3.9A-B**). Mutant worms also exhibited greater levels of paralysis and decreased lifespan when compared to N2 worms, and these phenotypes were exacerbated under osmotic stress (**Figure 3.10D-F**). Mutants also exhibited increased activity and wave initiation rates, but speed was unaltered, suggestive of hyperactive and uncoordinated movement (**Figure 3.11A-C**). The p.A95T variant which showed almost no V-

ATPase function when tested in yeast (**Figure 3.6, 3.7**), resulted in homozygous sterility in worms, suggestive of a greater impact on V-ATPase function.

Additionally, we functionally examined, in yeast, three variants from gnomAD (p.R48W, p.G103S, p.M131I) that had high CADD scores and were predicted to be damaging by SIFT or Poly-Phen2 which contrasts the assumption that population variants would be benign (**Table 3.2**). Expression of p.R48W and p.M131I resulted in growth that was more similar to the wild type rescue compared to patient variants at 5mM $CaCl_2$ and did not significantly differ from wild type rescue as higher concentrations of $CaCl_2$ were tested (**Figure 3.7, 3.8**). p.G103S, which had the highest CADD score of the gnomAD variants tested, resulted in significantly decreased growth compared to the wild type rescue across all $CaCl_2$ concentrations tested (**Figure 3.7, 3.8**). Nevertheless, p.G103S still yielded a significantly larger eAUC when compared to the best growing patient variant (p.F137L), suggesting the possibility of a level of decrease in V-ATPase activity that can be tolerated. However, we cannot exclude the possibility that a mild neurological deficit could be associated with the p.G103S variant. Further testing of population variants could establish the minimum level of V-ATPase activity required to maintain normal function.

The cellular mechanisms that underlie the clinical features in patients with pathogenic variants in *ATP6V0C* are not yet known. However, given the importance of acidification driven by the V-ATPase in many cellular processes, it is likely that several pathways are impacted. Firstly, the acidification of SVs within the central nervous system by the V-ATPase allows transporters, such as VGLUTs and VGAT, to load their respective cargo into SVs [142, 143]. Our modeling data predicted that the majority of patient variants would interfere with the interaction between ATP6V0C and ATP6V0A (**Figure 3.4**). Interestingly, patients with pathogenic *ATP6V0A1* variants (encoding ATP6V0A) present with developmental delay, epilepsy, intellectual disability and cerebellar atrophy, similar to patients with pathogenic *ATP6V0C* variants [101]. Recent work has shown that primary cortical neurons from

Atp6v0a1$^{A512P/A512P}$ mice, modeling a variant identified in a patient, have decreased SV neurotransmitter content and form fewer synapses [101]. Given the functional interaction between ATP6V0C and ATP6V0A1, we speculate that synaptic defects likely contribute to disease pathology in patients with pathogenic *ATP6V0C* variants. These defects may also be independent of the ATPase activity of the V_1 domain as the V_0 domain is involved in neurotransmitter release independent of its acidification of SVs [144-146].

The V-ATPase also plays an important role within the trans-Golgi network (TGN) and inhibition of V-ATPase driven acidification can lead to improper trafficking and sorting of various membrane bound proteins, including neuropeptides and neuropeptide receptors [147-149]. Additionally, pathogenic variants in *SLC9A7* in patients with developmental delay, intellectual disability, and muscle weakness have been shown to cause alkalization of the TGN [150]. Therefore V-ATPase dysfunction may also lead to altered synaptic signaling via disruption of trafficking and sorting of receptors to the synapse. Consistent with this prediction of altered synaptic signaling, worms expressing the p.G63A, p.F137L, and p.L150F variants displayed greater sensitivity to aldicarb and higher rates of paralysis when compared to N2 worms (**Figure 3.11D**).

Lastly, perturbations of lysosomal and autophagy pathways may also contribute to the epilepsy and neurodevelopmental phenotypes seen in these patients. The V-ATPase plays an essential role in acidifying endosomes, lysosomes, and autophagosomes which ultimately creates the environment needed for proper trafficking and maturation of endocytic organelles and acid hydrolase function within the autophagy pathway [88, 151]. Previous work by Nakamura *et al.* demonstrated that the V-ATPase is required for protein degradation from autophagic bodies in yeast vacuoles [152] and Fassio *et al.* showed impairments in autophagic flux caused by pathogenic variants in *ATP6V1A* [100]. In recent years, the contribution of impaired autophagy to neurodegenerative and neurodevelopmental disorders, including epilepsy, has risen in importance [100, 153, 154]. We hypothesize that the epilepsy and other

neurodevelopmental phenotypes seen in patients with pathogenic V-ATPase variants may be due to a combinatorial effect of impaired synaptic signaling, trafficking and sorting of various membrane bound proteins, and defects along the endomembrane system including the lysosomal/autophagy degradation pathway. Further research on *ATP6V0C* variants will be required to better understand underlying disease mechanisms.

The prevalence of neurodevelopmental disorders, including epilepsy, resulting from variants in *ATP6V0C* is likely underestimated as this gene, to the best of our knowledge, is not currently included on commercially available epilepsy or intellectual disability gene panels. Of the 23 genes that encode for a subunit of the V-ATPase, twelve, including *ATP6V0C*, are associated with disease. Ten additional members of the complex are expressed in the central nervous system but are currently not associated with disease (**Table A1**) [155]. Screening of these genes for potential pathogenic variants in patients with disorders such as epilepsy and intellectual disability should be undertaken. Additionally, longitudinal studies in patients with *ATP6V0C* mutations, and the identification of additional patients, will play an important role in resolving the full spectrum of co-morbidities associated with altered ATP6V0C and V-ATPase function.

In summary, we report 26 patients with heterozygous *ATP6V0C* variants who presented with developmental delay, early-onset epilepsy, and intellectual disability. *In silico* modeling suggests that the majority of patient missense variants disrupt the interaction between the ATP6V0C and ATP6V0A subunits, and functional testing revealed that these variants decrease V-ATPase activity in yeast, and impair motor function, growth and lifespan in worms. Further work is needed to fully elucidate the mechanism(s) by which altered *ATP6V0C* function lead to the range of observed clinical phenotypes, and whether other V-ATPase subunits not currently known to cause disease harbor pathogenic variants in patients with neurodevelopmental disorders without a current genetic diagnosis.

3.6 Acknowledgements and Funding

We would like to thank the patients and their families for their participation in this study. We would also like to thank Dr. Anita Corbett, and Dr. Meleah Hickman (Emory University) for helpful advice and for providing reagents and equipment. Figure 1 was adapted from “V-ATP synthase”, by BioRender.com (2022); retrieved from https://app.biorender.com/biorender-templates.

This study was supported by a training grant appointment to K.A.M (5T32GM008490), an Emory University Research Council grant to A.E. J.B. is funded by a senior clinical investigator fellowship of the FWO-Flanders. This study was also supported in part by the Emory University Integrated Cellular Imaging Core (EICIC) and the Emory Integrated Genomics Core (EIGC) shared resources of Winship Cancer Institute of Emory University, NIH/NCI (P30CA138292), the Canadian Rare Disease Models and Mechanisms Network (RDMM), Citizens United for Research in Epilepsy (339143), and the National Institute for Neurological Disorders and Stroke (U01-NS077303-04S1 and R01-NS058721). The content is solely the responsibility of the authors and does not necessarily reflect the official views of the National Institutes of Health.

Additionally, this research was made possible through access to the data and findings generated by the 100,000 Genomes Project. The 100,000 Genomes Project is managed by Genomics England Limited (a wholly owned company of the Department of Health and Social Care). The 100,000 Genomes Project is funded by the National Institute for Health Research and NHS England. The Wellcome Trust, Cancer Research UK and the Medical Research Council have also funded research infrastructure. The 100,000 Genomes Project uses data provided by patients and collected by the National Health Service as part of their care and support.

Chapter 4

Novel Missense *CNTNAP2* Variant Identified in Two Consanguineous Pakistani Families with Epilepsy and Intellectual Disability

4.1 Summary

This work stemmed from the analysis of 10 consanguineous families with epilepsy of suspected genetic origin. The families are of Afridi tribal heritage in Pakistan where there is a high prevalence of consanguineous marriages. Due to this we hypothesized that the causative variants would be recessive and shared between the affected individuals in each family. Analysis of the family described in this chapter revealed a homozygous p.G228R variant in *CNTNAP2* in the two affected brothers.

During the course of this study, we became aware of a second, unrelated family of Pakistani origin that had two sisters with epilepsy and intellectual disability that were homozygous for the same p.G228R variant in *CNTNAP2*. Segregation analysis of the *CNTNAP2* variant showed that it segregated with disease in the second family, and that none of the other variants identified in the first family were present in the second family.

CNTNAP2 has previously been associated with AR focal epilepsy and intellectual disability in other consanguineous families. Interestingly, in the previous cases reported, at least one allele resulted in loss of function (premature truncation, splice-site, frameshift, or indels). Our report is the first known case where a homozygous *CNTNAP2* missense variant is reported to be likely pathogenic. Further functional studies of the variant identified in this family as well as homozygous missense variants in *CNTNAP2* in general will prove useful towards gaining a better mechanistic understanding of how these variants affect protein function.

4.2 Introduction

Epilepsy is a neurological disorder characterized by the presence of recurrent and unprovoked seizures, and it is estimated that 70-80% of epilepsy cases have a genetic contribution [4]. Discovery of novel genetic variants contributing to epilepsy have largely used experimental designs that favor the discovery of dominant acting alleles [4, 156]. As a consequence, while targeted gene panels have yielded a diagnostic rate of approximately 18% in epilepsy patients, only 4-8% of these cases were due to variants in genes known to cause AR forms of epilepsy [54, 55, 157]. WES provides a higher diagnostic rate of 20-30% in families with suspected AR epilepsy, illustrating the utility of WES over gene panels when AR inheritance is suspected [39, 158, 159]. Several genes with broad biological functions have now been associated with AR epilepsy [18]. AR forms of epilepsy can be severe with early-onset, such as epilepsy caused by variants in *ALDH7A1*, or have onset in adolescence or later as with variants in *CSTB* [18].

We identified two brothers affected with epilepsy, autism and intellectual disability of unknown etiology from a consanguineous family (Family 1) belonging to the Afridi tribe within the Federally Administered Tribal Area of the province Khyber Pakhtunkhwa, Pakistan. Consanguineous marriages in this region are common, resulting in a high prevalence of genetically recessive disorders. Through WGS we identified 32 genomic regions that were identical by descent between the two brothers. Variant annotation and filtering reduced the candidate regions to six harboring 21 homozygous genetic variants of which four were located within exons of different genes. Segregation of the exonic variants among additional unaffected family members revealed homozygous inheritance of variants in *PDZD7, RBM20,* and *CNTNAP2* in only the affected siblings. A second consanguineous family (Family 2) of Pakistani origin was subsequently identified in which two siblings were also affected with refractory focal epilepsy and intellectual disability. The two affected members of Family 2 were found to be homozygous for the same *CNTNAP2* variant identified in Family 1.

4.3 Subjects and Methods

Family 1

Family 1 is from Pakistan and is of Afridi tribal heritage. Six members from this family participated in this study. Clinical information and blood samples were obtained from all participants after they or a parent/guardian had provided written informed consent. This study was approved by the Ethical Committee at the Institute of Biotechnology and Genetic Engineering, the University of Agriculture, Peshawar, Pakistan.

Genomic DNA was extracted from whole blood for all participants with the Omega SQ II Blood DNA kit following manufacturer's protocol (D0714, Omega Bio Tek). Paired-end sequencing libraries were generated for individuals V-2 and V-4 using a KAPA HyperPrep kit (Roche, KK8504) following manufacturer protocols. WGS was performed using an Illumina NovaSeq6000 instrument according to manufacturer protocols. Reads were mapped to hg38 using PEMapper and variants (single-nucleotide, SNVs, and indels) called using PECaller [160]. Average sequencing depth for both samples was greater than 34X. Runs of homozygosity were identified using the "homozyg group-verbose" function in PLINK (pngu.mgh.harvard.edu/Purcell/plink/) [161].

Variants were annotated using Bystro (bystro.io) [162]. Filtering of shared homozygous variants between the affected siblings using the following command: "CADD > 15, MAF < 0.001, exonic, splice" (**Figure 4.1**) resulted in the identification of exonic variants in *PDZD7*, *ALG6*, *RBM20*, and *CNTNAP2*.

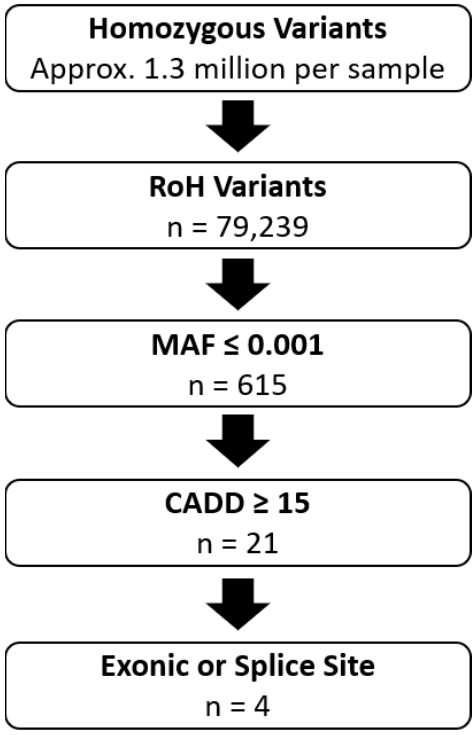


Figure 4.1. Overview of Variant Filtering in Family 1. RoH Variants: Variants that fall within the 103.8 MB of shared homozygous regions between the two affected individuals. MAF ≤ 0.001: Number of variants with a minor allele frequency ≤ 0.001 based on data from gnomAD. CADD ≥ 15: Number of variants with a CADD score ≥ 15. Exonic or Splice Site: Number of variants located within exons or splice acceptor or donor sites. n = indicates remaining number of variants after each step.

Family 2

Family 2 is also of Pakistani origin. Six members of this family participated in this study. Clinical information and blood samples from the affected siblings and salvia samples from four other family members were obtained after they or a parent/guardian had provided written informed consent.

Genomic DNA was extracted from whole blood (III-4 and III-5) or from saliva (II-10, II-11, III-2, III-6), Individuals III-4 and III-5 were screened using the National Health Services Genetic Epilepsy Syndromes gene panel (Test ID 59.1) which comprises 402 known and evidence-based epilepsy genes.

Segregation Analysis

Segregation of variants in *PDZD7*, *ALG6*, *RBM20*, and *CNTNAP2* was confirmed by sequencing in both families. The following primers were used for PCR and sequencing: PDZD7_F-ATGTGTGCCCTTCTCTAACTG and PDZD7_R-GTCCAGGCGAGGGTAAGTT, ALG6_F-TAAGTTGTCTGAGATTCCAGG and ALG6_R-GACAAACAGGCTCCAATC, RBM20_F-AAAGGGAACCGTCTTCTG and RBM20_R-GCATCGCCTCTTTATGTTAAG, CNTNAP2_F-AACAGAGGACTGTCAATTTC and CNTNAP2_R-CAGTCAGAACACACCTAAGT.

Lollipop Variant Plot

The lollipop variant plot was generated using freely available source code from https://github.com/pbnjay/lollipops [114]. Population variants in gnomAD for *CNTNAP2* were taken from v2.1.1. Human Gene Mutation Database (HGMD) variants for *CNTNAP2* were taken from HGMD Professional version 2021.3. Resulting lollipop diagrams of HGMD and gnomAD variants were merged into a single image. UniProt accession Q9UHC6 was used to draw the protein domains which were overlaid on the final image.

GSA Analysis

Genomic DNA samples from six members of Family 1 and four members of Family 2 were run on the Illumina Global Screening Array (GSA) v3.0. GSA variants were processed and

called using Genome Studio. High-quality, high-completeness SNPs called on both the GSA and by WGS in Family 1 members were merged into a single dataset (302,802 in total) to allow pairwise comparison of all Family 2 and Family 2 members. Identity by descent (IBD) fractions were estimated from observed identity by state (IBS) counts.

Data Availability

WGS data for individuals V-2 and V-4 in Family 1 was deposited in SRA (Accession #: PRJNA695560) and the *CNTNAP2* variant was deposited in ClinVar (Accession #: SCV001478474).

4.4 Results

4.4.1 Clinical Presentation and Family History

4.4.1.1 Family 1

The two affected brothers (V-2 and V-4) are members of a consanguineous family from the Afridi tribe in Pakistan (**Figure 4.2A**). Both individuals exhibited normal height, weight and head circumference during the perinatal period. Clinical history of the parents and full siblings of V-2 and V-4 revealed no indication of neurological disease.

Individual V-2 is currently 13 years old. He presented with primary generalized epilepsy onset at 3 years of age. Partial seizure control was achieved with AEDs; however, occasional breakthrough seizures still occur (**Table 4.1**). Aggressive behavior was observed until four years of age. Electroencephalogram (EEG) analysis performed under sedation at 4 years of age identified brief, recurrent, generalized spike and wave discharges.

Individual V-4 is currently 10 years old. He presented with focal-onset epilepsy at 2.5 years of age. AEDs were successful in temporarily decreasing seizure frequency; however, he currently has daily seizures of variable frequency (up to 2–3 seizures per day). Seizure onset was accompanied by language regression, and he currently speaks few words and is unable to communicate in full sentences. This individual also exhibits severe aggressive behaviors. EEG

analysis was conducted under sedation at 7 years of age and recurrent left, fronto-temporal spikes were observed, consistent with a diagnosis of focal epilepsy. Both affected siblings are also diagnosed with intellectual disability and autism spectrum disorder.

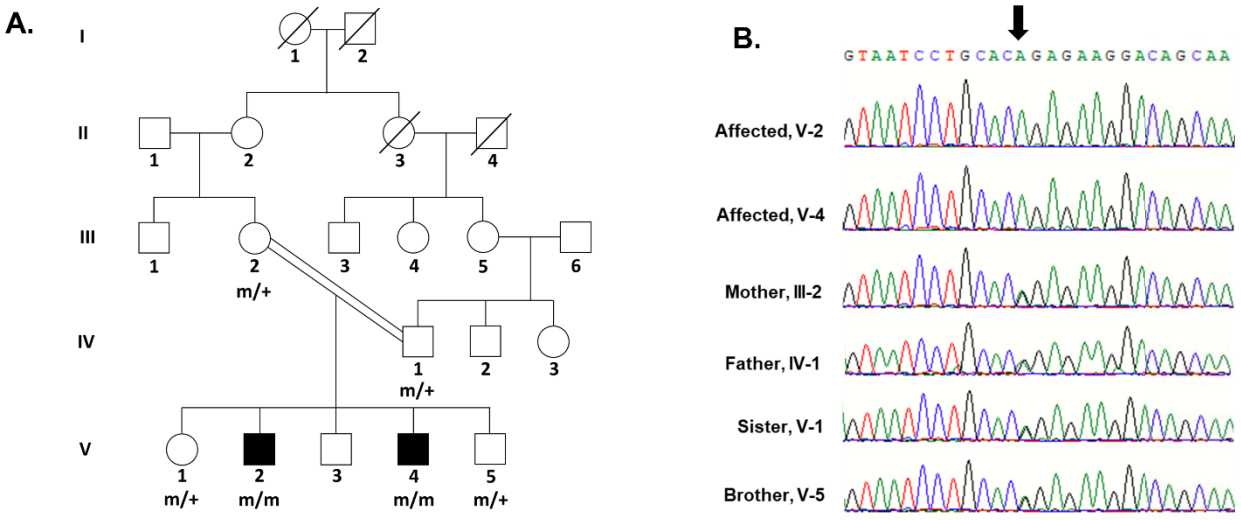


Figure 4.2. *CNTNAP2* p.Gly228Arg Variant Segregates with Disease in Family 1. A) Pedigree for Family 1 with affected status (filled symbol) shown. Genotype for *CNTNAP2* variant is shown under individuals who underwent Sanger sequencing. "+" wild-type allele, "m" mutant allele. **B**) Sequence traces from Sanger sequencing of *CNTNAP2* in members of Family 2. The position of the homozygous c.758G>A variant in the affected siblings is indicated by the arrow. Both parents in the family are heterozygous and the unaffected siblings are heterozygous for this nucleotide substitution.

	Family 1		Family 2	
	V-2	V-4	III-4	III-5
Current Age	13 years	10 years	7 years	5 years
Epilepsy Syndrome	Primary Generalized Epilepsy	Focal Onset Epilepsy	Focal Onset Epilepsy	Focal Onset Epilepsy
Age of Seizure Onset	3 years	2.5 years	3 years	2.5 years
Primary Seizure Type	Generalized Tonic-Clonic	Focal Onset Aware (Partial)	Focal	Left side Tonic followed by Generalized Tonic-Clonic
Current AEDs	Valproic Acid, Diazepam, Clonazepam, Risperidone	Carbamazepine, Valproic Acid, Levetiracetam, Clonazepam, Methylphenidate	Carbamazepine and Lamotrigine	Carbamazepine, Valproic Acid
Seizures Controlled?	Some breakthrough seizures	No	No (daily seizures)	No
Intellectual Disability	Moderate	Moderate	Yes	Yes
Developmental Delay	Yes, with speech regression	Yes, with speech regression	Yes, poor speech and language	Yes, poor speech abilities
Autism spectrum disorder	Yes	Yes	Poor social abilities	Poor social abilities
EEG	4 years: frequent brief generalized high amplitude SWDs	7 years: frequent high amplitude spike discharges in left fronto-temporal region	Confirmed seizures of temporal lobe origin	Interictal frontotemporal epileptiform abnormalities and post-ictal temporal slowing on left side
Other Behavioral Features	Aggression, until 4 years	Aggression	Aggression	Aggression
Other Features	Body movement imbalance (mild)	Body movement imbalance (mild)	Stereotypic hand flapping movements, post-natal microcephaly	Stereotypic hand flapping movements

Table 4.1. Clinical Presentation of Affected Siblings.

Abbreviations: AEDs = antiepileptic drugs, SWD = spike-wave discharges

4.4.1.2 Family 2

The two affected sisters (III-4 and III-5) are member of a consanguineous family of Pakistani origin (**Figure 4.3A, Table 4.1**). Antenatal, birth and early development during the first year were normal. Routine blood tests, and serum, urine and cerebrospinal fluid metabolic analyses were normal. No structural brain abnormalities were visible via MRI in either sibling. Clinical history of the parents and full siblings of III-4 and III-5 revealed no indication of neurological disease.

III-4 is currently 7 years old. She presented with focal onset seizures at 3 years of age. She experiences hypermotor movement of limbs, and eye deviation to the right during her seizures which typically last 1 minute. She currently has daily seizures despite AED treatment. EEG analysis confirmed epileptic seizures of temporal lobe origin.

III-5 is currently 5 years old. She presented with an explosive onset of seizures at 2.5 years of age. She experiences pallor, eye deviation to the left, tonic seizures affecting the left side followed by generalized tonic-clonic seizures. Her seizures last from 2-5 minutes and typically occur in clusters. Partial seizure control was obtained through AED treatment. EEG analysis showed interictal frontotemporal epileptiform abnormalities and post ictal temporal slowing on the left side, but seizures have not been captured. Both affected sisters exhibited significant learning difficulties, poor social abilities, and aggressive behaviors by 4 years of age.

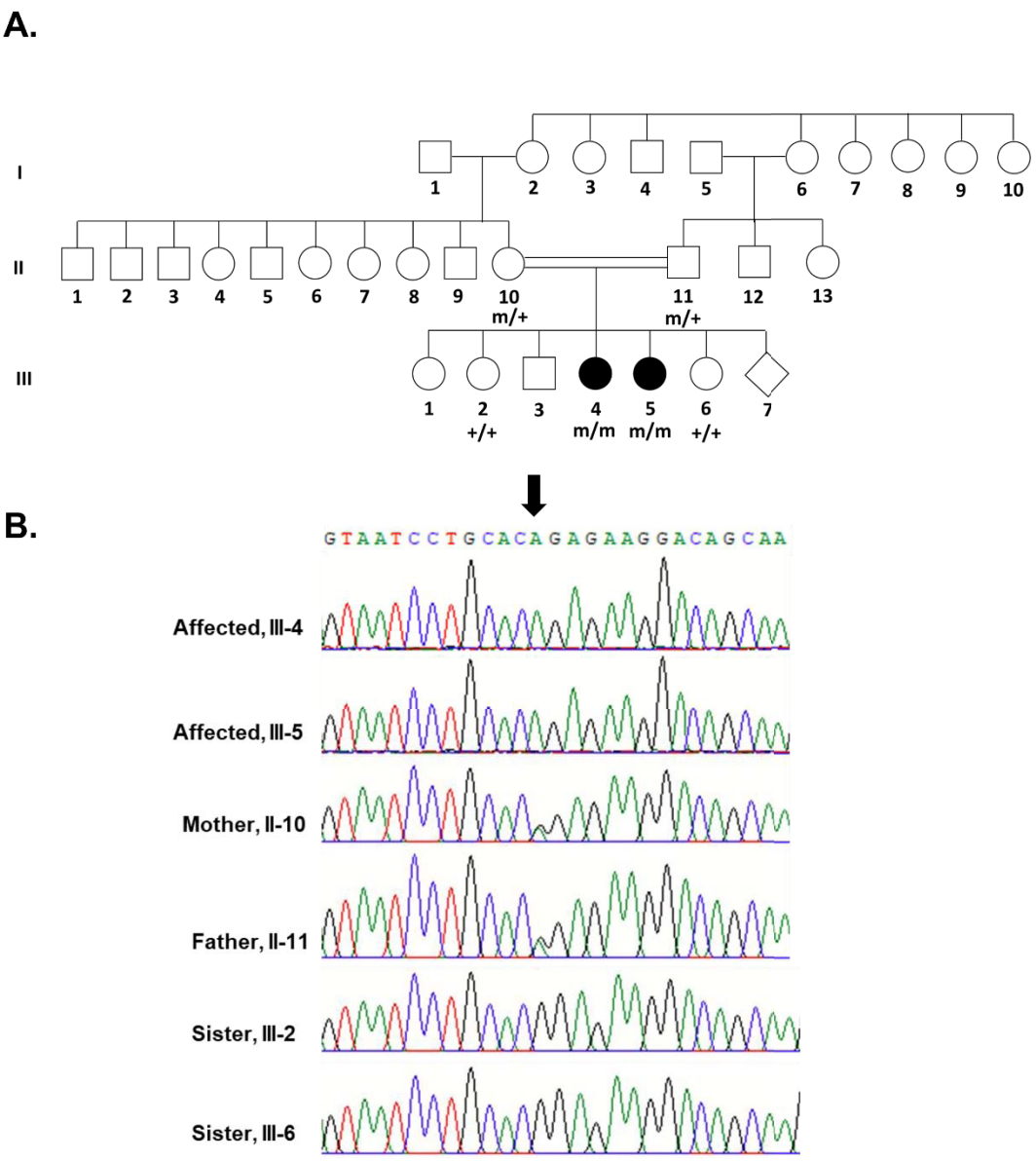


Figure 4.3. *CNTNAP2* p.Gly228Arg Variant Segregates with Disease in Family 2. A) Pedigree for Family 2 with affected status (filled symbol) shown. Genotype for *CNTNAP2* variant is shown under individuals who underwent Sanger sequencing. "+" wild-type allele, "m" mutant allele. **B**) Sequence traces from Sanger sequencing of *CNTNAP2* in members of Family 2. The position of the homozygous c.758G>A variant in the affected siblings is indicated by the arrow. Both parents in the family are heterozygous and the unaffected siblings are wild type for this nucleotide substitution.

4.4.2 Genomic Analyses

4.4.2.1 Family 1

The affected siblings had not previously undergone testing to determine the genetic basis for their disease. Given the consanguineous nature of the family, we hypothesized that a homozygous variant shared by the two affected siblings was the most likely explanation for their clinical features. We performed WGS on genomic DNA isolated from whole-blood samples from the two affected brothers.

Homozygosity analysis revealed that 32 genomic regions, totaling 103.8 MB, were shared by the two affected brothers and homozygous variants within these regions were identified. Variants, either SNVs or indels with CADD scores ≥ 15 and minor allele frequency (MAF) ≤ 0.001 were selected for further analysis (**Figure 4.1, Table 4.2**). After filtering for variants in exons or splice sites, four candidate variants were identified: *PDZD7* (c.1348_1350delGAG, p.Glu450del), *ALG6* (c.1033G>C, p.Glu345Gln), *RBM20* (c.1587C>G, p.Ser529Arg), and *CNTNAP2* (c.785G>A, p.Gly228Arg) (**Table 4.3**).

Chr	Position	Ref	Alt	Variant Type	Site Type	Gene	dbSNP ID	CADD Score	gnomAD frequency
1	63419415	G	C	SNV	exonic	*ALG6*	rs142707911	28.2	0
1	69324851	C	T	SNV	intergenic		rs559114975	19	0
3	174840870	T	A	SNV	intergenic		rs545360340	15.4	0.000031859
3	176594852	C	T	SNV	intergenic		rs369457267	21.8	0.000031849
4	13371261	C	A	SNV	intronic		rs937001005	15.5	0
7	147108278	G	A	SNV	exonic	*CNTNAP2*	rs371512835	21.4	0
10	101018271	CTC	-	DEL	exonic	*PDZD7*	rs555444131	23.3	0.000010614
10	101249392	G	A	SNV	intergenic		rs548280240	22	0
10	101901487	T	G	SNV	intronic		rs569668688	15.2	0
10	110797567	C	G	SNV	exonic	*RBM20*	rs539932441	32	0
10	110851117	T	A	SNV	intergenic		rs372591825	19.3	0
10	110858145	G	A	SNV	intergenic		rs368141988	16.2	0
10	110921574	T	G	SNV	intronic		rs187395797	16.3	0
10	110985777	G	A	SNV	intronic		rs201258692	17.3	0.000548420
10	111004594	T	-	DEL	intronic		rs730881016	16.1	0.000556479
10	111010957	C	T	SNV	intronic		rs370486167	16.3	0
21	24018297	G	A	SNV	intergenic		rs189938410	22.5	0
21	25549608	T	A	SNV	intergenic		rs551677808	17.9	0
21	31174041	G	A	SNV	intronic		rs184146552	16	0
21	32678632	A	T	SNV	intronic		rs75688000	15.7	0.000009427
21	36067727	T	G	SNV	intergenic		rs914317800	15.1	0

Table 4.2. Shared Variants Identified in Affected Brothers from Family 1.

Homozygous variants with MAF ≤ 0.001, CADD ≥ 15. Genomic position is based on hg38. gnomAD frequency represents homozygous allele frequency.

Abbreviations: Alt = alternate allele, Chr = chromosome, DEL = Deletion, Ref = reference allele, SNV = single nucleotide variant

Gene		*CNTNAP2*	*ALG6*	*RBM20*	*PDZD7*
Chromosome		7	1	10	10
HGVS DNA		NM_014141.6: c.682G>A	NM_013339.4: c.1033G>C	NM_001134363.3: c.1587C>G	NM_001195263.2: c.1348_1350delGAG
HGVS Protein		p.Gly228Arg	p.Glu345Gln	p.Ser529Arg	p.Glu450del
CADD Score		21.4	28.2	32	23.3
Population Frequency	**gnomAD[a]**	0	0	0	0.000010614
	GMEVariome[b]	0	0	0	0
	GenomeAsia 100K[c]	0	0	0	0
Family 1	**V:2**	Hom	Hom	Hom	Hom
	V:4	Hom	Hom	Hom	Hom
	III:2	Het	Het	Het	Het
	IV:1	Het	Het	Het	Het
	V:1	Het	Hom	Het	Het
	V:5	Het	Het	Het	Het
Family 2	**III:4**	Hom	WT	WT	WT
	III:5	Hom	WT	WT	WT
	II:10	Het	WT	WT	WT
	II:11	Het	WT	WT	WT
	III:2	WT	WT	WT	WT
	III:6	WT	WT	WT	WT

Table 4.3. Segregation of Candidate Variants.

[a]gnomAD homozygous allele frequency is reported

[b]GMEVariome is a collection of sequence data from individuals in the Greater Middle East, including Pakistan [163]. Homozygous allele frequency is reported.

[c]GenomeAsia 100K is a database of Asian individuals, including South Asians [164]. Homozygous allele frequency is reported

Abbreviations: Het = heterozygous, Hom = homozygous

Sanger sequencing was used to determine the segregation of the four candidate variants in the parents (III-2 and IV-1), an unaffected brother (V-5), and an unaffected sister (V-1). The variants in *PDZD7, RMB20*, and *CNTNAP2* co-segregated with disease status, occurring in a homozygous state in the two affected brothers and a heterozygous state in the four other family members (**Figure 4.2B, Table 4.3**). The *ALG6* variant was homozygous in unaffected sibling (V-1), indicating that it was unlikely to be pathogenic.

4.4.2.2 Family 2

Gene panel testing in the two affected sisters (III-4 and III-5) revealed that both siblings were homozygous for the same *CNTNAP2* c.682G>A (p.Gly228Arg) variant that was first found in the affected brothers in Family 1. Sanger sequencing for this variant in the other available family members confirmed co-segregation with disease status, with it only being detected in the homozygous state in the affected siblings. As expected, each parent (II-10 and II-11) was heterozygous for this variant, and it was absent in the two unaffected siblings (III-2 and III-6, **Figure 4.3B**). The variants detected in *PDZD7, RMB20*, and *ALG6* in Family 1 were not observed in Family 2 (**Table 4.3**).

4.4.3 Relatedness and Linkage Analysis

Given the rarity of this variant in other populations, and that it was observed twice in the same relatively isolated population, both times as a homozygote only found in affected individuals from a consanguineous family, it seems likely that the *CNTNAP2* p.Gly228Arg variant has been inherited by both Family 1 and 2 from some common ancestor, rather than being the result of two different *de novo* mutational events. Nevertheless, we know Family 1 and 2 are two independent observations of the association between genotype and phenotype as the

affected individuals are different sexes and ages and cannot be the result of the same probands being discovered by two different investigators.

However, given the likelihood that *CNTNAP2* p.Gly228Arg is a shared mutation from a common ancestor, it is important to understand the extent to which the rest of the genome is also shared. IBD analysis via PLINK suggests that no more than 1/32 of the genome is shared IBD between any member of Family 1 and Family 2, with insufficient power to rule out less of the genome being shared. Thus, while these Families could be as close as 2-3rd cousins, there is no evidence that any individual in Family 1 is closely related to any individual in Family 2. As a positive control, within families all members had estimated relatedness consistent with known relatedness. Thus, since overall genomic sharing is quite low between Family 1 and Family 2, we should view the fact two families show the same phenotype-genotype pattern at *CNTNAP2* p.Gly228Arg as fundamentally independent replications.

Although these families are small and only two in number, we can formally perform a linkage analysis. Under the hypothesis that *CNTNAP2* p.Gly228Arg is the only recessive, 100% penetrant (no recombination distance) variant contributing to disease in these families, the likelihood of the observation under this hypothesis in each family is 1.0, and the likelihood of the data under chance segregation is $\frac{1}{4} * \frac{1}{4} * \frac{3}{4} * \frac{3}{4} = \frac{9}{256} = 0.03156$ for each family. Thus, both families provide a LOD score of ~1.45 with a combined LOD of 2.9. This is the maximum possible LOD score for a recessive hypothesis with 2 affected and 2 unaffected sibs in each family.

4.5 Discussion

We first performed genetic analysis on two brothers from a consanguineous family of Afridi tribal heritage who presented with epilepsy, intellectual disability, and autism spectrum disorder. Analysis of 32 regions of homozygosity shared by the affected siblings revealed four

candidate variants: *PDZD7* (p.Glu450del), *ALG6* (p.Glu345Gln), *RBM20* (p.Ser529Arg), and *CNTNAP2* (p.Gly228Arg).

PDZD7 encodes a scaffolding protein expressed in the cilia of inner ear hair cells and photoreceptors [165, 166]. Pathogenic *PDZD7* variants have been identified in individuals with AR non-syndromic hearing loss and it has been suggested that *PDZD7* variants may also modify the severity of Usher syndrome [166-168]. The lack of overlap between the clinical phenotypes associated with *PDZD7* variants and the affected siblings suggests that the p.Glu450del variant is unlikely to be pathogenic. Furthermore, this variant is seen three times in a homozygous state in gnomAD which excludes individuals with severe pediatric-onset illnesses.

ALG6 encodes the enzyme $Man_9GlcNAc_2$-P-P-Dol α-1,3-glucosyltransferase. Patients with pathogenic *ALG6* variants have abnormal accumulation of $Man_9GlcNAc_2$-P-P-Dol and hypoglycosylation of proteins which contribute to autosomal recessive congenital disorder of glycosylation (CDG) Type Ic (or ALG6-CDG) [169, 170]. ALG6-CDG is typically characterized by developmental delay, seizures and speech disabilities [171]. However, an unaffected sibling (V-1) was also homozygous for the *ALG6* p.Glu345Gln variant, indicating that it is unlikely to be pathogenic.

RMB20 encodes the RNA-binding motif protein #20 which is highly expressed in striated muscle and functions as a splicing factor [172, 173]. Pathogenic variants in *RBM20* have been identified in patients with autosomal dominant (AD) dilated cardiomyopathy [174, 175]. The identified p.Ser529Arg variant falls within the RNA recognition motif (RRM) which plays an important role in splicing regulation, and mutations within the RRM have been previously identified in individuals with AD dilated cardiomyopathy [176-179]. The cardiac phenotypes associated with RBM20 variants are inconsistent with the neurological features observed in the affected siblings. However, since this gene has been shown to regulate splicing of some neuronally expressed transcripts, such as *CACNA1C* in which variants have been reported in

some patients with epilepsy, we cannot exclude the possibility that the p.Ser529Arg variant might contribute to the clinical features in the affected siblings from Family 1 [173, 180].

CNTNAP2 encodes the contactin-associated protein-like 2 which is a transmembrane protein that facilitates axo-glial contacts in mature myelinated axons and is expressed throughout the brain with highest expression in the cortex, hippocampus, substantia nigra, interpeduncle nucleus, pontine nucleus and medial mammillary nucleus [31, 181]. The homozygous *CNTNAP2* frameshift variant, c.3709delG, identified in a consanguineous Old-Order Amish family, was the first reported mutation for cortical dysplasia-focal epilepsy syndrome [182]. Zweier and colleagues identified the homozygous splice acceptor variant, c.1671-1G>T, in a consanguineous family with Pakistani ancestry exhibiting Pitt-Hopkins-like syndrome [183]. Additionally, homozygous deletions of various sizes within *CNTNAP2* have also been identified in patients with epilepsy, intellectual disability and/or autism spectrum disorder [184-188]. Segregation analysis in Family 1 demonstrated that the c.785G>A (p.Gly228Arg) variant, located in the first laminin-G domain, only occurred in a homozygous state in the affected siblings, consistent with a recessive mode of inheritance (**Figure 4.2B, Figure 4.4** and **Table 4.3**). Furthermore, this variant was not observed in a homozygous state in gnomAD, GMEVariome, or GenomeAsia100K (**Table 4.3**). Approximately 20% of individuals in gnomAD are of Asian descent and may not fully represent the ancestry of the Families in this study. GMEVariome contains 2,497 unrelated individuals of Middle Eastern ancestry, and GenomeAsia100K currently contains 1,736 unrelated individuals with 113 from Pakistan [163, 164]. Therefore, these two population databases better reflect the ancestry of the Families in this study and provide evidence that the p.Gly228Arg variant in *CNTNAP2* is not seen in a homozygous state within healthy individuals of Asian/Middle Eastern descent.

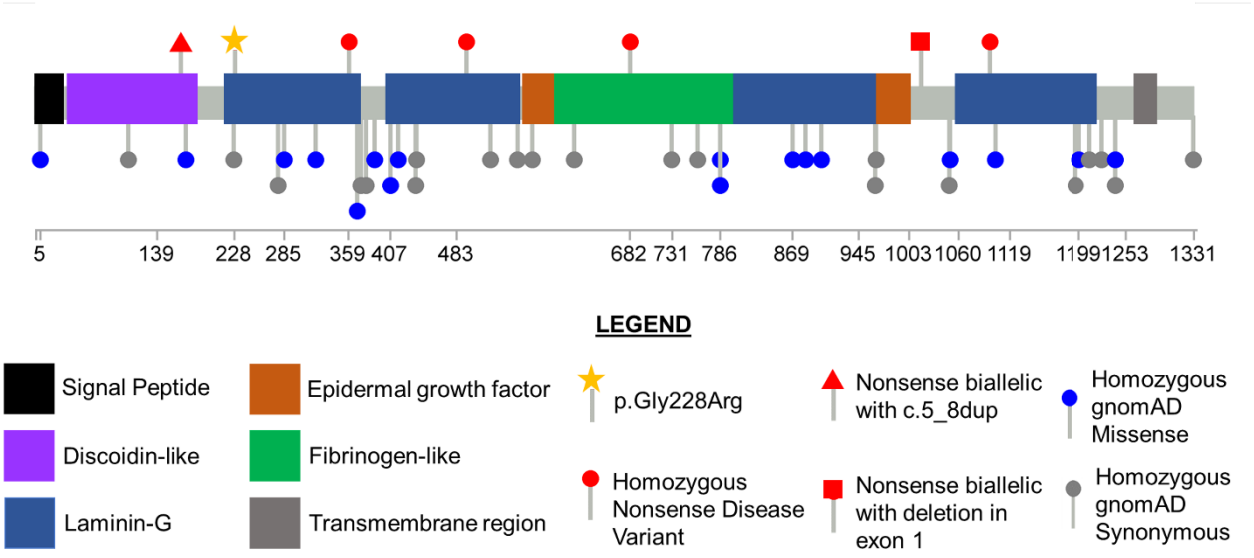


Figure 4.4. Previously Reported Disease Variants in *CNTNAP2*. Lollipop plot of previously reported homozygous or biallelic nonsense disease variants in HGMD (professional version 2021.3) and homozygous variants in gnomAD (v2.1.1). Star indicates novel homozygous p.Gly228Arg variant identified in this study. Domain and variant locations based on UniProt accession Q9UHC6.

During the course of this study, we became aware of a second consanguineous family of Pakistani origin (Family 2) in which the two affected siblings were found to be also homozygous for the c.785G>A (p.Gly228Arg) variant in *CNTNAP2*. Clinical presentation in the two affected sisters in Family 2 showed striking overlap with the affected siblings from Family 1, with all four affected individuals exhibiting epilepsy, intellectual disability, and aggressive behaviors (**Table 4.1**). Segregation of the *CNTNAP2* variant within Family 2 confirmed co-segregation with disease status (**Figure 2B** and **2D**). The additional candidate variants observed in Family 1 were not detected in Family 2, providing additional support for the pathogenicity of the *CNTNAP2* variant (**Table 4.3**). Together these families give a LOD score of 2.9 in support of p.Gly228Arg being a recessive, fully penetrant disease-causing variant.

Prior to the current study, all reported disease-associated *CNTNAP2* variants were homozygous or biallelic deletions, splice site variants, or nonsense variants which would be predicted to abolish protein function [182, 183, 186, 187]. Mice lacking *Cntnap2* (*Cntnap2*$^{-/-}$) recapitulate the seizure and core autism-related behavioral phenotypes seen in patients and reveal a number of pathophysiological alterations including neuronal migration abnormalities, asynchronous neuron firing patterns, and a reduced number of GABAergic interneurons [189]. Recent work in *in vitro* cell culture models has shown that loss of one *CNTNAP2* allele is sufficient to impact axonal growth and heterozygous missense variants identified in patients with autism were associated with deficits in protein trafficking [31].

In addition to epilepsy and intellectual disability, there are other striking similarities between other clinical features of the affected siblings in the current families and previous reports (**Table 4.4**). For example, Smogavec and colleagues described eight patients with homozygous *CNTNAP2* variants of which six patients were non-verbal or had very limited speech ability [186]. Additionally, two patients exhibited aggressive behaviors and three were reported to have notable temper tantrums [186].

	This Study	Strauss [182]	Gregor [188]	Smogavec [186]
Current Age	5-13 years	NA	NA	NA
Epilepsy Syndrome	3/4 Focal Onset Epilepsy	Focal Epilepsy	5/7 with epilepsy	Complex Focal, Generalized Epilepsy
Age of Seizure Onset	2.5-3 years	14-20 months	3 months – 3 years	5 months – 3 years
Primary Seizure Type	2/4 – Focal 1/4 – Generalized 1/4 – Secondary Generalized	9/9 - Focal Onset Aware 4-9 – Secondary generalized	Complex	ND
Current AEDs	Valproic Acid (3/4), Carbamazepine (3/4), Clonazepam (2/4), Diazepam (1/4), Risperidone (1/4), Levetiracetam (1/4), Methylphenidate (1/4), Lamotrigine (1/4)	ND	ND	Lamotrigine (3/8), Cabamazepine (2/8), Valproate (2/8), Levetiracetam (1/8), Oxcarbazepine (1/8), Phenobarbitone (1/8), Zonisamide (1/8)
Seizures Controlled?	3/4 – No, 1/4 – with breakthrough seizures	ND	2/5 – No	2/8 - No
Intellectual Disability	4/4	9/9	7/7 (Moderate to Severe)	8/8 (Moderate to Severe)
Developmental Delay	4/4 – with speech regression/poor speech	9/9 - with speech regression	7/7 4/7 – No speech 1/7 – Simple Sentences	6/8 – Less than 10 words 2/8 – Simple sentences
Autism spectrum disorder	2/4 – Yes 2/4 – Poor social abilities	6/9	ND	ND
EEG	Frequent brief generalized high amplitude SWDs (1/4), frequent left frontotemporal high amplitude discharges (1/4), interictal frontotemporal epileptiform abnormalities (1/4)	7/9 – normal background, seizures arising from temporal (occasionally frontal) regions with unilateral high-amplitude spike-slow-wave discharges or focal slowing	ND	Generalized slowing (1/8), occasional spikes and slowing left temporal (1/8), frequent left frontotemporal epileptiform discharges (1/8), slow rhythm, sometimes with epileptic discharges (2/8), diffuse cerebral dysfunction (1/8)
Other Behavioral Features	4/4 – Aggression	ND	1/7 – Autoagressive behavior	2/8 – Aggression 2/8 – Notable temper tantrums
Other Features	2/4 – Body movement imbalance 2/4 – stereotypic hand flapping movements	ND	1/7 - Hypotonia	2/8 – Hypotonia 2/8 - Ataxia

Table 4.4. Comparison of Patients Reported in this Study to Previously Reported Patients. Abbreviations: NA = not applicable, ND = no data, SWD = spike-wave discharges

Taken together, the genetic analysis of the two families and clinical similarities between the affected siblings and other patients with *CNTNAP2* variants strongly support that the identified p.Gly228Arg variant is likely causative. Functional analysis of this variant will be required to establish the impact on CNTNAP2 function and the underlying disease mechanism. This is the first demonstration of co-segregation of a homozygous *CNTNAP2* missense variant within an affected pedigree. Further work to functionally evaluate both homozygous and heterozygous missense *CNTNAP2* variants will aid in the identification of genotype-phenotype correlations and may contribute to the development of more efficacious treatments for these patients.

4.6 Acknowledgements

This study was supported in part by the Emory Integrated Genomics Core (EIGC), which is subsidized by the Emory University School of Medicine and is one of the Emory Integrated Core Facilities. We thank the family and clinicians for their participation in this study. This work was partially supported by an International Research Support Initiative Program award from the Higher Education Commission of Pakistan to N.B. Additional support was provided by the Georgia Clinical & Translational Science Alliance of the National Institutes of Health [UL1TR002378].

Chapter 5

Discussion

5.1 Summary

The overarching goal of my dissertation research was to gain a better understanding of the genetic contributions to epilepsy. I did this, in part, by using existing sequence data from epilepsy patients that had undergone gene panel or WES testing at EGL Genetics. As shown in **Chapters 2** and **3,** along with Kameryn Butler, a former graduate student in the Escayg lab, I was able to utilize this source of data in different ways. As detailed in **Chapter 2,** we identified eight patients with *SLC6A1* variants out of 460 patients analyzed. This large number of patients identified with *SLC6A1* variants was surprising as *SCN8A* and *CDKL5,* both well-established epilepsy genes, had similar diagnostic yields from the same dataset. We showed that the identified *SLC6A1* variants lead to decreased GABA re-uptake (**Figure 5.1A**). This was the first-time patient *SLC6A1* variants were tested for an impact on protein function, and our study identified a possible disease mechanism. In contrast, as outlined in **Chapter 3,** a single *ATP6V0C* variant identified from WES data led us to establish new collaborations which resulted in the largest cohort of patients with variants in *ATP6V0C*. Through functional testing, I was able to demonstrate that these variants result in decreased V-ATPase function (**Figure 5.1B**) and I hypothesized that missense variants likely act via a dominant-negative mechanism.

In **Chapter 4,** I used an unbiased screening approach to identify a likely pathogenic variant**.** In that study, we performed WGS and used ROH and segregation analysis to identify likely causative variants in two brothers with epilepsy, intellectual disability, and autism spectrum disorder from a consanguineous Pakistani family. A homozygous p.G228R variant in *CNTNAP2* was seen in both brothers and was not present in a homozygous state in other unaffected family members. A second family with two similarly affected sisters carrying the same *CNTNAP2* variant was subsequently identified providing further evidence for the pathogenicity of this recessive variant (**Figure 5.1C**).

In this chapter I will expand upon the previously presented topics in my dissertation, provide future directions for work relating to *ATP6V0C* and epilepsy, and discuss ideas for how

genetic testing can continue to evolve in order to find answers for those with suspected genetic disease(s).

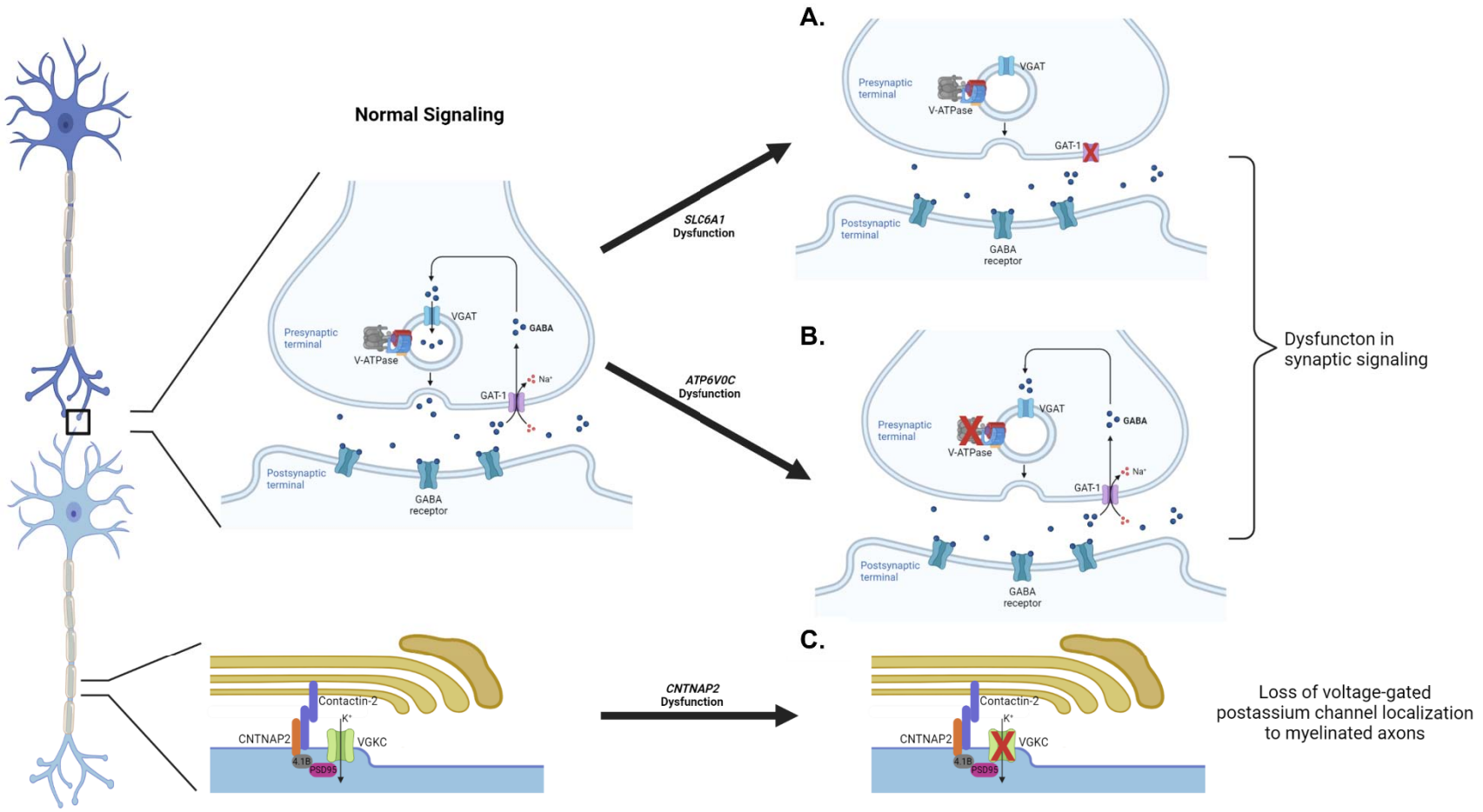


Figure 5.1. Overview of the Effect of the Identified Variants on Neuronal Signaling. A) In **Chapter 2**, *SLC6A1* variants were shown to decrease GABA uptake by GAT-1. **B**) In **Chapter 3**, *ATP6V0C* variants were shown to decrease the function of the V-ATPase which is predicted to lead to a loss of neurotransmitter loading into synaptic vesicles. **C**) In **Chapter 4**, a homozygous variant in *CNTNAP2* was identified. Loss of *CNTNAP2* has been shown to lead to a loss of voltage-gated potassium channels (VGKCs) on myelinated axons [32]. Figure created with BioRender.com.

5.2 Considerations in Genetic Testing and Variant Evaluation

When deciding which genetic test(s) should be used to screen individuals with suspected genetic disorders, there are many components, often antagonistic of each other, that must be considered. A clinician and patient must take several factors into consideration: such as turnaround time, the likelihood that the individual has a genetic disorder, and, especially in the United States, whether the cost of the test will be covered by insurance. Discussion of the major considerations that must be accounted for in genetic testing and variant evaluation is provided below.

5.2.1 Types of Tests Used

As discussed in **Chapter 1**, there are several different types of tests that can be used to achieve a genetic diagnosis. Each test has its pros and cons related to the balance between the likelihood of finding a causative variant vs. the number of variants that must be analyzed/curated before results can be returned to the patient. The phenotypic presentation of the individual undergoing testing as well as a history of disease in their family can help provide insight as to what test is best to use. If an individual has a very specific phenotype for a well-established genetic disorder, then performing targeted gene sequencing or using a gene panel can avoid the cost and long-turnaround time associated with WES/WGS. Further, if the patient has an affected family member with the same, or a similar, phenotype with a known genetic diagnosis, targeted sequencing for the specific variant can be performed as the first-tier genetic test. On the other hand, if the phenotype does not fit a particular diagnosis, then the more unbiased approaches of WES or WGS would increase the chance of finding a causative variant. While WES and WGS have a longer turnaround time, if they are the only genetic test needed it leads to a shorter overall diagnostic timeline.

5.2.2 Burden for Pathogenicity

In **Chapter 1** I described the ACMG/AMP criteria for variant classification and how it is used to classify variants on a spectrum from pathogenic to benign. With these criteria there is a much higher burden to classify a variant as pathogenic compared to benign (**Figure 1.1**). This likely reflects a cautious approach to apply a label that can feel as definitive as “pathogenic” or “likely pathogenic” to a variant and report it back to an affected individual, but one must also be equally as cautious to call a variant (likely) benign and potentially miss an important diagnosis. Many of the variant classification criteria can be assigned by obtaining an extensive family history of disease, using publicly available datasets to score variants, and identifying whether the variant has been previously associated with disease. Functional testing may also be used to provide support for variant pathogenicity, and it is the only criteria that involves the direct testing of a variant *in vitro* or *in vivo*.

In **Chapters 2** and **3** of this dissertation, I demonstrated the utility of functional analysis for the classification of variants found through standard genetic testing. In both chapters, my work was the first to report functional effects of variants identified in patients with epilepsy on protein function. This work was important in establishing likely disease mechanisms and will guide future work on therapeutic interventions.

While functional testing is an important piece of information that can be used to classify variants, it is primarily used when a VUS in a known gene is identified or if a GUS is identified with strong animal model and expression data to support the effort to undertake functional testing. This leads to the question of what if there is a VUS in a GUS?

5.2.3 Ignoring the “Unknown”

When VUSs are identified, the ACMG/AMP classification guidelines provide a framework for evaluation and potential reclassification, but when variants are identified in a GUS there is

less guidance on what to do. Some classification criteria, such as PVS1, PM1, BP1, and BP3, even rely on the function of the gene being known to apply the evidence to a variant (**Table 1.2**). Variants in GUSs are likely to be classified as VUSs until research into the function of a gene is performed.

Additionally, the knowledge of what genes are already associated with a disease can bias variant analysis. This can come in the form of what genes are included on a gene panel or whether variants in known genes for a particular disease are evaluated first when WES or WGS is performed. By focusing on the "low-hanging fruit" of known disease genes, one might miss a variant in a GUS that is contributing to the phenotype. To avoid biases that come from evaluating known disease genes first, the criteria used to evaluate variants can be shifted. Instead of looking at whether a variant occurs in a known disease gene, or not, one can focus instead on factors that can be evaluated separately from gene function. By using MAF thresholds, CADD scores, and starting with exonic or splice site variants, variants that are likely to be pathogenic are identified irrespective of the gene they are in. I employed this type of unbiased variant analysis in **Chapter 4** and **Appendix B** when evaluating WGS data from eight consanguineous families with epilepsy.

5.3 How do we Find the "Missing" Variants?

One of the most important future directions regarding genetic testing in epilepsy is finding the "missing" variants. In epilepsy, even when generally unbiased approaches like WES and WGS are used, diagnostic rates are still only 20% - 50% indicating there are still variants and other factors at play that have yet to be identified. In the following subsections I outline the three main areas I believe should be explored to find missing pathogenic variants.

5.3.1 Non-coding Variants

The vast majority of the human genome does not code for proteins. Previously it was thought that this non-coding portion of the genome was just “junk” and provided little to no function; however, our understanding of the functions of the non-coding genome continues to improve with advances in sequencing and analysis technologies. The idea that “junk” DNA is functional, paired with the increase of WGS in genetic testing provides the opportunity to look at the contribution of non-coding variants in patients with epilepsy. In addition, use of datasets, such as ChIP-seq and ATAC-seq can provide information about the temporal and spatial expression of regions of the genome containing non-coding variants. Lastly, as the function of non-coding variants is unlikely to be previously known, use of functional testing as described in **Chapter 1** will be an important part of the process to verify that identified non-coding variants have an effect that would explain the patient’s phenotype.

5.3.1.1 Untranslated Regions

Untranslated regions (UTRs) occur at both the 5’ and 3’ of genes and are defined as regions of mRNA that are not translated into protein. UTRs play an important role in mRNA stability, secondary structure, localization, and translation. Using WES and gene panels, certain regions of UTRs may be screened as they are considered exonic sequences, but WGS would provide full coverage of UTRs. The consequence of a variant in an UTR might not be easy to determine, and functional assays to look at translation and/or mRNA stability will be important to provide evidence for pathogenicity. Doing these functional assays in biologically relevant systems is especially important given the likely interplay between cell or tissue specific factors and translation.

5.3.1.2 DNA Regulatory Elements

Screening for pathogenic variants in DNA regulatory elements such as promoters, enhancers, silencers, and insulators should also be done when WGS data is available. With the increasing availability of publicly available ChIP-seq and ATAC-seq datasets, the location of these regulatory elements can be readily predicted, and variants located within these regions can be assayed for a functional impact.

Promoters are located near the transcription start site (generally within 500bp) and recruit transcription factors along with RNA polymerase to initiate transcription. Enhancers help with RNA polymerase and transcription staging and can be in close proximity to promoters depending on the 3D chromosomal organization providing either temporal or spatial transcription specificity. Silencers and insulators function to silence gene transcription. Variants that fall within any of these regions could impact transcriptional efficiency and/or lead to mRNA expression at the wrong time or place.

5.3.1.3 non-coding RNAs

The last major type of non-coding variants that should be prioritized for analysis are those that fall within non-coding RNAs (ncRNAs). Micro RNAs (miRNAs), which inhibit mRNA expression by binding the 3’ UTR of a mRNA, are the most studied ncRNAs to date. Studies have shown that miRNAs play an important role in epilepsy, influencing processes such as neuroinflammation, apoptosis, glial cell dysfunction, and oxidative stress [190]. In addition to miRNAs, long non-coding RNAs (lncRNA) should be looked at. lncRNAs are ncRNAs that are greater than 200bp in length and have been shown to be involved in various cellular processes including transcription and facilitating protein-protein interactions [191].

5.3.1.4 Therapeutic Strategies for Targeting Non-coding Variants

The identification of non-coding variants also provides the opportunity for precision therapy. ASOs as discussed in **Chapter 1** provide the ability to modulate mRNA and subsequent protein expression in a way that restores more normal function based on the known disease mechanism. ASOs could be used to specifically target mutant transcripts for degradation, thereby preventing aberrant expression of transcripts resulting from pathogenic UTR variants. The use of deactivated (or nuclease deficient) Cas9 (dCas9) could be used when variants are located within DNA regulatory elements. Functional assays would be necessary to determine whether the pathogenic variant results in an increase or decrease of transcription and/or in transcription that is taking place at the wrong time or place. Once this is known, dCas9 can be fused with a protein that either activates or represses transcription and dCas9 would be able to "deliver" this protein in a sequence specific manner [192]. Strategies utilizing dCas9 fused to a transcriptional activator (dCas9-VP160) to increase *Scn1a* activation have been successful in reducing hyperthermia-induced seizures in *Scn1a*$^{+/-}$ mice [193].

5.3.2 Alternative and Poison Exons

Alternative splicing is a mechanism by which multiple proteins can be produced from one gene and can occur in a tissue- or developmental stage-specific manner. Screening for variants in non-canonical or alternative exons may identify new pathogenic variants in known epilepsy genes. For example, *SCN8A* encodes transcripts with either a neonatal or adult version of exon 5, and variants in the neonatal exon have been identified in patients with epilepsy [194]. This highlights the importance of routinely screening alternative and non-canonical exons in genetic testing.

Poison Exons (PEs) are a special class of alternative exons that contain a premature stop codon leading to NMD of the transcript. PEs are typically stretches of conserved sequence

within an intron and smaller in size than a regularly spliced exon [195]. Identifying and screening PEs for pathogenic variants will be an important future direction when individuals undergo WGS. The utility of analyzing PEs has already been demonstrated for two epilepsy genes, *FLNA* and *SCN1A*, in which pathogenic variants result in the inclusion of a PE at a higher frequency than normal and lead to haploinsufficiency [196, 197].

PEs also make compelling targets for treatments. In **Chapter 1**, I described how the use of an ASO in *Scn1a*$^{+/-}$ mice was able to reduce seizure frequency and increase survival by preventing the inclusion of a PE in the otherwise wild-type allele thereby increasing wild type *Scn1a* mRNA levels [51].

5.3.3 Mosaic Variants

Mosaic variants are often overlooked because they can be difficult to detect through traditional sequencing techniques for a variety of reasons. Mosaic variants arise from a mutational event that occurs during development, and the developmental stage at which this event occurs can greatly impact the number of cells affected. Variants that arise prior to blastulation are likely to affect many tissues throughout the body, but if the mutational event occurs after gastrulation it is possible that only one germ layer, or even just one tissue, is affected. The timing of the mutational event also has consequences for the ability of sequencing to detect mosaic variants. If the gDNA source from a patient escaped the mutational event, the variant may only be present at very low levels, if at all. This can make it difficult to decern if the variant called is real or a sequencing artifact. Being able to use gDNA isolated from the affected tissue leads to the best possible chance of detecting mosaic variants; however, in patients with epilepsy this is particularly challenging as brain tissue is not accessible outside of surgical procedures. The use of deep and ultra-deep coverage along with trained and validated variant calling programs can detect low levels of somatic mosaicism [198]. However, it is currently not

cost effective to sequence every individual at the >250x coverage needed to detect low levels of mosaicism.

To further complicate the detection of mosaic variants and classification of variant pathogenicity, there is evidence that some forms of epilepsy are caused by “two-hits” with one variant being germline and the other somatic mosaic. The two variants can either be in the same gene, within the same pathway, or in pathways that interact with each other. In epilepsy with focal cortical dysplasia (FCD), this has been shown to be true. Pathogenic variants in *DEPDC5* are associated with refractory focal epilepsy, but in some patients FCD is also seen. Screening of brain tissue (resected or post-mortem) from these patients revealed a second, mosaic variant in *DEPDC5* [199]. In a separate study, patients with epilepsy and FCD were found to have germline *NPRL3* variants and somatic mosaic variants in *WNT2* [200]. NPRL3 and WNT2 are both involved in WNT signaling which has a well-established role in cell migration and neural patterning during embryonic development [201].

5.4 Future Directions for *ATP6V0C* and Epilepsy

While my dissertation research encompassed several different genes, the primary focus of my work was uncovering the role of *ATP6V0C* in epilepsy and neurodevelopmental disorders. Additionally, *ATP6V0C* is a new epilepsy gene providing substantial opportunities for further research. Below I outline three major areas that I believe should pursued in regard to *ATP6V0C*-derived epilepsy and the impacts they would have for patients and potentially other disorders caused by similar mechanisms.

5.4.1 Mechanistic Studies

In **Chapter 3**, I proposed that missense *ATP6V0C* variants, along with those that escape NMD, likely act via a dominant-negative mechanism. This hypothesis was based on modeling

data which suggested that the identified patient variants were likely to interfere with the rotary mechanism of the c-ring and its interface with ATP6V0A during the process of proton translocation. Further support for this proposed mechanism is provided by the knowledge that known V-ATPase inhibitors, such as bafilomycin and archazolid, have binding sites within the c-ring that create steric hinderance effectively blocking the transfer of protons from the c-ring to ATP6V0A [202, 203]. While there is strong evidence for a dominant-negative mechanism, providing experimental evidence for this mechanism will be an important next step to guide the development of precision therapy strategies.

Experimental evidence for this mechanism can be provided via growth assays in yeast. To date, we have only used haploid yeast where the endogenous copy of *VMA3* has been deleted. In order to look at a dominant-negative effect, both wild-type and mutant copies of *VMA3* need to be present. By placing the patient variants under an inducible promoter, such as GAL or TET0, and transforming plasmids into a diploid strain in which only one endogenous copy of wild-type *VMA3* is present, we can look for a difference in growth when the inducible promoter is on vs. off. Seeing a decrease in growth/V-ATPase function when a patient variant is expressed would support a dominant-negative effect.

In addition to determining the mechanism by which patient variants decrease V-ATPase function, it is also important to understand how this decrease in V-ATPase function leads to the clinical phenotypes observed in patients. In **Chapter 3** I proposed two mechanisms for how V-ATPase dysfunction leads to neurodevelopmental disorders. The first is through a loss of synaptic signaling and the second is through dysfunction in autophagy.

A loss of synaptic signaling is likely given that the V-ATPase creates the necessary proton gradient for both VGLUTs and VGAT to load glutamate and GABA, respectively, into SVs [142, 143]. It has been previously shown that loss of V-ATPase function leads to an overall decrease in neurotransmitter content in SVs and synapse formation [101]. Additionally, V-ATPase driven acidification within the TGN is important for proper trafficking and sorting of

various membrane bound proteins. Inhibition of this acidification may also lead to a loss of synaptic signaling through improper trafficking of neurotransmitter and neuropeptide receptors [149].

Looking at the uptake of both glutamate and GABA in mammalian cell culture, either in differentiated neuroblastoma cell lines expressing different variants, induced pluripotent stem cell lines from patients, or primary neurons from genetic mouse models, would provide insight into perturbations of SV loading and synapse formation related to decreased V-ATPase function. Additionally, looking for the relative abundance presence of glutamate and GABA receptors at the post-synaptic neuron, and if they are at the correct type (i.e., excitatory or inhibitory) of synapse could provide important insights to the neurological phenotypes seen in patients with *ATP6V0C* variants.

The second proposed mechanism is that decreased V-ATPase function negatively impacts autophagy, an important cellular process that degrades and recycles various cellular components. The V-ATPase plays a critical role in acidifying organelles, such as lysosomes and endosomes, which allows for cleavage of propeptides into functional enzymes [151]. Previous studies have shown that loss of V-ATPase function prevents fusion of lysosomes with autophagosomes which is a critical step in protein degradation [152]. Further, defects in autophagy are associated with pathogenic variants in genes encoding for other subunits of the V-ATPase. For example, variants in *ATP6V1A* that cause epileptic encephalopathy result in decreased expression of LAMP1 (a marker of lysosomes) indicating a decrease in overall number of lysosomes and reduce the recruitment of the V-ATPase to phagosomal structures [100]. Similar defects in autophagy were seen in *Atp6v0a1*$^{A512P/A512P}$ mice with accumulation of undegraded autophagosomes and lysosomes in the hippocampus due to a lack of lysosomal acidification [101]. Failure to properly and efficiently degrade cellular waste can be especially detrimental in post-mitotic neurons as it will lead to accumulation of toxic proteins and/or damaged organelles [153]. Assessing intracellular pH of various organelles as well the

accumulation of undegraded autophagosomes/lysosomes will be an important step in determining if autophagy is impaired with ATP6V0C dysfunction.

Taken together, understanding both the genetic and functional mechanisms behind ATP6V0C dysfunction and disease will provide the opportunity to identify and/or develop specific therapies for patients.

5.4.2 Therapeutic Strategies

As discussed in **Chapter 1** the development of precision therapies for epilepsy is an area of high interest. Many of the strategies discussed involve increasing or decreasing expression of the wild-type or mutant allele, respectively. For *ATP6V0C* precision therapy strategies are less straightforward as haploinsufficiency is a known disease mechanism and there is strong evidence that a dominant-negative mechanism is also likely for missense variants. Patients with haploinsufficiency of *ATP6V0C* caused by a 16p13.3 microdeletion do not show any significant difference in clinical presentation with respect to age of seizure onset when compared to patients with germline missense or frameshifting variants (19.75 ± 3.25 months vs. 13.47 ± 2.25 months). Therefore, strategies to decrease mutant transcript levels in patients with missense variants are unlikely to lead to a less severe phenotype and would need to be combined with strategies to also increase wild-type transcript levels.

5.4.2.1 Overexpression of wild-type ATP6V0C

It is possible that increasing wild-type levels of *ATP6V0C* may be therapeutic either by rescuing the haploinsufficiency or by potentially outcompeting mutant subunits during V-ATPase formation. One obvious concern with increasing wild-type *ATP6V0C* levels is whether too much ATP6V0C is pathogenic, but there is evidence to suggest that an increase in wild-type ATP6V0C would not be pathogenic. First, one patient with generalized atonic seizures has been

reported with a duplication of *ATP6V0C* and 117 other genes encompassing a 16 Mb region [204]. Two additional patients with similar phenotypes were previously reported with duplications of 16p13.3-p12.3; however, neither of these duplications included *ATP6V0C* [205, 206]. The minimal overlapping region of these patients included *GRIN2A,* a known epilepsy gene, and the three patients with 16p13.3-p12.3 phenocopied patients with pathogenic missense *GRIN2A* variants [204-206]. Therefore, it was determined that the duplication of *GRIN2A* underlies the epilepsy phenotype seen in these patients. Secondly, when establishing the yeast system to functionally evaluate patient variants, I transformed a plasmid that constitutively expressed *VMA3* in a wild-type yeast strain. Spotting assays showed no difference in growth after 48 hours compared to the wild-type strain transformed with the empty plasmid suggesting that increased VMA3 is not detrimental to overall fitness of yeast (**Figure 5.2**). Lastly, ATP6V0C is part of a multisubunit complex. So, increasing wild-type ATP6V0C might not result in an overall increase of V-ATPase formation or activity since the protein levels of the other subunits needed to form a functional V-ATPase would not be altered. Instead, it would provide additional wild-type ATP6V0C to form functional V-ATPase complexes. The idea that overexpression of any subunit of the V-ATPase is not pathogenic is further supported by a review of the literature and entries in HGMD and ClinGen that revealed no pathogenic duplications, in patients with known phenotypes related to V-ATPase dysfunction, involving any genes that encode for a subunit in the V-ATPase.

Similar to ATP6V0C, GABRG2 is one subunit of the larger $GABA_A$ receptor and a nonsense variant that escapes NMD (p.Q390X) was identified in a family with GEFS+ [207]. In a *Gabrg2*$^{+/Q390X}$ knock-in mouse overexpression of wild-type *Gabrg2* was able to increase seizure resistance and survival after pentylenetetrazol (PTZ) administration [208]. This study demonstrates overexpression of a wild-type subunit can have therapeutic benefits and is an area that should be explored as a treatment option in patients with pathogenic *ATP6V0C* variants.

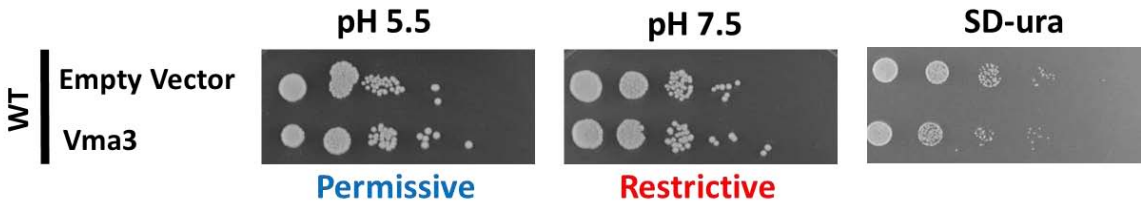


Figure 5.2. Expression of Vma3 in a Wild-Type Yeast Strain Does Not Affect Growth. A centromeric plasmid containing a constitutive GAPDH promoter with the multiple cloning sites intact (Empty Vector) or expressing the wild-type *VMA3* ORF (Vma3). $1OD_{600}$ of overnight cultures was collected and a serial dilution assay was performed on three types of media: YPD at pH 5.5, YPD at pH 7.5, and SD-ura. YPD at pH 7.5 selects for V-ATPase function.

5.4.2.2 Gene Editing

Another strategy that could be therapeutic in patients with missense or frameshift variants that escape NMD is direct editing of *ATP6V0C* to revert it to the wild-type sequence. Direct editing of the genome is possible through use of CRISPR-Cas9, but there are many caveats that must be addressed for it to be an efficacious treatment. Specifically, for *ATP6V0C*, there are four known pseudogenes [209]. Depending on the location a variant, efficient targeting and editing of *ATP6V0C* might be difficult if the region is conserved between any of the pseudogenes. As with any genome editing there is the inherent risk of off-target editing which could result in pathogenic variants in a new location of the genome. However, as gene editing to revert pathogenic variants to wild-type is likely to be performed after birth, off target effects could be localized to specific tissues through the use of tissue specific promoters to control expression of different components needed for CRISPR-Cas9 editing.

5.4.2.3 High Throughput Screening of AEDs and Novel Compounds

Another avenue to develop treatments for patients with ATP6V0C (and V-ATPase) dysfunction would be to convert the yeast assays into a tool for high throughput screening of known AEDs as well as novel compounds. As demonstrated in **Chapter 3**, expression of patient variants results in decreased growth in the presence of $CaCl_2$ and decreased LysoSensor fluorescence compared to rescue of a *vma3Δ* strain with wild-type *VMA3*. This could provide a powerful tool for the initial screening of compounds, especially considering that V-ATPase acidifies organelles in yeast and that the autophagy pathway is well conserved between yeast and humans [210]. While growth at different $CaCl_2$ concentrations could be used in a screen, taking advantage of pH sensitive probes to measure intraorganellar pH offers the opportunity for a higher throughput screen. However, initial screening could also be done using the *C. elegans* strains developed in **Chapter 3.** Previous work has demonstrated the ability to perform high

throughput drug screens in *C. elegans* [211]. Regardless of what system is used, compounds that show rescue of V-ATPase function in initial screens would need to be further verified in more physiologically relevant models, which would allow for assessment of any changes to neurological phenotypes after treatment.

5.4.3 Longitudinal Studies

Another area that I feel is important to explore is the natural course of disease in patients with ATP6V0C dysfunction. Now that a cohort of 26 patients has been identified, it would be informative to follow them in a longitudinal study. Our cohort consists of patients that are primarily under the age of 18 and prior to our study many clinicians did not know the phenotype(s) of patients with pathogenic *ATP6V0C* variants. Understanding the natural course of disease can have important implications for understanding treatment efficacy and the healthcare resources needed [212].

Additionally, it is important to understand what, if any, late-onset phenotypes might be seen in patients. Other organ systems might become affected later in life given that ATP6V0C is a ubiquitously expressed subunit of the V-ATPase, and other subunits of the V-ATPase have been associated with conditions other than neurological disease (**Table A.1**). Based on the presentation of those patients it would be of particular interest to look at the development of kidney issues, particularly tubular acidosis and kidney stones, and increased bone density (osteopetrosis). If these organ systems appear to be unaffected in patients with *ATP6V0C* variants, then it would be interesting to look further into why they are not affected given the ubiquitous expression of *ATP6V0C* and the V-ATPase.

5.5 Future Directions for Genetic Testing

In addition to future directions specific to my work on *ATP6V0C*, I believe there are more general areas that are of particular interest related to genetic testing. These future directions will result in more robust datasets and understanding of how variants impact function and aid in the future classification of variants identified through genetic testing.

5.5.1 Gene-specific Databases

One area of great interest is the development and curation of gene-specific databases which would provide information such as functional effect of a particular variant, known phenotypes of patients with specific variants, and predicted (or known) disease mechanism. These databases will be useful to researchers developing therapies, to genetic testing labs that may be evaluating variants for pathogenicity, and to clinicians who could use this information to compare the phenotype of their patient to other patients with known variants. Such databases already exist for *SCN1A* and other sodium channels (scn-portal.broadinstitute.org) and are in development for *SLC6A1* (slc6a1-portal.broadinstitute.org).

In order to increase the utility of these databases, saturation mutagenesis should be performed and data from functional analysis of these variants added to the databases. This will not only be a valuable tool for classification of variants during genetic testing but may provide insights into items such as the minimum level of protein function needed for normal function, and whether variants have different pathogenic effects on protein function. Both of these pieces of information can be further used when evaluating potential therapies and whether one approach will work for all patients. The yeast system described in **Chapter 3**, could be used for a such an effort to characterize functional effects of all possible variants in ATP6V0C as well as other subunits of the V-ATPase.

5.5.2 Reevaluation of data

Another area that is important regarding genetic testing data is reevaluation of variants, especially VUSs and GUSs, and I believe as sequencing of non-coding variants becomes more commonplace, reevaluation will become even more important. Reevaluation allows for the inclusion of new knowledge on gene function and how specific variants might impact function. Studies where variants have been reevaluated found that in approximately 10% of cases, they were able to reclassify a variant based on new information [213, 214]. The timing of this reevaluation can vary based on individual testing site policies, but in general, most sites wait 12-24 months after initial sequencing to reevaluate any data.

I believe that 12-24 months after initial sequencing is an ideal timeframe for reevaluation of variants as it balances the likelihood of new information being published, the amount of people needed (and certified) to clinically reclassify variants, and the ability to provide answers to patients, their families, and the clinical team. Reevaluating variants earlier than 12 months limits the amount of new information that is available and would likely require an increased workforce to keep up with variant reclassification. However, waiting too long after initial testing to reevaluate variants can delay the start of precision therapies that have been developed and may leave patients and families without answers for longer than necessary.

5.5.3 Long-read WGS

Just as gene panels replaced linkage analysis and targeted sequencing, and WES has replaced gene panels, WGS, especially long-read WGS, will become the new “gold standard” for genetic testing. A limitation of WGS (and WES) is that the current Illumina NGS platforms that are predominantly used have average read lengths of 300bp. This read length is too short to capture most structural variants longer than 50bp and creates difficulty with calling and mapping repetitive regions of the genome (e.g., telomeres and centromeres). Long-read

sequencing can overcome these challenges with read lengths starting at 5Kb (5000bp) and the ability to go up to several megabases in length [215]. Therefore, long-read WGS has the potential to identify point variants and also call CNVs, thus replacing the need to use both WES/WGS and aCGH in diagnostic testing.

There are still several items that need to be addressed before long-read WGS becomes the standard sequencing strategy. The first is the limited number of existing genomes sequenced via long-read WGS that can be used as reference genomes. When mapping to current “short-read” reference genomes there is always the possibility that sequence from long-read WGS will be called as sequencing artifact because that particular region was inaccurately called using current methods. This also relates to the need for new mapping and variant calling algorithms to process long-read WGS data. The need for new algorithms is due in part to the longer read lengths, but also to deal with the fact repetitive regions will be captured and as such different quality control metrics are needed to avoid accidently calling these regions as sequencing artifacts/errors.

5.6 Overall Conclusions

In conclusion, the data presented in this dissertation highlight the diverse nature of epilepsy genes, but also the various ways in which one can identify novel genetic causes of epilepsy. In addition, I demonstrated the power genetic testing provides to 1) identify disease causing genes/variants, 2) understand disease mechanisms, and 3) develop precision therapies for patients with epilepsy. The techniques and strategies used can be applied to future studies on the genetics of epilepsy but can also be applied more broadly to other genetic disorders.

References

1. ILAE. Guidelines for epidemiologic studies on epilepsy. Commission on Epidemiology and Prognosis, International League Against Epilepsy. Epilepsia. 1993;34(4):592-6. Epub 1993/07/01. doi: 10.1111/j.1528-1157.1993.tb00433.x. PubMed PMID: 8330566.

2. Zack MM, Kobau R. National and State Estimates of the Numbers of Adults and Children with Active Epilepsy - United States, 2015. MMWR Morb Mortal Wkly Rep. 2017;66(31):821-5. Epub 2017/08/11. doi: 10.15585/mmwr.mm6631a1. PubMed PMID: 28796763; PMCID: PMC5687788.

3. WHO. Epilepsy [Internet]. World Health Organization. 2019. Available from: https://www.who.int/news-room/factsheets/detail/epilepsy.

4. Myers CT, Mefford HC. Advancing epilepsy genetics in the genomic era. Genome Med. 2015;7:91. Epub 2015/08/26. doi: 10.1186/s13073-015-0214-7. PubMed PMID: 26302787; PMCID: PMC4549122.

5. Rao VR, Lowenstein DH. Epilepsy. Curr Biol. 2015;25(17):R742-6. Epub 2015/09/02. doi: 10.1016/j.cub.2015.07.072. PubMed PMID: 26325130.

6. Forsgren L, Hauser WA, Olafsson E, Sander JW, Sillanpaa M, Tomson T. Mortality of epilepsy in developed countries: a review. Epilepsia. 2005;46 Suppl 11:18-27. Epub 2006/01/06. doi: 10.1111/j.1528-1167.2005.00403.x. PubMed PMID: 16393174.

7. Seidenberg M, Pulsipher DT, Hermann B. Association of epilepsy and comorbid conditions. Future Neurol. 2009;4(5):663-8. Epub 2010/02/18. doi: 10.2217/fnl.09.32. PubMed PMID: 20161538; PMCID: PMC2802344.

8. Patel P, Moshe SL. The evolution of the concepts of seizures and epilepsy: What's in a name? Epilepsia Open. 2020;5(1):22-35. Epub 2020/03/07. doi: 10.1002/epi4.12375. PubMed PMID: 32140641; PMCID: PMC7049807.

9. Wilson JV, Reynolds EH. Texts and documents. Translation and analysis of a cuneiform text forming part of a Babylonian treatise on epilepsy. Med Hist. 1990;34(2):185-98. Epub 1990/04/01. doi: 10.1017/s0025727300050651. PubMed PMID: 2187129; PMCID: PMC1036070.

10. Lennox WG. The genetics of epilepsy. Am J Psychiatry. 1947;103(4):457-62. Epub 1947/01/01. doi: 10.1176/ajp.103.4.457. PubMed PMID: 20287556.

11. Lennox WG. The heredity of epilepsy as told by relatives and twins. J Am Med Assoc. 1951;146(6):529-36. Epub 1951/06/09. doi: 10.1001/jama.1951.03670060005002. PubMed PMID: 14832012.

12. Berkovic SF, Howell RA, Hay DA, Hopper JL. Epilepsies in twins: genetics of the major epilepsy syndromes. Ann Neurol. 1998;43(4):435-45. Epub 1998/04/18. doi: 10.1002/ana.410430405. PubMed PMID: 9546323.

13. Kjeldsen MJ, Corey LA, Christensen K, Friis ML. Epileptic seizures and syndromes in twins: the importance of genetic factors. Epilepsy Res. 2003;55(1-2):137-46. Epub 2003/09/02. doi: 10.1016/s0920-1211(03)00117-7. PubMed PMID: 12948623.

14. Hauser WA, Annegers JF, Kurland LT. Incidence of epilepsy and unprovoked seizures in Rochester, Minnesota: 1935-1984. Epilepsia. 1993;34(3):453-68. Epub 1993/05/01. doi: 10.1111/j.1528-1157.1993.tb02586.x. PubMed PMID: 8504780.

15. Jallon P, Loiseau P, Loiseau J. Newly diagnosed unprovoked epileptic seizures: presentation at diagnosis in CAROLE study. Coordination Active du Reseau Observatoire Longitudinal de l' Epilepsie. Epilepsia. 2001;42(4):464-75. Epub 2001/07/07. doi: 10.1046/j.1528-1157.2001.31400.x. PubMed PMID: 11440341.

16. Steinlein OK, Mulley JC, Propping P, Wallace RH, Phillips HA, Sutherland GR, Scheffer IE, Berkovic SF. A missense mutation in the neuronal nicotinic acetylcholine receptor alpha 4 subunit is associated with autosomal dominant nocturnal frontal lobe epilepsy. Nat Genet. 1995;11(2):201-3. Epub 1995/10/01. doi: 10.1038/ng1095-201. PubMed PMID: 7550350.

17. Vesa J, Hellsten E, Verkruyse LA, Camp LA, Rapola J, Santavuori P, Hofmann SL, Peltonen L. Mutations in the palmitoyl protein thioesterase gene causing infantile neuronal ceroid lipofuscinosis. Nature. 1995;376(6541):584-7. Epub 1995/08/17. doi: 10.1038/376584a0. PubMed PMID: 7637805.

18. Calhoun JD, Carvill GL. Unravelling the genetic architecture of autosomal recessive epilepsy in the genomic era. J Neurogenet. 2018;32(4):295-312. Epub 2018/09/25. doi: 10.1080/01677063.2018.1513509. PubMed PMID: 30247086.

19. Escayg A, Goldin AL. Sodium channel SCN1A and epilepsy: mutations and mechanisms. Epilepsia. 2010;51(9):1650-8. Epub 2010/09/14. doi: 10.1111/j.1528-1167.2010.02640.x. PubMed PMID: 20831750; PMCID: PMC2937162.

20. Rosanoff MJ, Ottman R. Penetrance of LGI1 mutations in autosomal dominant partial epilepsy with auditory features. Neurology. 2008;71(8):567-71. Epub 2008/08/20. doi: 10.1212/01.wnl.0000323926.77565.ee. PubMed PMID: 18711109; PMCID: PMC2652575.

21. Acuna-Hidalgo R, Bo T, Kwint MP, van de Vorst M, Pinelli M, Veltman JA, Hoischen A, Vissers LE, Gilissen C. Post-zygotic Point Mutations Are an Underrecognized Source of De Novo Genomic Variation. Am J Hum Genet. 2015;97(1):67-74. Epub 2015/06/10. doi: 10.1016/j.ajhg.2015.05.008. PubMed PMID: 26054435; PMCID: PMC4571017.

22. Helbig I, Heinzen EL, Mefford HC, Commission IG. Primer Part 1-The building blocks of epilepsy genetics. Epilepsia. 2016;57(6):861-8. Epub 2016/05/27. doi: 10.1111/epi.13381. PubMed PMID: 27226047.

23. Kacser H, Burns JA. The molecular basis of dominance. Genetics. 1981;97(3-4):639-66. Epub 1981/03/01. doi: 10.1093/genetics/97.3-4.639. PubMed PMID: 7297851; PMCID: PMC1214416.

24. Yarham JW, Lamichhane TN, Pyle A, Mattijssen S, Baruffini E, Bruni F, Donnini C, Vassilev A, He L, Blakely EL, Griffin H, Santibanez-Koref M, Bindoff LA, Ferrero I, Chinnery PF, McFarland R, Maraia RJ, Taylor RW. Defective i6A37 modification of mitochondrial and

cytosolic tRNAs results from pathogenic mutations in TRIT1 and its substrate tRNA. PLoS Genet. 2014;10(6):e1004424. Epub 2014/06/06. doi: 10.1371/journal.pgen.1004424. PubMed PMID: 24901367; PMCID: PMC4046958.

25. Chang IJ, He M, Lam CT. Congenital disorders of glycosylation. Ann Transl Med. 2018;6(24):477. Epub 2019/02/12. doi: 10.21037/atm.2018.10.45. PubMed PMID: 30740408; PMCID: PMC6331365.

26. Inoue K, Zhuang L, Ganapathy V. Human Na+ -coupled citrate transporter: primary structure, genomic organization, and transport function. Biochem Biophys Res Commun. 2002;299(3):465-71. Epub 2002/11/26. doi: 10.1016/s0006-291x(02)02669-4. PubMed PMID: 12445824.

27. Thevenon J, Milh M, Feillet F, St-Onge J, Duffourd Y, Juge C, Roubertie A, Heron D, Mignot C, Raffo E, Isidor B, Wahlen S, Sanlaville D, Villeneuve N, Darmency-Stamboul V, Toutain A, Lefebvre M, Chouchane M, Huet F, Lafon A, de Saint Martin A, Lesca G, El Chehadeh S, Thauvin-Robinet C, Masurel-Paulet A, Odent S, Villard L, Philippe C, Faivre L, Riviere JB. Mutations in SLC13A5 cause autosomal-recessive epileptic encephalopathy with seizure onset in the first days of life. Am J Hum Genet. 2014;95(1):113-20. Epub 2014/07/06. doi: 10.1016/j.ajhg.2014.06.006. PubMed PMID: 24995870; PMCID: PMC4085634.

28. Carvill GL, Heavin SB, Yendle SC, McMahon JM, O'Roak BJ, Cook J, Khan A, Dorschner MO, Weaver M, Calvert S, Malone S, Wallace G, Stanley T, Bye AM, Bleasel A, Howell KB, Kivity S, Mackay MT, Rodriguez-Casero V, Webster R, Korczyn A, Afawi Z, Zelnick N, Lerman-Sagie T, Lev D, Moller RS, Gill D, Andrade DM, Freeman JL, Sadleir LG, Shendure J, Berkovic SF, Scheffer IE, Mefford HC. Targeted resequencing in epileptic encephalopathies identifies de novo mutations in CHD2 and SYNGAP1. Nat Genet. 2013;45(7):825-30. Epub 2013/05/28. doi: 10.1038/ng.2646. PubMed PMID: 23708187; PMCID: PMC3704157.

29. Stromme P, Mangelsdorf ME, Shaw MA, Lower KM, Lewis SM, Bruyere H, Lutcherath V, Gedeon AK, Wallace RH, Scheffer IE, Turner G, Partington M, Frints SG, Fryns JP, Sutherland GR, Mulley JC, Gecz J. Mutations in the human ortholog of Aristaless cause X-linked mental retardation and epilepsy. Nat Genet. 2002;30(4):441-5. Epub 2002/03/13. doi: 10.1038/ng862. PubMed PMID: 11889467.

30. Fox JW, Lamperti ED, Eksioglu YZ, Hong SE, Feng Y, Graham DA, Scheffer IE, Dobyns WB, Hirsch BA, Radtke RA, Berkovic SF, Huttenlocher PR, Walsh CA. Mutations in filamin 1 prevent migration of cerebral cortical neurons in human periventricular heterotopia. Neuron. 1998;21(6):1315-25. Epub 1999/01/12. doi: 10.1016/s0896-6273(00)80651-0. PubMed PMID: 9883725.

31. Canali G, Garcia M, Hivert B, Pinatel D, Goullancourt A, Oguievetskaia K, Saint-Martin M, Girault JA, Faivre-Sarrailh C, Goutebroze L. Genetic variants in autism-related CNTNAP2 impair axonal growth of cortical neurons. Hum Mol Genet. 2018;27(11):1941-54. Epub 2018/05/23. doi: 10.1093/hmg/ddy102. PubMed PMID: 29788201.

32. Scott R, Sanchez-Aguilera A, van Elst K, Lim L, Dehorter N, Bae SE, Bartolini G, Peles E, Kas MJH, Bruining H, Marin O. Loss of Cntnap2 Causes Axonal Excitability Deficits, Developmental Delay in Cortical Myelination, and Abnormal Stereotyped Motor Behavior. Cereb Cortex. 2019;29(2):586-97. Epub 2018/01/05. doi: 10.1093/cercor/bhx341. PubMed PMID: 29300891.

33. Dong X, Tan NB, Howell KB, Barresi S, Freeman JL, Vecchio D, Piccione M, Radio FC, Calame D, Zong S, Eggers S, Scheffer IE, Tan TY, Van Bergen NJ, Tartaglia M, Christodoulou J, White SM. Bi-allelic LoF NRROS Variants Impairing Active TGF-beta1 Delivery Cause a Severe Infantile-Onset Neurodegenerative Condition with Intracranial Calcification. Am J Hum Genet. 2020;106(4):559-69. Epub 2020/03/21. doi: 10.1016/j.ajhg.2020.02.014. PubMed PMID: 32197075; PMCID: PMC7118692.

34. Smith C, McColl BW, Patir A, Barrington J, Armishaw J, Clarke A, Eaton J, Hobbs V, Mansour S, Nolan M, Rice GI, Rodero MP, Seabra L, Uggenti C, Livingston JH, Bridges LR, Jeffrey IJM, Crow YJ. Biallelic mutations in NRROS cause an early onset lethal microgliopathy. Acta Neuropathol. 2020;139(5):947-51. Epub 2020/02/27. doi: 10.1007/s00401-020-02137-7. PubMed PMID: 32100099; PMCID: PMC7181551.

35. Steinlein O, Smigrodzki R, Lindstrom J, Anand R, Kohler M, Tocharoentanaphol C, Vogel F. Refinement of the localization of the gene for neuronal nicotinic acetylcholine receptor alpha 4 subunit (CHRNA4) to human chromosome 20q13.2-q13.3. Genomics. 1994;22(2):493-5. Epub 1994/07/15. doi: 10.1006/geno.1994.1420. PubMed PMID: 7806245.

36. Phillips HA, Scheffer IE, Berkovic SF, Hollway GE, Sutherland GR, Mulley JC. Localization of a gene for autosomal dominant nocturnal frontal lobe epilepsy to chromosome 20q 13.2. Nat Genet. 1995;10(1):117-8. Epub 1995/05/01. doi: 10.1038/ng0595-117. PubMed PMID: 7647781.

37. Helbig I. New technologies in molecular genetics: the impact on epilepsy research. Prog Brain Res. 2014;213:253-78. Epub 2014/09/10. doi: 10.1016/B978-0-444-63326-2.00013-2. PubMed PMID: 25194493.

38. Sheidley BR, Malinowski J, Bergner AL, Bier L, Gloss DS, Mu W, Mulhern MM, Partack EJ, Poduri A. Genetic testing for the epilepsies: A systematic review. Epilepsia. 2021. Epub 2021/12/12. doi: 10.1111/epi.17141. PubMed PMID: 34893972.

39. Helbig KL, Farwell Hagman KD, Shinde DN, Mroske C, Powis Z, Li S, Tang S, Helbig I. Diagnostic exome sequencing provides a molecular diagnosis for a significant proportion of patients with epilepsy. Genet Med. 2016;18(9):898-905. Epub 2016/01/23. doi: 10.1038/gim.2015.186. PubMed PMID: 26795593.

40. Retterer K, Juusola J, Cho MT, Vitazka P, Millan F, Gibellini F, Vertino-Bell A, Smaoui N, Neidich J, Monaghan KG, McKnight D, Bai R, Suchy S, Friedman B, Tahiliani J, Pineda-Alvarez D, Richard G, Brandt T, Haverfield E, Chung WK, Bale S. Clinical application of whole-exome sequencing across clinical indications. Genet Med. 2016;18(7):696-704. Epub 2015/12/04. doi: 10.1038/gim.2015.148. PubMed PMID: 26633542.

41. Palmer EE, Schofield D, Shrestha R, Kandula T, Macintosh R, Lawson JA, Andrews I, Sampaio H, Johnson AM, Farrar MA, Cardamone M, Mowat D, Elakis G, Lo W, Zhu Y, Ying K, Morris P, Tao J, Dias KR, Buckley M, Dinger ME, Cowley MJ, Roscioli T, Kirk EP, Bye A, Sachdev RK. Integrating exome sequencing into a diagnostic pathway for epileptic encephalopathy: Evidence of clinical utility and cost effectiveness. Mol Genet Genomic Med. 2018;6(2):186-99. Epub 2018/01/10. doi: 10.1002/mgg3.355. PubMed PMID: 29314763; PMCID: PMC5902395.

42. Stark Z, Schofield D, Alam K, Wilson W, Mupfeki N, Macciocca I, Shrestha R, White SM, Gaff C. Prospective comparison of the cost-effectiveness of clinical whole-exome sequencing with that of usual care overwhelmingly supports early use and reimbursement. Genet Med. 2017;19(8):867-74. Epub 2017/01/27. doi: 10.1038/gim.2016.221. PubMed PMID: 28125081.

43. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL, Committee ALQA. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17(5):405-24. Epub 2015/03/06. doi: 10.1038/gim.2015.30. PubMed PMID: 25741868; PMCID: PMC4544753.

44. Brnich SE, Rivera-Munoz EA, Berg JS. Quantifying the potential of functional evidence to reclassify variants of uncertain significance in the categorical and Bayesian interpretation frameworks. Hum Mutat. 2018;39(11):1531-41. Epub 2018/08/11. doi: 10.1002/humu.23609. PubMed PMID: 30095857; PMCID: PMC6548460.

45. Brnich SE, Abou Tayoun AN, Couch FJ, Cutting GR, Greenblatt MS, Heinen CD, Kanavy DM, Luo X, McNulty SM, Starita LM, Tavtigian SV, Wright MW, Harrison SM, Biesecker LG, Berg JS, Clinical Genome Resource Sequence Variant Interpretation Working G. Recommendations for application of the functional evidence PS3/BS3 criterion using the ACMG/AMP sequence variant interpretation framework. Genome Med. 2019;12(1):3. Epub 2020/01/02. doi: 10.1186/s13073-019-0690-2. PubMed PMID: 31892348; PMCID: PMC6938631.

46. Sobreira N, Schiettecatte F, Valle D, Hamosh A. GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. Hum Mutat. 2015;36(10):928-30. Epub 2015/07/30. doi: 10.1002/humu.22844. PubMed PMID: 26220891; PMCID: PMC4833888.

47. Thomas RH, Berkovic SF. The hidden genetics of epilepsy-a clinically important new paradigm. Nat Rev Neurol. 2014;10(5):283-92. Epub 2014/04/16. doi: 10.1038/nrneurol.2014.62. PubMed PMID: 24733163.

48. Subbarao BS, Silverman A, Eapen BC. Seizure Medications. [Updated 2021 Jul 13]. In: StatPearls [Internet]. Treasure Island (FL); StatPearls Publishing; 2022 Jan -. Available from: https://www.ncbi.nlm.nih.gov/books/NBK482269/.

49. Meng H, Xu HQ, Yu L, Lin GW, He N, Su T, Shi YW, Li B, Wang J, Liu XR, Tang B, Long YS, Yi YH, Liao WP. The SCN1A mutation database: updating information and analysis of the relationships among genotype, functional alteration, and phenotype. Hum Mutat. 2015;36(6):573-80. Epub 2015/03/11. doi: 10.1002/humu.22782. PubMed PMID: 25754450.

50. Rinaldi C, Wood MJA. Antisense oligonucleotides: the next frontier for treatment of neurological disorders. Nat Rev Neurol. 2018;14(1):9-21. Epub 2017/12/02. doi: 10.1038/nrneurol.2017.148. PubMed PMID: 29192260.

51. Wengert ER, Wagley PK, Strohm SM, Reza N, Wenker IC, Gaykema RP, Christiansen A, Liau G, Patel MK. Targeted Augmentation of Nuclear Gene Output (TANGO) of Scn1a rescues parvalbumin interneuron excitability and reduces seizures in a mouse model of Dravet Syndrome. Brain Res. 2022;1775:147743. Epub 2021/11/30. doi: 10.1016/j.brainres.2021.147743. PubMed PMID: 34843701.

52. Meisler MH. SCN8A encephalopathy: Mechanisms and models. Epilepsia. 2019;60 Suppl 3:S86-S91. Epub 2020/01/07. doi: 10.1111/epi.14703. PubMed PMID: 31904118; PMCID: PMC6953611.

53. Lenk GM, Jafar-Nejad P, Hill SF, Huffman LD, Smolen CE, Wagnon JL, Petit H, Yu W, Ziobro J, Bhatia K, Parent J, Giger RJ, Rigo F, Meisler MH. Scn8a Antisense Oligonucleotide Is Protective in Mouse Models of SCN8A Encephalopathy and Dravet Syndrome. Ann Neurol. 2020;87(3):339-46. Epub 2020/01/17. doi: 10.1002/ana.25676. PubMed PMID: 31943325; PMCID: PMC7064908.

54. Butler KM, da Silva C, Alexander JJ, Hegde M, Escayg A. Diagnostic Yield From 339 Epilepsy Patients Screened on a Clinical Gene Panel. Pediatr Neurol. 2017;77:61-6. Epub 2017/10/24. doi: 10.1016/j.pediatrneurol.2017.09.003. PubMed PMID: 29056246.

55. Trump N, McTague A, Brittain H, Papandreou A, Meyer E, Ngoh A, Palmer R, Morrogh D, Boustred C, Hurst JA, Jenkins L, Kurian MA, Scott RH. Improving diagnosis and broadening the phenotypes in early-onset seizure and severe developmental delay disorders through gene panel analysis. J Med Genet. 2016;53(5):310-7. Epub 2016/03/20. doi: 10.1136/jmedgenet-2015-103263. PubMed PMID: 26993267; PMCID: PMC4862068.

56. Parrini E, Marini C, Mei D, Galuppi A, Cellini E, Pucatti D, Chiti L, Rutigliano D, Bianchini C, Virdo S, De Vita D, Bigoni S, Barba C, Mari F, Montomoli M, Pisano T, Rosati A, Clinical Study G, Guerrini R. Diagnostic Targeted Resequencing in 349 Patients with Drug-Resistant Pediatric Epilepsies Identifies Causative Mutations in 30 Different Genes. Hum Mutat. 2017;38(2):216-25. Epub 2016/11/20. doi: 10.1002/humu.23149. PubMed PMID: 27864847.

57. Durkin MM, Smith KE, Borden LA, Weinshank RL, Branchek TA, Gustafson EL. Localization of messenger RNAs encoding three GABA transporters in rat brain: an in situ hybridization study. Brain Res Mol Brain Res. 1995;33(1):7-21. Epub 1995/10/01. PubMed PMID: 8774941.

58. Carvill GL, McMahon JM, Schneider A, Zemel M, Myers CT, Saykally J, Nguyen J, Robbiano A, Zara F, Specchio N, Mecarelli O, Smith RL, Leventer RJ, Moller RS, Nikanorova M, Dimova P, Jordanova A, Petrou S, Euro ERESM-AE, Dravet working g, Helbig I, Striano P, Weckhuysen S, Berkovic SF, Scheffer IE, Mefford HC. Mutations in the GABA Transporter SLC6A1 Cause Epilepsy with Myoclonic-Atonic Seizures. Am J Hum Genet. 2015;96(5):808-15. Epub 2015/04/14. doi: 10.1016/j.ajhg.2015.02.016. PubMed PMID: 25865495; PMCID: PMC4570550.

59. Halvorsen M, Petrovski S, Shellhaas R, Tang Y, Crandall L, Goldstein D, Devinsky O. Mosaic mutations in early-onset genetic diseases. Genet Med. 2016;18(7):746-9. Epub 2015/12/31. doi: 10.1038/gim.2015.155. PubMed PMID: 26716362; PMCID: PMC4929028.

60. Johannesen KM, Gardella E, Linnankivi T, Courage C, de Saint Martin A, Lehesjoki AE, Mignot C, Afenjar A, Lesca G, Abi-Warde MT, Chelly J, Piton A, Merritt JL, 2nd, Rodan LH, Tan WH, Bird LM, Nespeca M, Gleeson JG, Yoo Y, Choi M, Chae JH, Czapansky-Beilman D, Reichert SC, Pendziwiat M, Verhoeven JS, Schelhaas HJ, Devinsky O, Christensen J, Specchio N, Trivisano M, Weber YG, Nava C, Keren B, Doummar D, Schaefer E, Hopkins S, Dubbs H, Shaw JE, Pisani L, Myers CT, Tang S, Tang S, Pal DK, Millichap JJ, Carvill GL, Helbig KL, Mecarelli O, Striano P, Helbig I, Rubboli G, Mefford HC, Moller RS. Defining the phenotypic

spectrum of SLC6A1 mutations. Epilepsia. 2018;59(2):389-402. Epub 2018/01/10. doi: 10.1111/epi.13986. PubMed PMID: 29315614.

61. Palmer S, Towne MC, Pearl PL, Pelletier RC, Genetti CA, Shi J, Beggs AH, Agrawal PB, Brownstein CA. SLC6A1 Mutation and Ketogenic Diet in Epilepsy With Myoclonic-Atonic Seizures. Pediatr Neurol. 2016;64:77-9. Epub 2016/09/08. doi: 10.1016/j.pediatrneurol.2016.07.012. PubMed PMID: 27600546; PMCID: PMC5223550.

62. Zech M, Jech R, Wagner M, Mantel T, Boesch S, Nocker M, Jochim A, Berutti R, Havrankova P, Fecikova A, Kemlink D, Roth J, Strom TM, Poewe W, Ruzicka E, Haslinger B, Winkelmann J. Molecular diversity of combined and complex dystonia: insights from diagnostic exome sequencing. Neurogenetics. 2017. Epub 2017/08/30. doi: 10.1007/s10048-017-0521-9. PubMed PMID: 28849312.

63. Kunkel TA, Roberts JD, Zakour RA. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. 1987;154:367-82. PubMed PMID: 3323813.

64. Kleinberger-Doron N, Kanner BI. Identification of tryptophan residues critical for the function and targeting of the gamma-aminobutyric acid transporter (subtype A). J Biol Chem. 1994;269(4):3063-7. PubMed PMID: 8300640.

65. Fuerst TR, Niles EG, Studier FW, Moss B. Eukaryotic transient-expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase. Proc Natl Acad Sci U S A. 1986;83(21):8122-6. PubMed PMID: 3095828.

66. Keynan S, Suh YJ, Kanner BI, Rudnick G. Expression of a cloned gamma-aminobutyric acid transporter in mammalian cells. Biochemistry. 1992;31(7):1974-9. PubMed PMID: 1536839.

67. Kishore S, Khanna A, Stamm S. Rapid generation of splicing reporters with pSpliceExpress. Gene. 2008;427(1-2):104-10. Epub 2008/10/22. doi: 10.1016/j.gene.2008.09.021. PubMed PMID: 18930792; PMCID: PMC2821805.

68. Ko A, Youn SE, Kim SH, Lee JS, Kim S, Choi JR, Kim HD, Lee ST, Kang HC. Targeted gene panel and genotype-phenotype correlation in children with developmental and epileptic encephalopathy. Epilepsy Res. 2018;141:48-55. Epub 2018/02/20. doi: 10.1016/j.eplepsyres.2018.02.003. PubMed PMID: 29455050.

69. Helseth A, Mohamad M. New Clinical Phenotype Caused by a Novel SLC6A1 Mutation with Marked Responsiveness to Zonisamide (P4. 179). Neurology. 2017;88(16 Supplement):P4. 179.

70. Lucariello M, Vidal E, Vidal S, Saez M, Roa L, Huertas D, Pineda M, Dalfo E, Dopazo J, Jurado P, Armstrong J, Esteller M. Whole exome sequencing of Rett syndrome-like patients reveals the mutational diversity of the clinical phenotype. Hum Genet. 2016;135(12):1343-54. Epub 2016/08/20. doi: 10.1007/s00439-016-1721-3. PubMed PMID: 27541642; PMCID: PMC5065581.

71. Posey JE, Harel T, Liu P, Rosenfeld JA, James RA, Coban Akdemir ZH, Walkiewicz M, Bi W, Xiao R, Ding Y, Xia F, Beaudet AL, Muzny DM, Gibbs RA, Boerwinkle E, Eng CM, Sutton VR, Shaw CA, Plon SE, Yang Y, Lupski JR. Resolution of Disease Phenotypes Resulting from

Multilocus Genomic Variation. N Engl J Med. 2017;376(1):21-31. Epub 2016/12/14. doi: 10.1056/NEJMoa1516767. PubMed PMID: 27959697; PMCID: PMC5335876.

72. Rauch A, Wieczorek D, Graf E, Wieland T, Endele S, Schwarzmayr T, Albrecht B, Bartholdi D, Beygo J, Di Donato N, Dufke A, Cremer K, Hempel M, Horn D, Hoyer J, Joset P, Ropke A, Moog U, Riess A, Thiel CT, Tzschach A, Wiesener A, Wohlleber E, Zweier C, Ekici AB, Zink AM, Rump A, Meisinger C, Grallert H, Sticht H, Schenck A, Engels H, Rappold G, Schrock E, Wieacker P, Riess O, Meitinger T, Reis A, Strom TM. Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. Lancet. 2012;380(9854):1674-82. Epub 2012/10/02. doi: 10.1016/S0140-6736(12)61480-9. PubMed PMID: 23020937.

73. Sanders SJ, Murtha MT, Gupta AR, Murdoch JD, Raubeson MJ, Willsey AJ, Ercan-Sencicek AG, DiLullo NM, Parikshak NN, Stein JL, Walker MF, Ober GT, Teran NA, Song Y, El-Fishawy P, Murtha RC, Choi M, Overton JD, Bjornson RD, Carriero NJ, Meyer KA, Bilguvar K, Mane SM, Sestan N, Lifton RP, Gunel M, Roeder K, Geschwind DH, Devlin B, State MW. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. Nature. 2012;485(7397):237-41. Epub 2012/04/13. doi: 10.1038/nature10945. PubMed PMID: 22495306; PMCID: PMC3667984.

74. Wang TY, Guo H, Xiong B, Stessman HAF, Wu HD, Coe BP, Turner TN, Liu YL, Zhao WJ, Hoekzema K, Vives L, Xia L, Tang MN, Ou JJ, Chen BY, Shen YD, Xun GL, Long M, Lin J, Kronenberg ZN, Peng Y, Bai T, Li HH, Ke XY, Hu ZM, Zhao JP, Zou XB, Xia K, Eichler EE. De novo genic mutations among a Chinese autism spectrum disorder cohort. Nature Communications. 2016;7. doi:10.1038/ncomms13316. PubMed PMID: 27824329; PMCID: PMC5105161.

75. Yoo Y, Jung J, Lee YN, Lee Y, Cho H, Na E, Hong J, Kim E, Lee JS, Lee JS, Hong CS, Park SY, Wie JH, Miller K, Shur N, Clow C, Ebel RS, DeBrosse SD, Henderson LB, Willaert R, Castaldi C, Tikhonova I, Bilguvar K, Mane S, Kim KJ, Hwang YS, Lee SG, So I, Lim BC, Choi HJ, Seong JY, Shin YB, Jung H, Chae JH, Choi M. GABBR2 Mutations Determine Phenotype in Rett Syndrome and Epileptic Encephalopathy. Annals of Neurology. 2017;82(3):466-78. doi: 10.1002/ana.25032. PubMed PMID: 28856709.

76. Yamashita A, Singh SK, Kawate T, Jin Y, Gouaux E. Crystal structure of a bacterial homologue of Na+/Cl--dependent neurotransmitter transporters. Nature. 2005;437(7056):215-23. Epub 2005/07/26. doi: 10.1038/nature03978. PubMed PMID: 16041361.

77. Krishnamurthy H, Gouaux E. X-ray structures of LeuT in substrate-free outward-open and apo inward-open states. Nature. 2012;481(7382):469-74. Epub 2012/01/11. doi: 10.1038/nature10737. PubMed PMID: 22230955; PMCID: PMC3306218.

78. Ulmschneider MB, Sansom MS. Amino acid distributions in integral membrane protein structures. Biochim Biophys Acta. 2001;1512(1):1-14. Epub 2001/05/04. PubMed PMID: 11334619.

79. Jensen K, Chiu CS, Sokolova I, Lester HA, Mody I. GABA transporter-1 (GAT1)-deficient mice: differential tonic activation of GABAA versus GABAB receptors in the hippocampus. J Neurophysiol. 2003;90(4):2690-701. Epub 2003/06/20. doi: 10.1152/jn.00240.2003. PubMed PMID: 12815026.

80. Han HA, Cortez MA, Snead OC, III. GABAB Receptor and Absence Epilepsy. In: th, Noebels JL, Avoli M, Rogawski MA, Olsen RW, Delgado-Escueta AV, editors. Jasper's Basic Mechanisms of the Epilepsies. Bethesda (MD)2012.

81. Cope DW, Di Giovanni G, Fyson SJ, Orban G, Errington AC, Lorincz ML, Gould TM, Carter DA, Crunelli V. Enhanced tonic GABAA inhibition in typical absence epilepsy. Nat Med. 2009;15(12):1392-8. Epub 2009/12/08. doi: 10.1038/nm.2058. PubMed PMID: 19966779; PMCID: PMC2824149.

82. Walker MC, Kullmann DM. Tonic GABAA Receptor-Mediated Signaling in Epilepsy. In: th, Noebels JL, Avoli M, Rogawski MA, Olsen RW, Delgado-Escueta AV, editors. Jasper's Basic Mechanisms of the Epilepsies. Bethesda (MD)2012.

83. Khosravani H, Zamponi GW. Voltage-gated calcium channels and idiopathic generalized epilepsies. Physiol Rev. 2006;86(3):941-66. Epub 2006/07/04. doi: 10.1152/physrev.00002.2006. PubMed PMID: 16816142.

84. Snead OC, 3rd. Evidence for GABAB-mediated mechanisms in experimental generalized absence seizures. Eur J Pharmacol. 1992;213(3):343-9. Epub 1992/03/31. PubMed PMID: 1319918.

85. Zhang L, Wither RG, Lang M, Wu C, Sidorova-Darmos E, Netchev H, Matolcsy CB, Snead OC, Eubanks JH. A Role for Diminished GABA Transporter Activity in the Cortical Discharge Phenotype of MeCP2-Deficient Mice. Neuropsychopharmacology. 2016;41(6):1467-76. Epub 2015/10/27. doi: 10.1038/npp.2015.323. PubMed PMID: 26499511; PMCID: PMC4832024.

86. Macdonald RL, Kang JQ, Gallagher MJ. Mutations in GABAA receptor subunits associated with genetic epilepsies. J Physiol. 2010;588(Pt 11):1861-9. Epub 2010/03/24. doi: 10.1113/jphysiol.2010.186999. PubMed PMID: 20308251; PMCID: PMC2901974.

87. Forgac M. Vacuolar ATPases: rotary proton pumps in physiology and pathophysiology. Nat Rev Mol Cell Biol. 2007;8(11):917-29. Epub 2007/10/04. doi: 10.1038/nrm2272. PubMed PMID: 17912264.

88. Maxson ME, Grinstein S. The vacuolar-type H(+)-ATPase at a glance - more than a proton pump. J Cell Sci. 2014;127(Pt 23):4987-93. Epub 2014/12/03. doi: 10.1242/jcs.158550. PubMed PMID: 25453113.

89. Van Damme T, Gardeitchik T, Mohamed M, Guerrero-Castillo S, Freisinger P, Guillemyn B, Kariminejad A, Dalloyaux D, van Kraaij S, Lefeber DJ, Syx D, Steyaert W, De Rycke R, Hoischen A, Kamsteeg EJ, Wong SY, van Scherpenzeel M, Jamali P, Brandt U, Nijtmans L, Korenke GC, Chung BHY, Mak CCY, Hausser I, Kornak U, Fischer-Zirnsak B, Strom TM, Meitinger T, Alanay Y, Utine GE, Leung PKC, Ghaderi-Sohi S, Coucke P, Symoens S, De Paepe A, Thiel C, Haack TB, Malfait F, Morava E, Callewaert B, Wevers RA. Mutations in ATP6V1E1 or ATP6V1A Cause Autosomal-Recessive Cutis Laxa. Am J Hum Genet. 2017;100(2):216-27. Epub 2017/01/10. doi: 10.1016/j.ajhg.2016.12.010. PubMed PMID: 28065471; PMCID: PMC5294668.

90. Alazami AM, Al-Qattan SM, Faqeih E, Alhashem A, Alshammari M, Alzahrani F, Al-Dosari MS, Patel N, Alsagheir A, Binabbas B, Alzaidan H, Alsiddiky A, Alharbi N, Alfadhel M,

Kentab A, Daza RM, Kircher M, Shendure J, Hashem M, Alshahrani S, Rahbeeni Z, Khalifa O, Shaheen R, Alkuraya FS. Expanding the clinical and genetic heterogeneity of hereditary disorders of connective tissue. Hum Genet. 2016;135(5):525-40. Epub 2016/03/30. doi: 10.1007/s00439-016-1660-z. PubMed PMID: 27023906.

91. Kornak U, Reynders E, Dimopoulou A, van Reeuwijk J, Fischer B, Rajab A, Budde B, Nurnberg P, Foulquier F, Group AD-tS, Lefeber D, Urban Z, Gruenewald S, Annaert W, Brunner HG, van Bokhoven H, Wevers R, Morava E, Matthijs G, Van Maldergem L, Mundlos S. Impaired glycosylation and cutis laxa caused by mutations in the vesicular H+-ATPase subunit ATP6V0A2. Nat Genet. 2008;40(1):32-4. Epub 2007/12/25. doi: 10.1038/ng.2007.45. PubMed PMID: 18157129.

92. Smith AN, Skaug J, Choate KA, Nayir A, Bakkaloglu A, Ozen S, Hulton SA, Sanjad SA, Al-Sabban EA, Lifton RP, Scherer SW, Karet FE. Mutations in ATP6N1B, encoding a new kidney vacuolar proton pump 116-kD subunit, cause recessive distal renal tubular acidosis with preserved hearing. Nat Genet. 2000;26(1):71-5. Epub 2000/09/06. doi: 10.1038/79208. PubMed PMID: 10973252.

93. Karet FE, Finberg KE, Nelson RD, Nayir A, Mocan H, Sanjad SA, Rodriguez-Soriano J, Santos F, Cremers CW, Di Pietro A, Hoffbrand BI, Winiarski J, Bakkaloglu A, Ozen S, Dusunsel R, Goodyer P, Hulton SA, Wu DK, Skvorak AB, Morton CC, Cunningham MJ, Jha V, Lifton RP. Mutations in the gene encoding B1 subunit of H+-ATPase cause renal tubular acidosis with sensorineural deafness. Nat Genet. 1999;21(1):84-90. Epub 1999/01/23. doi: 10.1038/5022. PubMed PMID: 9916796.

94. Jobst-Schwan T, Klambt V, Tarsio M, Heneghan JF, Majmundar AJ, Shril S, Buerger F, Ottlewski I, Shmukler BE, Topaloglu R, Hashmi S, Hafeez F, Emma F, Greco M, Laube GF, Fathy HM, Pohl M, Gellermann J, Milosevic D, Baum MA, Mane S, Lifton RP, Kane PM, Alper SL, Hildebrandt F. Whole exome sequencing identified ATP6V1C2 as a novel candidate gene for recessive distal renal tubular acidosis. Kidney Int. 2020;97(3):567-79. Epub 2020/01/22. doi: 10.1016/j.kint.2019.09.026. PubMed PMID: 31959358; PMCID: PMC7039771.

95. Steward CG. Neurological aspects of osteopetrosis. Neuropathol Appl Neurobiol. 2003;29(2):87-97. Epub 2003/03/29. PubMed PMID: 12662317.

96. Frattini A, Orchard PJ, Sobacchi C, Giliani S, Abinun M, Mattsson JP, Keeling DJ, Andersson AK, Wallbrandt P, Zecca L, Notarangelo LD, Vezzoni P, Villa A. Defects in TCIRG1 subunit of the vacuolar proton pump are responsible for a subset of human autosomal recessive osteopetrosis. Nat Genet. 2000;25(3):343-6. Epub 2000/07/11. doi: 10.1038/77131. PubMed PMID: 10888887.

97. Yuan Y, Zhang J, Chang Q, Zeng J, Xin F, Wang J, Zhu Q, Wu J, Lu J, Guo W, Yan X, Jiang H, Zhou B, Li Q, Gao X, Yuan H, Yang S, Han D, Mao Z, Chen P, Lin X, Dai P. De novo mutation in ATP6V1B2 impairs lysosome acidification and causes dominant deafness-onychodystrophy syndrome. Cell Res. 2014;24(11):1370-3. Epub 2014/06/11. doi: 10.1038/cr.2014.77. PubMed PMID: 24913193; PMCID: PMC4220163.

98. Kortum F, Caputo V, Bauer CK, Stella L, Ciolfi A, Alawi M, Bocchinfuso G, Flex E, Paolacci S, Dentici ML, Grammatico P, Korenke GC, Leuzzi V, Mowat D, Nair LD, Nguyen TT, Thierry P, White SM, Dallapiccola B, Pizzuti A, Campeau PM, Tartaglia M, Kutsche K. Mutations

in KCNH1 and ATP6V1B2 cause Zimmermann-Laband syndrome. Nat Genet. 2015;47(6):661-7. Epub 2015/04/29. doi: 10.1038/ng.3282. PubMed PMID: 25915598.

99. Zhang Y, Huang H, Zhao G, Yokoyama T, Vega H, Huang Y, Sood R, Bishop K, Maduro V, Accardi J, Toro C, Boerkoel CF, Lyons K, Gahl WA, Duan X, Malicdan MC, Lin S. ATP6V1H Deficiency Impairs Bone Development through Activation of MMP9 and MMP13. PLoS Genet. 2017;13(2):e1006481. Epub 2017/02/06. doi: 10.1371/journal.pgen.1006481. PubMed PMID: 28158191; PMCID: PMC5291374.

100. Fassio A, Esposito A, Kato M, Saitsu H, Mei D, Marini C, Conti V, Nakashima M, Okamoto N, Olmez Turker A, Albuz B, Semerci Gunduz CN, Yanagihara K, Belmonte E, Maragliano L, Ramsey K, Balak C, Siniard A, Narayanan V, Group CRR, Ohba C, Shiina M, Ogata K, Matsumoto N, Benfenati F, Guerrini R. De novo mutations of the ATP6V1A gene cause developmental encephalopathy with epilepsy. Brain. 2018;141(6):1703-18. Epub 2018/04/19. doi: 10.1093/brain/awy092. PubMed PMID: 29668857; PMCID: PMC5972584.

101. Aoto K, Kato M, Akita T, Nakashima M, Mutoh H, Akasaka N, Tohyama J, Nomura Y, Hoshino K, Ago Y, Tanaka R, Epstein O, Ben-Haim R, Heyman E, Miyazaki T, Belal H, Takabayashi S, Ohba C, Takata A, Mizuguchi T, Miyatake S, Miyake N, Fukuda A, Matsumoto N, Saitsu H. ATP6V0A1 encoding the a1-subunit of the V0 domain of vacuolar H(+)-ATPases is essential for brain development in humans and mice. Nat Commun. 2021;12(1):2107. Epub 2021/04/10. doi: 10.1038/s41467-021-22389-5. PubMed PMID: 33833240; PMCID: PMC8032687.

102. Bott LC, Forouhan M, Lieto M, Sala AJ, Ellerington R, Johnson JO, Speciale AA, Criscuolo C, Filla A, Chitayat D, Nemeth AH, Network IUD, Angelucci F, Lim WF, Striano P, Zara F, Helbig I, Muona M, Courage C, Lehesjoki A-E, Berkovic SF, Consortium V-A, Fischbeck KH, Brancati F, Morimoto RI, Wood MJA, Rinaldi C. Variants in ATP6V0A1 cause progressive myoclonus epilepsy and developmental and epileptic encephalopathy. Brain Communications. 2021. doi: 10.1093/braincomms/fcab245.

103. Shaw M, Winczewska-Wiktor A, Badura-Stronka M, Koirala S, Gardner A, Kuszel L, Kowal P, Steinborn B, Starczewska M, Garry S, Scheffer IE, Berkovic SF, Gecz J. EXOME REPORT: Novel mutation in ATP6V1B2 segregating with autosomal dominant epilepsy, intellectual disability and mild gingival and nail abnormalities. Eur J Med Genet. 2019:103799. Epub 2019/10/28. doi: 10.1016/j.ejmg.2019.103799. PubMed PMID: 31655144.

104. Beauregard-Lacroix E, Pacheco-Cuellar G, Ajeawung NF, Tardif J, Dieterich K, Dabir T, Vind-Kezunovic D, White SM, Zadori D, Castiglioni C, Tranebjaerg L, Torring PM, Blair E, Wisniewska M, Camurri MV, van Bever Y, Molidperee S, Taylor J, Dionne-Laporte A, Sisodiya SM, Hennekam RCM, Campeau PM. DOORS syndrome and a recurrent truncating ATP6V1B2 variant. Genet Med. 2021;23(1):149-54. Epub 2020/09/03. doi: 10.1038/s41436-020-00950-9. PubMed PMID: 32873933.

105. Ramser J, Abidi FE, Burckle CA, Lenski C, Toriello H, Wen G, Lubs HA, Engert S, Stevenson RE, Meindl A, Schwartz CE, Nguyen G. A unique exonic splice enhancer mutation in a family with X-linked mental retardation and epilepsy points to a novel role of the renin receptor. Hum Mol Genet. 2005;14(8):1019-27. Epub 2005/03/05. doi: 10.1093/hmg/ddi094. PubMed PMID: 15746149.

106. Gupta HV, Vengoechea J, Sahaya K, Virmani T. A splice site mutation in ATP6AP2 causes X-linked intellectual disability, epilepsy, and parkinsonism. Parkinsonism Relat Disord. 2015;21(12):1473-5. Epub 2015/10/16. doi: 10.1016/j.parkreldis.2015.10.001. PubMed PMID: 26467484.

107. Wang L, Wu D, Robinson CV, Wu H, Fu TM. Structures of a Complete Human V-ATPase Reveal Mechanisms of Its Assembly. Mol Cell. 2020;80(3):501-11 e3. Epub 2020/10/17. doi: 10.1016/j.molcel.2020.09.029. PubMed PMID: 33065002; PMCID: PMC7655608.

108. Mucha BE, Banka S, Ajeawung NF, Molidperee S, Chen GG, Koenig MK, Adejumo RB, Till M, Harbord M, Perrier R, Lemyre E, Boucher RM, Skotko BG, Waxler JL, Thomas MA, Hodge JC, Gecz J, Nicholl J, McGregor L, Linden T, Sisodiya SM, Sanlaville D, Cheung SW, Ernst C, Campeau PM. A new microdeletion syndrome involving TBC1D24, ATP6V0C, and PDPK1 causes epilepsy, microcephaly, and developmental delay. Genet Med. 2019;21(5):1058-64. Epub 2018/09/25. doi: 10.1038/s41436-018-0290-3. PubMed PMID: 30245510.

109. Tinker RJ, Burghel GJ, Garg S, Steggall M, Cuvertino S, Banka S. Haploinsufficiency of ATP6V0C possibly underlies 16p13.3 deletions that cause microcephaly, seizures, and neurodevelopmental disorder. Am J Med Genet A. 2020. Epub 2020/10/23. doi: 10.1002/ajmg.a.61905. PubMed PMID: 33090716.

110. Ittiwut C, Poonmaksatit S, Boonsimma P, Desudchit T, Suphapeetiporn K, Ittiwut R, Shotelersuk V. Novel de novo mutation substantiates ATP6V0C as a gene causing epilepsy with intellectual disability. Brain Dev. 2021;43(3):490-4. Epub 2020/11/17. doi: 10.1016/j.braindev.2020.10.016. PubMed PMID: 33190975.

111. Deciphering Developmental Disorders S. Prevalence and architecture of de novo mutations in developmental disorders. Nature. 2017;542(7642):433-8. Epub 2017/01/31. doi: 10.1038/nature21062. PubMed PMID: 28135719.

112. Investigators GPP, Smedley D, Smith KR, Martin A, Thomas EA, McDonagh EM, Cipriani V, Ellingford JM, Arno G, Tucci A, Vandrovcova J, Chan G, Williams HJ, Ratnaike T, Wei W, Stirrups K, Ibanez K, Moutsianas L, Wielscher M, Need A, Barnes MR, Vestito L, Buchanan J, Wordsworth S, Ashford S, Rehmstrom K, Li E, Fuller G, Twiss P, Spasic-Boskovic O, Halsall S, Floto RA, Poole K, Wagner A, Mehta SG, Gurnell M, Burrows N, James R, Penkett C, Dewhurst E, Graf S, Mapeta R, Kasanicki M, Haworth A, Savage H, Babcock M, Reese MG, Bale M, Baple E, Boustred C, Brittain H, de Burca A, Bleda M, Devereau A, Halai D, Haraldsdottir E, Hyder Z, Kasperaviciute D, Patch C, Polychronopoulos D, Matchan A, Sultana R, Ryten M, Tavares ALT, Tregidgo C, Turnbull C, Welland M, Wood S, Snow C, Williams E, Leigh S, Foulger RE, Daugherty LC, Niblock O, Leong IUS, Wright CF, Davies J, Crichton C, Welch J, Woods K, Abulhoul L, Aurora P, Bockenhauer D, Broomfield A, Cleary MA, Lam T, Dattani M, Footitt E, Ganesan V, Grunewald S, Compeyrot-Lacassagne S, Muntoni F, Pilkington C, Quinlivan R, Thapar N, Wallis C, Wedderburn LR, Worth A, Bueser T, Compton C, Deshpande C, Fassihi H, Haque E, Izatt L, Josifova D, Mohammed S, Robert L, Rose S, Ruddy D, Sarkany R, Say G, Shaw AC, Wolejko A, Habib B, Burns G, Hunter S, Grocock RJ, Humphray SJ, Robinson PN, Haendel M, Simpson MA, Banka S, Clayton-Smith J, Douzgou S, Hall G, Thomas HB, O'Keefe RT, Michaelides M, Moore AT, Malka S, Pontikos N, Browning AC, Straub V, Gorman GS, Horvath R, Quinton R, Schaefer AM, Yu-Wai-Man P, Turnbull DM, McFarland R, Taylor RW, O'Connor E, Yip J, Newland K, Morris HR, Polke J, Wood NW, Campbell C, Camps C, Gibson K, Koelling N, Lester T, Nemeth AH, Palles C, Patel S, Roy

NBA, Sen A, Taylor J, Cacheiro P, Jacobsen JO, Seaby EG, Davison V, Chitty L, Douglas A, Naresh K, McMullan D, Ellard S, Temple IK, Mumford AD, Wilson G, Beales P, Bitner-Glindzicz M, Black G, Bradley JR, Brennan P, Burn J, Chinnery PF, Elliott P, Flinter F, Houlden H, Irving M, Newman W, Rahman S, Sayer JA, Taylor JC, Webster AR, Wilkie AOM, Ouwehand WH, Raymond FL, Chisholm J, Hill S, Bentley D, Scott RH, Fowler T, Rendon A, Caulfield M. 100,000 Genomes Pilot on Rare-Disease Diagnosis in Health Care - Preliminary Report. N Engl J Med. 2021;385(20):1868-80. Epub 2021/11/11. doi: 10.1056/NEJMoa2035790. PubMed PMID: 34758253.

113. Carvill GL, Weckhuysen S, McMahon JM, Hartmann C, Moller RS, Hjalgrim H, Cook J, Geraghty E, O'Roak BJ, Petrou S, Clarke A, Gill D, Sadleir LG, Muhle H, von Spiczak S, Nikanorova M, Hodgson BL, Gazina EV, Suls A, Shendure J, Dibbens LM, De Jonghe P, Helbig I, Berkovic SF, Scheffer IE, Mefford HC. GABRA1 and STXBP1: novel genetic causes of Dravet syndrome. Neurology. 2014;82(14):1245-53. Epub 2014/03/14. doi: 10.1212/WNL.0000000000000291. PubMed PMID: 24623842; PMCID: PMC4001207.

114. Jay JJ, Brouwer C. Lollipops in the Clinic: Information Dense Mutation Plots for Precision Medicine. PLoS One. 2016;11(8):e0160519. Epub 2016/08/05. doi: 10.1371/journal.pone.0160519. PubMed PMID: 27490490; PMCID: PMC4973895.

115. Butler KM, Moody OA, Schuler E, Coryell J, Alexander JJ, Jenkins A, Escayg A. De novo variants in GABRA2 and GABRA5 alter receptor function and contribute to early-onset epilepsy. Brain. 2018;141(8):2392-405. Epub 2018/07/03. doi: 10.1093/brain/awy171. PubMed PMID: 29961870; PMCID: PMC6061692.

116. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alfoldi J, Wang Q, Collins RL, Laricchia KM, Ganna A, Birnbaum DP, Gauthier LD, Brand H, Solomonson M, Watts NA, Rhodes D, Singer-Berk M, England EM, Seaby EG, Kosmicki JA, Walters RK, Tashman K, Farjoun Y, Banks E, Poterba T, Wang A, Seed C, Whiffin N, Chong JX, Samocha KE, Pierce-Hoffman E, Zappala Z, O'Donnell-Luria AH, Minikel EV, Weisburd B, Lek M, Ware JS, Vittal C, Armean IM, Bergelson L, Cibulskis K, Connolly KM, Covarrubias M, Donnelly S, Ferriera S, Gabriel S, Gentry J, Gupta N, Jeandet T, Kaplan D, Llanwarne C, Munshi R, Novod S, Petrillo N, Roazen D, Ruano-Rubio V, Saltzman A, Schleicher M, Soto J, Tibbetts K, Tolonen C, Wade G, Talkowski ME, Genome Aggregation Database C, Neale BM, Daly MJ, MacArthur DG. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature. 2020;581(7809):434-43. Epub 2020/05/29. doi: 10.1038/s41586-020-2308-7. PubMed PMID: 32461654; PMCID: PMC7334197.

117. Traynelis J, Silk M, Wang Q, Berkovic SF, Liu L, Ascher DB, Balding DJ, Petrovski S. Optimizing genomic medicine in epilepsy through a gene-customized approach to missense variant interpretation. Genome Res. 2017;27(10):1715-29. Epub 2017/09/03. doi: 10.1101/gr.226589.117. PubMed PMID: 28864458; PMCID: PMC5630035.

118. Marley R, Baines RA. Increased persistent Na+ current contributes to seizure in the slamdance bang-sensitive Drosophila mutant. J Neurophysiol. 2011;106(1):18-29. Epub 2011/04/01. doi: 10.1152/jn.00808.2010. PubMed PMID: 21451059; PMCID: PMC3129721.

119. Berman H, Henrick K, Nakamura H, Markley JL. The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data. Nucleic Acids Res. 2007;35(Database issue):D301-3. Epub 2006/12/05. doi: 10.1093/nar/gkl971. PubMed PMID: 17142228; PMCID: PMC1669775.

120. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis. 1997;18(15):2714-23. Epub 1998/03/21. doi: 10.1002/elps.1150181505. PubMed PMID: 9504803.

121. Green MR, Sambrook J, Sambrook J. Molecular cloning : a laboratory manual. 4th ed. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press; 2012.

122. Sikorski RS, Hieter P. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. Genetics. 1989;122(1):19-27. Epub 1989/05/01. PubMed PMID: 2659436; PMCID: PMC1203683.

123. Perzov N, Padler-Karavani V, Nelson H, Nelson N. Characterization of yeast V-ATPase mutants lacking Vph1p or Stv1p and the effect on endocytosis. J Exp Biol. 2002;205(Pt 9):1209-19. Epub 2002/04/12. PubMed PMID: 11948198.

124. Sprouffske K, Wagner A. Growthcurver: an R package for obtaining interpretable metrics from microbial growth curves. BMC Bioinformatics. 2016;17:172. Epub 2016/04/21. doi: 10.1186/s12859-016-1016-7. PubMed PMID: 27094401; PMCID: PMC4837600.

125. Restif C, Ibanez-Ventoso C, Vora MM, Guo S, Metaxas D, Driscoll M. CeleST: computer vision software for quantitative analysis of C. elegans swim behavior reveals novel features of locomotion. PLoS Comput Biol. 2014;10(7):e1003702. Epub 2014/07/18. doi: 10.1371/journal.pcbi.1003702. PubMed PMID: 25033081; PMCID: PMC4102393.

126. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet. 2014;46(3):310-5. Epub 2014/02/04. doi: 10.1038/ng.2892. PubMed PMID: 24487276; PMCID: PMC3992975.

127. Yamamoto GL, Aguena M, Gos M, Hung C, Pilch J, Fahiminiya S, Abramowicz A, Cristian I, Buscarilli M, Naslavsky MS, Malaquias AC, Zatz M, Bodamer O, Majewski J, Jorge AA, Pereira AC, Kim CA, Passos-Bueno MR, Bertola DR. Rare variants in SOS2 and LZTR1 are associated with Noonan syndrome. J Med Genet. 2015;52(6):413-21. Epub 2015/03/22. doi: 10.1136/jmedgenet-2015-103018. PubMed PMID: 25795793.

128. de Goede C, Yue WW, Yan G, Ariyaratnam S, Chandler KE, Downes L, Khan N, Mohan M, Lowe M, Banka S. Role of reverse phenotyping in interpretation of next generation sequencing data and a review of INPP5E related disorders. Eur J Paediatr Neurol. 2016;20(2):286-95. Epub 2016/01/11. doi: 10.1016/j.ejpn.2015.11.012. PubMed PMID: 26748598.

129. Bockelmann S, Menche D, Rudolph S, Bender T, Grond S, von Zezschwitz P, Muench SP, Wieczorek H, Huss M. Archazolid A binds to the equatorial region of the c-ring of the vacuolar H+-ATPase. J Biol Chem. 2010;285(49):38304-14. Epub 2010/10/05. doi: 10.1074/jbc.M110.137539. PubMed PMID: 20884613; PMCID: PMC2992264.

130. Kawasaki-Nishi S, Nishi T, Forgac M. Interacting helical surfaces of the transmembrane segments of subunits a and c' of the yeast V-ATPase defined by disulfide-mediated cross-linking. J Biol Chem. 2003;278(43):41908-13. Epub 2003/08/15. doi: 10.1074/jbc.M308026200. PubMed PMID: 12917411.

131. Nelson H, Nelson N. Disruption of genes encoding subunits of yeast vacuolar H(+)-ATPase causes conditional lethality. Proc Natl Acad Sci U S A. 1990;87(9):3503-7. Epub 1990/05/01. PubMed PMID: 2139726; PMCID: PMC53929.

132. Su Y, Blake-Palmer KG, Sorrell S, Javid B, Bowers K, Zhou A, Chang SH, Qamar S, Karet FE. Human H+ATPase a4 subunit mutations causing renal tubular acidosis reveal a role for interaction with phosphofructokinase-1. Am J Physiol Renal Physiol. 2008;295(4):F950-8. Epub 2008/07/18. doi: 10.1152/ajprenal.90258.2008. PubMed PMID: 18632794; PMCID: PMC2576143.

133. Zirngibl RA, Wang A, Yao Y, Manolson MF, Krueger J, Dupuis L, Mendoza-Londono R, Voronov I. Novel c.G630A TCIRG1 mutation causes aberrant splicing resulting in an unusually mild form of autosomal recessive osteopetrosis. J Cell Biochem. 2019;120(10):17180-93. Epub 2019/05/22. doi: 10.1002/jcb.28979. PubMed PMID: 31111556.

134. Ochotny N, Van Vliet A, Chan N, Yao Y, Morel M, Kartner N, von Schroeder HP, Heersche JN, Manolson MF. Effects of human a3 and a4 mutations that result in osteopetrosis and distal renal tubular acidosis on yeast V-ATPase expression and activity. J Biol Chem. 2006;281(36):26102-11. Epub 2006/07/15. doi: 10.1074/jbc.M601118200. PubMed PMID: 16840787.

135. Ohya Y, Umemoto N, Tanida I, Ohta A, Iida H, Anraku Y. Calcium-sensitive cls mutants of Saccharomyces cerevisiae showing a Pet- phenotype are ascribable to defects of vacuolar membrane H(+)-ATPase activity. J Biol Chem. 1991;266(21):13971-7. Epub 1991/07/25. PubMed PMID: 1830311.

136. Vaccaro A, Tauffenberger A, Aggad D, Rouleau G, Drapeau P, Parker JA. Mutant TDP-43 and FUS cause age-dependent paralysis and neurodegeneration in C. elegans. PLoS One. 2012;7(2):e31321. Epub 2012/03/01. doi: 10.1371/journal.pone.0031321. PubMed PMID: 22363618; PMCID: PMC3283630.

137. Mahoney TR, Luo S, Nonet ML. Analysis of synaptic transmission in Caenorhabditis elegans using an aldicarb-sensitivity assay. Nat Protoc. 2006;1(4):1772-7. Epub 2007/05/10. doi: 10.1038/nprot.2006.281. PubMed PMID: 17487159.

138. Vashlishan AB, Madison JM, Dybbs M, Bai J, Sieburth D, Ch'ng Q, Tavazoie M, Kaplan JM. An RNAi screen identifies genes that regulate GABA synapses. Neuron. 2008;58(3):346-61. Epub 2008/05/10. doi: 10.1016/j.neuron.2008.02.019. PubMed PMID: 18466746.

139. Hawasli AH, Saifee O, Liu C, Nonet ML, Crowder CM. Resistance to volatile anesthetics by mutations enhancing excitatory neurotransmitter release in Caenorhabditis elegans. Genetics. 2004;168(2):831-43. Epub 2004/10/30. doi: 10.1534/genetics.104.030502. PubMed PMID: 15514057; PMCID: PMC1448830.

140. Saifee O, Wei L, Nonet ML. The Caenorhabditis elegans unc-64 locus encodes a syntaxin that interacts genetically with synaptobrevin. Mol Biol Cell. 1998;9(6):1235-52. Epub 1998/06/17. doi: 10.1091/mbc.9.6.1235. PubMed PMID: 9614171; PMCID: PMC25346.

141. Dyle MC, Kolakada D, Cortazar MA, Jagannathan S. How to get away with nonsense: Mechanisms and consequences of escape from nonsense-mediated RNA decay. Wiley

Interdiscip Rev RNA. 2020;11(1):e1560. Epub 2019/07/31. doi: 10.1002/wrna.1560. PubMed PMID: 31359616.

142. McIntire SL, Reimer RJ, Schuske K, Edwards RH, Jorgensen EM. Identification and characterization of the vesicular GABA transporter. Nature. 1997;389(6653):870-6. Epub 1997/12/31 23:16. doi: 10.1038/39908. PubMed PMID: 9349821.

143. Bellocchio EE, Reimer RJ, Fremeau RT, Jr., Edwards RH. Uptake of glutamate into synaptic vesicles by an inorganic phosphate transporter. Science. 2000;289(5481):957-60. Epub 2000/08/11. doi: 10.1126/science.289.5481.957. PubMed PMID: 10938000.

144. Morel N, Poea-Guyon S. The membrane domain of vacuolar H(+)ATPase: a crucial player in neurotransmitter exocytotic release. Cell Mol Life Sci. 2015;72(13):2561-73. Epub 2015/03/22. doi: 10.1007/s00018-015-1886-2. PubMed PMID: 25795337.

145. Peters C, Bayer MJ, Buhler S, Andersen JS, Mann M, Mayer A. Trans-complex formation by proteolipid channels in the terminal phase of membrane fusion. Nature. 2001;409(6820):581-8. Epub 2001/02/24. doi: 10.1038/35054500. PubMed PMID: 11214310.

146. Wang D, Epstein D, Khalaf O, Srinivasan S, Williamson WR, Fayyazuddin A, Quiocho FA, Hiesinger PR. Ca2+-Calmodulin regulates SNARE assembly and spontaneous neurotransmitter release via v-ATPase subunit V0a1. J Cell Biol. 2014;205(1):21-31. Epub 2014/04/16. doi: 10.1083/jcb.201312109. PubMed PMID: 24733584; PMCID: PMC3987144.

147. Maxfield FR, McGraw TE. Endocytic recycling. Nat Rev Mol Cell Biol. 2004;5(2):121-32. Epub 2004/03/26. doi: 10.1038/nrm1315. PubMed PMID: 15040445.

148. Presley JF, Mayor S, McGraw TE, Dunn KW, Maxfield FR. Bafilomycin A1 treatment retards transferrin receptor recycling more than bulk membrane recycling. J Biol Chem. 1997;272(21):13929-36. Epub 1997/05/23. doi: 10.1074/jbc.272.21.13929. PubMed PMID: 9153255.

149. Zhang X, Bao L, Ma GQ. Sorting of neuropeptides and neuropeptide receptors into secretory pathways. Prog Neurobiol. 2010;90(2):276-83. Epub 2009/10/27. doi: 10.1016/j.pneurobio.2009.10.011. PubMed PMID: 19853638.

150. Khayat W, Hackett A, Shaw M, Ilie A, Dudding-Byth T, Kalscheuer VM, Christie L, Corbett MA, Juusola J, Friend KL, Kirmse BM, Gecz J, Field M, Orlowski J. A recurrent missense variant in SLC9A7 causes nonsyndromic X-linked intellectual disability with alteration of Golgi acidification and aberrant glycosylation. Hum Mol Genet. 2019;28(4):598-614. Epub 2018/10/20. doi: 10.1093/hmg/ddy371. PubMed PMID: 30335141; PMCID: PMC6360272.

151. Sobota JA, Back N, Eipper BA, Mains RE. Inhibitors of the V0 subunit of the vacuolar H+-ATPase prevent segregation of lysosomal- and secretory-pathway proteins. J Cell Sci. 2009;122(Pt 19):3542-53. Epub 2009/09/10. doi: 10.1242/jcs.034298. PubMed PMID: 19737820; PMCID: PMC2746133.

152. Nakamura N, Matsuura A, Wada Y, Ohsumi Y. Acidification of vacuoles is required for autophagic degradation in the yeast, Saccharomyces cerevisiae. J Biochem. 1997;121(2):338-44. Epub 1997/02/01. doi: 10.1093/oxfordjournals.jbchem.a021592. PubMed PMID: 9089409.

153. Fassio A, Falace A, Esposito A, Aprile D, Guerrini R, Benfenati F. Emerging Role of the Autophagy/Lysosomal Degradative Pathway in Neurodevelopmental Disorders With Epilepsy. Front Cell Neurosci. 2020;14:39. Epub 2020/04/02. doi: 10.3389/fncel.2020.00039. PubMed PMID: 32231521; PMCID: PMC7082311.

154. Saha S, Panigrahi DP, Patil S, Bhutia SK. Autophagy in health and disease: A comprehensive review. Biomed Pharmacother. 2018;104:485-95. Epub 2018/05/26. doi: 10.1016/j.biopha.2018.05.007. PubMed PMID: 29800913.

155. Fagerberg L, Hallstrom BM, Oksvold P, Kampf C, Djureinovic D, Odeberg J, Habuka M, Tahmasebpoor S, Danielsson A, Edlund K, Asplund A, Sjostedt E, Lundberg E, Szigyarto CA, Skogs M, Takanen JO, Berling H, Tegel H, Mulder J, Nilsson P, Schwenk JM, Lindskog C, Danielsson F, Mardinoglu A, Sivertsson A, von Feilitzen K, Forsberg M, Zwahlen M, Olsson I, Navani S, Huss M, Nielsen J, Ponten F, Uhlen M. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. Mol Cell Proteomics. 2014;13(2):397-406. Epub 2013/12/07. doi: 10.1074/mcp.M113.035600. PubMed PMID: 24309898; PMCID: PMC3916642.

156. Epi KC, Epilepsy Phenome/Genome P, Allen AS, Berkovic SF, Cossette P, Delanty N, Dlugos D, Eichler EE, Epstein MP, Glauser T, Goldstein DB, Han Y, Heinzen EL, Hitomi Y, Howell KB, Johnson MR, Kuzniecky R, Lowenstein DH, Lu YF, Madou MR, Marson AG, Mefford HC, Esmaeeli Nieh S, O'Brien TJ, Ottman R, Petrovski S, Poduri A, Ruzzo EK, Scheffer IE, Sherr EH, Yuskaitis CJ, Abou-Khalil B, Alldredge BK, Bautista JF, Berkovic SF, Boro A, Cascino GD, Consalvo D, Crumrine P, Devinsky O, Dlugos D, Epstein MP, Fiol M, Fountain NB, French J, Friedman D, Geller EB, Glauser T, Glynn S, Haut SR, Hayward J, Helmers SL, Joshi S, Kanner A, Kirsch HE, Knowlton RC, Kossoff EH, Kuperman R, Kuzniecky R, Lowenstein DH, McGuire SM, Motika PV, Novotny EJ, Ottman R, Paolicchi JM, Parent JM, Park K, Poduri A, Scheffer IE, Shellhaas RA, Sherr EH, Shih JJ, Singh R, Sirven J, Smith MC, Sullivan J, Lin Thio L, Venkat A, Vining EP, Von Allmen GK, Weisenberg JL, Widdess-Walsh P, Winawer MR. De novo mutations in epileptic encephalopathies. Nature. 2013;501(7466):217-21. Epub 2013/08/13. doi: 10.1038/nature12439. PubMed PMID: 23934111; PMCID: PMC3773011.

157. Zhang Y, Kong W, Gao Y, Liu X, Gao K, Xie H, Wu Y, Zhang Y, Wang J, Gao F, Wu X, Jiang Y. Gene Mutation Analysis in 253 Chinese Children with Unexplained Epilepsy and Intellectual/Developmental Disabilities. PLoS One. 2015;10(11):e0141782. Epub 2015/11/07. doi: 10.1371/journal.pone.0141782. PubMed PMID: 26544041; PMCID: PMC4636363.

158. Papuc SM, Abela L, Steindl K, Begemann A, Simmons TL, Schmitt B, Zweier M, Oneda B, Socher E, Crowther LM, Wohlrab G, Gogoll L, Poms M, Seiler M, Papik M, Baldinger R, Baumer A, Asadollahi R, Kroell-Seger J, Schmid R, Iff T, Schmitt-Mechelke T, Otten K, Hackenberg A, Addor MC, Klein A, Azzarello-Burri S, Sticht H, Joset P, Plecko B, Rauch A. The role of recessive inheritance in early-onset epileptic encephalopathies: a combined whole-exome sequencing and copy number study. Eur J Hum Genet. 2019;27(3):408-21. Epub 2018/12/16. doi: 10.1038/s41431-018-0299-8. PubMed PMID: 30552426; PMCID: PMC6460568.

159. Minardi R, Licchetta L, Baroni MC, Pippucci T, Stipa C, Mostacci B, Severi G, Toni F, Bergonzini L, Carelli V, Seri M, Tinuper P, Bisulli F. Whole-exome sequencing in adult patients with developmental and epileptic encephalopathy: It is never too late. Clin Genet. 2020. Epub 2020/07/30. doi: 10.1111/cge.13823. PubMed PMID: 32725632.

160. Johnston HR, Chopra P, Wingo TS, Patel V, International Consortium on B, Behavior in 22q11.2 Deletion S, Epstein MP, Mulle JG, Warren ST, Zwick ME, Cutler DJ. PEMapper and PECaller provide a simplified approach to whole-genome sequencing. Proc Natl Acad Sci U S A. 2017;114(10):E1923-E32. Epub 2017/02/23. doi: 10.1073/pnas.1618065114. PubMed PMID: 28223510; PMCID: PMC5347547.

161. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559-75. Epub 2007/08/19. doi: 10.1086/519795. PubMed PMID: 17701901; PMCID: PMC1950838.

162. Kotlar AV, Trevino CE, Zwick ME, Cutler DJ, Wingo TS. Bystro: rapid online variant annotation and natural-language filtering at whole-genome scale. Genome Biol. 2018;19(1):14. Epub 2018/02/08. doi: 10.1186/s13059-018-1387-3. PubMed PMID: 29409527; PMCID: PMC5801807.

163. Scott EM, Halees A, Itan Y, Spencer EG, He Y, Azab MA, Gabriel SB, Belkadi A, Boisson B, Abel L, Clark AG, Greater Middle East Variome C, Alkuraya FS, Casanova JL, Gleeson JG. Characterization of Greater Middle Eastern genetic variation for enhanced disease gene discovery. Nat Genet. 2016;48(9):1071-6. Epub 2016/07/19. doi: 10.1038/ng.3592. PubMed PMID: 27428751; PMCID: PMC5019950.

164. GenomeAsia KC. The GenomeAsia 100K Project enables genetic discoveries across Asia. Nature. 2019;576(7785):106-11. Epub 2019/12/06. doi: 10.1038/s41586-019-1793-z. PubMed PMID: 31802016; PMCID: PMC7054211.

165. Liu H, Pecka JL, Zhang Q, Soukup GA, Beisel KW, He DZ. Characterization of transcriptomes of cochlear inner and outer hair cells. J Neurosci. 2014;34(33):11085-95. Epub 2014/08/15. doi: 10.1523/JNEUROSCI.1690-14.2014. PubMed PMID: 25122905; PMCID: PMC4131018.

166. Ebermann I, Phillips JB, Liebau MC, Koenekoop RK, Schermer B, Lopez I, Schafer E, Roux AF, Dafinger C, Bernd A, Zrenner E, Claustres M, Blanco B, Nurnberg G, Nurnberg P, Ruland R, Westerfield M, Benzing T, Bolz HJ. PDZD7 is a modifier of retinal disease and a contributor to digenic Usher syndrome. J Clin Invest. 2010;120(6):1812-23. Epub 2010/05/05. doi: 10.1172/JCI39715. PubMed PMID: 20440071; PMCID: PMC2877930.

167. Vona B, Lechno S, Hofrichter MA, Hopf S, Laig AK, Haaf T, Keilmann A, Zechner U, Bartsch O. Confirmation of PDZD7 as a Nonsyndromic Hearing Loss Gene. Ear Hear. 2016;37(4):e238-46. Epub 2016/02/06. doi: 10.1097/AUD.0000000000000278. PubMed PMID: 26849169.

168. Booth KT, Azaiez H, Kahrizi K, Simpson AC, Tollefson WT, Sloan CM, Meyer NC, Babanejad M, Ardalani F, Arzhangi S, Schnieders MJ, Najmabadi H, Smith RJ. PDZD7 and hearing loss: More than just a modifier. Am J Med Genet A. 2015;167A(12):2957-65. Epub 2015/09/30. doi: 10.1002/ajmg.a.37274. PubMed PMID: 26416264; PMCID: PMC4741280.

169. Lefeber DJ, Morava E, Jaeken J. How to find and diagnose a CDG due to defective N-glycosylation. J Inherit Metab Dis. 2011;34(4):849-52. Epub 2011/07/09. doi: 10.1007/s10545-011-9370-0. PubMed PMID: 21739167; PMCID: PMC3137781.

170. Jaeken J, Lefeber D, Matthijs G. Clinical utility gene card for: ALG6 defective congenital disorder of glycosylation. Eur J Hum Genet. 2015;23(2). Epub 2014/07/24. doi: 10.1038/ejhg.2014.146. PubMed PMID: 25052310; PMCID: PMC4297893.

171. Morava E, Tiemes V, Thiel C, Seta N, de Lonlay P, de Klerk H, Mulder M, Rubio-Gozalbo E, Visser G, van Hasselt P, Horovitz DDG, de Souza CFM, Schwartz IVD, Green A, Al-Owain M, Uziel G, Sigaudy S, Chabrol B, van Spronsen FJ, Steinert M, Komini E, Wurm D, Bevot A, Ayadi A, Huijben K, Dercksen M, Witters P, Jaeken J, Matthijs G, Lefeber DJ, Wevers RA. ALG6-CDG: a recognizable phenotype with epilepsy, proximal muscle weakness, ataxia and behavioral and limb anomalies. J Inherit Metab Dis. 2016;39(5):713-23. Epub 2016/06/12. doi: 10.1007/s10545-016-9945-x. PubMed PMID: 27287710.

172. Guo W, Schafer S, Greaser ML, Radke MH, Liss M, Govindarajan T, Maatz H, Schulz H, Li S, Parrish AM, Dauksaite V, Vakeel P, Klaassen S, Gerull B, Thierfelder L, Regitz-Zagrosek V, Hacker TA, Saupe KW, Dec GW, Ellinor PT, MacRae CA, Spallek B, Fischer R, Perrot A, Ozcelik C, Saar K, Hubner N, Gotthardt M. RBM20, a gene for hereditary cardiomyopathy, regulates titin splicing. Nat Med. 2012;18(5):766-73. Epub 2012/04/03. doi: 10.1038/nm.2693. PubMed PMID: 22466703; PMCID: PMC3569865.

173. Lennermann D, Backs J, van den Hoogenhof MMG. New Insights in RBM20 Cardiomyopathy. Curr Heart Fail Rep. 2020;17(5):234-46. Epub 2020/08/14. doi: 10.1007/s11897-020-00475-x. PubMed PMID: 32789749; PMCID: PMC7495990.

174. Parikh VN, Caleshu C, Reuter C, Lazzeroni LC, Ingles J, Garcia J, McCaleb K, Adesiyun T, Sedaghat-Hamedani F, Kumar S, Graw S, Gigli M, Stolfo D, Dal Ferro M, Ing AY, Nussbaum R, Funke B, Wheeler MT, Hershberger RE, Cook S, Steinmetz LM, Lakdawala NK, Taylor MRG, Mestroni L, Merlo M, Sinagra G, Semsarian C, Meder B, Judge DP, Ashley E. Regional Variation in RBM20 Causes a Highly Penetrant Arrhythmogenic Cardiomyopathy. Circ Heart Fail. 2019;12(3):e005371. Epub 2019/03/16. doi: 10.1161/CIRCHEARTFAILURE.118.005371. PubMed PMID: 30871351; PMCID: PMC6422044.

175. Brauch KM, Karst ML, Herron KJ, de Andrade M, Pellikka PA, Rodeheffer RJ, Michels VV, Olson TM. Mutations in ribonucleic acid binding protein gene cause familial dilated cardiomyopathy. J Am Coll Cardiol. 2009;54(10):930-41. Epub 2009/08/29. doi: 10.1016/j.jacc.2009.05.038. PubMed PMID: 19712804; PMCID: PMC2782634.

176. van den Hoogenhof MMG, Beqqali A, Amin AS, van der Made I, Aufiero S, Khan MAF, Schumacher CA, Jansweijer JA, van Spaendonck-Zwarts KY, Remme CA, Backs J, Verkerk AO, Baartscheer A, Pinto YM, Creemers EE. RBM20 Mutations Induce an Arrhythmogenic Dilated Cardiomyopathy Related to Disturbed Calcium Handling. Circulation. 2018;138(13):1330-42. Epub 2018/04/14. doi: 10.1161/CIRCULATIONAHA.117.031947. PubMed PMID: 29650543.

177. Liu S, Xie Y, Zhang H, Feng Z, Huang J, Huang J, Hu S, Wei Y. Multiple genetic variants in adolescent patients with left ventricular noncompaction cardiomyopathy. Int J Cardiol. 2020;302:117-23. Epub 2020/01/11. doi: 10.1016/j.ijcard.2019.12.001. PubMed PMID: 31918855.

178. Hazebroek MR, Krapels I, Verdonschot J, van den Wijngaard A, Vanhoutte E, Hoos M, Snijders L, van Montfort L, Witjens M, Dennert R, Crijns H, Brunner-La Rocca HP, Brunner HG, Heymans S. Prevalence of Pathogenic Gene Mutations and Prognosis Do Not Differ in Isolated

Left Ventricular Dysfunction Compared With Dilated Cardiomyopathy. Circ Heart Fail. 2018;11(3):e004682. Epub 2018/03/16. doi: 10.1161/CIRCHEARTFAILURE.117.004682. PubMed PMID: 29540472.

179. Watanabe T, Kimura A, Kuroyanagi H. Alternative Splicing Regulator RBM20 and Cardiomyopathy. Front Mol Biosci. 2018;5:105. Epub 2018/12/14. doi: 10.3389/fmolb.2018.00105. PubMed PMID: 30547036; PMCID: PMC6279932.

180. Bozarth X, Dines JN, Cong Q, Mirzaa GM, Foss K, Lawrence Merritt J, 2nd, Thies J, Mefford HC, Novotny E. Expanding clinical phenotype in CACNA1C related disorders: From neonatal onset severe epileptic encephalopathy to late-onset epilepsy. Am J Med Genet A. 2018;176(12):2733-9. Epub 2018/12/05. doi: 10.1002/ajmg.a.40657. PubMed PMID: 30513141; PMCID: PMC6312477.

181. Gordon A, Salomon D, Barak N, Pen Y, Tsoory M, Kimchi T, Peles E. Expression of Cntnap2 (Caspr2) in multiple levels of sensory systems. Mol Cell Neurosci. 2016;70:42-53. Epub 2015/12/10. doi: 10.1016/j.mcn.2015.11.012. PubMed PMID: 26647347.

182. Strauss KA, Puffenberger EG, Huentelman MJ, Gottlieb S, Dobrin SE, Parod JM, Stephan DA, Morton DH. Recessive symptomatic focal epilepsy and mutant contactin-associated protein-like 2. N Engl J Med. 2006;354(13):1370-7. Epub 2006/03/31. doi: 10.1056/NEJMoa052773. PubMed PMID: 16571880.

183. Zweier C, de Jong EK, Zweier M, Orrico A, Ousager LB, Collins AL, Bijlsma EK, Oortveld MA, Ekici AB, Reis A, Schenck A, Rauch A. CNTNAP2 and NRXN1 are mutated in autosomal-recessive Pitt-Hopkins-like mental retardation and determine the level of a common synaptic protein in Drosophila. Am J Hum Genet. 2009;85(5):655-66. Epub 2009/11/10. doi: 10.1016/j.ajhg.2009.10.004. PubMed PMID: 19896112; PMCID: PMC2775834.

184. Rodenas-Cuadrado P, Pietrafusa N, Francavilla T, La Neve A, Striano P, Vernes SC. Characterisation of CASPR2 deficiency disorder--a syndrome involving autism, epilepsy and language impairment. BMC Med Genet. 2016;17:8. Epub 2016/02/05. doi: 10.1186/s12881-016-0272-8. PubMed PMID: 26843181; PMCID: PMC4739328.

185. Watson CM, Crinnion LA, Tzika A, Mills A, Coates A, Pendlebury M, Hewitt S, Harrison SM, Daly C, Roberts P, Carr IM, Sheridan EG, Bonthron DT. Diagnostic whole genome sequencing and split-read mapping for nucleotide resolution breakpoint identification in CNTNAP2 deficiency syndrome. Am J Med Genet A. 2014;164A(10):2649-55. Epub 2014/07/22. doi: 10.1002/ajmg.a.36679. PubMed PMID: 25045150.

186. Smogavec M, Cleall A, Hoyer J, Lederer D, Nassogne MC, Palmer EE, Deprez M, Benoit V, Maystadt I, Noakes C, Leal A, Shaw M, Gecz J, Raymond L, Reis A, Shears D, Brockmann K, Zweier C. Eight further individuals with intellectual disability and epilepsy carrying bi-allelic CNTNAP2 aberrations allow delineation of the mutational and phenotypic spectrum. J Med Genet. 2016;53(12):820-7. Epub 2016/07/22. doi: 10.1136/jmedgenet-2016-103880. PubMed PMID: 27439707.

187. Lindy AS, Stosser MB, Butler E, Downtain-Pickersgill C, Shanmugham A, Retterer K, Brandt T, Richard G, McKnight DA. Diagnostic outcomes for genetic testing of 70 genes in 8565 patients with epilepsy and neurodevelopmental disorders. Epilepsia. 2018;59(5):1062-71. Epub 2018/04/15. doi: 10.1111/epi.14074. PubMed PMID: 29655203.

188. Gregor A, Albrecht B, Bader I, Bijlsma EK, Ekici AB, Engels H, Hackmann K, Horn D, Hoyer J, Klapecki J, Kohlhase J, Maystadt I, Nagl S, Prott E, Tinschert S, Ullmann R, Wohlleber E, Woods G, Reis A, Rauch A, Zweier C. Expanding the clinical spectrum associated with defects in CNTNAP2 and NRXN1. BMC Med Genet. 2011;12:106. Epub 2011/08/11. doi: 10.1186/1471-2350-12-106. PubMed PMID: 21827697; PMCID: PMC3162517.

189. Penagarikano O, Abrahams BS, Herman EI, Winden KD, Gdalyahu A, Dong H, Sonnenblick LI, Gruver R, Almajano J, Bragin A, Golshani P, Trachtenberg JT, Peles E, Geschwind DH. Absence of CNTNAP2 leads to epilepsy, neuronal migration abnormalities, and core autism-related deficits. Cell. 2011;147(1):235-46. Epub 2011/10/04. doi: 10.1016/j.cell.2011.08.040. PubMed PMID: 21962519; PMCID: PMC3390029.

190. Wang J, Zhao J. MicroRNA Dysregulation in Epilepsy: From Pathogenetic Involvement to Diagnostic Biomarker and Therapeutic Agent Development. Front Mol Neurosci. 2021;14:650372. Epub 2021/03/30. doi: 10.3389/fnmol.2021.650372. PubMed PMID: 33776649; PMCID: PMC7994516.

191. Yao RW, Wang Y, Chen LL. Cellular functions of long noncoding RNAs. Nat Cell Biol. 2019;21(5):542-51. Epub 2019/05/03. doi: 10.1038/s41556-019-0311-8. PubMed PMID: 31048766.

192. Matharu N, Ahituv N. Modulating gene regulation to treat genetic disorders. Nat Rev Drug Discov. 2020;19(11):757-75. Epub 2020/10/07. doi: 10.1038/s41573-020-0083-7. PubMed PMID: 33020616.

193. Colasante G, Lignani G, Brusco S, Di Berardino C, Carpenter J, Giannelli S, Valassina N, Bido S, Ricci R, Castoldi V, Marenna S, Church T, Massimino L, Morabito G, Benfenati F, Schorge S, Leocani L, Kullmann DM, Broccoli V. dCas9-Based Scn1a Gene Activation Restores Inhibitory Interneuron Excitability and Attenuates Seizures in Dravet Syndrome Mice. Mol Ther. 2020;28(1):235-53. Epub 2019/10/15. doi: 10.1016/j.ymthe.2019.08.018. PubMed PMID: 31607539; PMCID: PMC6952031.

194. Butler KM, da Silva C, Shafir Y, Weisfeld-Adams JD, Alexander JJ, Hegde M, Escayg A. De novo and inherited SCN8A epilepsy mutations detected by gene panel analysis. Epilepsy Res. 2017;129:17-25. Epub 2016/11/23. doi: 10.1016/j.eplepsyres.2016.11.002. PubMed PMID: 27875746; PMCID: PMC5321682.

195. Yan Q, Weyn-Vanhentenryck SM, Wu J, Sloan SA, Zhang Y, Chen K, Wu JQ, Barres BA, Zhang C. Systematic discovery of regulated and conserved alternative exons in the mammalian brain reveals NMD modulating chromatin regulators. Proc Natl Acad Sci U S A. 2015;112(11):3445-50. Epub 2015/03/05. doi: 10.1073/pnas.1502849112. PubMed PMID: 25737549; PMCID: PMC4371929.

196. Zhang X, Chen MH, Wu X, Kodani A, Fan J, Doan R, Ozawa M, Ma J, Yoshida N, Reiter JF, Black DL, Kharchenko PV, Sharp PA, Walsh CA. Cell-Type-Specific Alternative Splicing Governs Cell Fate in the Developing Cerebral Cortex. Cell. 2016;166(5):1147-62 e15. Epub 2016/08/28. doi: 10.1016/j.cell.2016.07.025. PubMed PMID: 27565344; PMCID: PMC5248659.

197. Carvill GL, Engel KL, Ramamurthy A, Cochran JN, Roovers J, Stamberger H, Lim N, Schneider AL, Hollingsworth G, Holder DH, Regan BM, Lawlor J, Lagae L, Ceulemans B, Bebin EM, Nguyen J, EuroEpinomics Rare Epilepsy Syndrome M-AE, Dravet Working G, Barsh GS,

Weckhuysen S, Meisler M, Berkovic SF, De Jonghe P, Scheffer IE, Myers RM, Cooper GM, Mefford HC. Aberrant Inclusion of a Poison Exon Causes Dravet Syndrome and Related SCN1A-Associated Genetic Epilepsies. Am J Hum Genet. 2018;103(6):1022-9. Epub 2018/12/12. doi: 10.1016/j.ajhg.2018.10.023. PubMed PMID: 30526861; PMCID: PMC6288405.

198. Dou Y, Kwon M, Rodin RE, Cortes-Ciriano I, Doan R, Luquette LJ, Galor A, Bohrson C, Walsh CA, Park PJ. Accurate detection of mosaic variants in sequencing data without matched controls. Nat Biotechnol. 2020;38(3):314-9. Epub 2020/01/08. doi: 10.1038/s41587-019-0368-8. PubMed PMID: 31907404; PMCID: PMC7065972.

199. Ribierre T, Deleuze C, Bacq A, Baldassari S, Marsan E, Chipaux M, Muraca G, Roussel D, Navarro V, Leguern E, Miles R, Baulac S. Second-hit mosaic mutation in mTORC1 repressor DEPDC5 causes focal cortical dysplasia-associated epilepsy. J Clin Invest. 2018;128(6):2452-8. Epub 2018/05/01. doi: 10.1172/JCI99384. PubMed PMID: 29708508; PMCID: PMC5983335.

200. Bennett MF, Hildebrand MS, Kayumi S, Corbett MA, Gupta S, Ye Z, Krivanek M, Burgess R, Henry OJ, Damiano JA, Boys A, Gecz J, Bahlo M, Scheffer IE, Berkovic SF. Evidence for a Dual-Pathway, 2-Hit Genetic Model for Focal Cortical Dysplasia and Epilepsy. Neurol Genet. 2022;8(1):e652. Epub 2022/02/01. doi: 10.1212/NXG.0000000000000652. PubMed PMID: 35097204; PMCID: PMC8789218.

201. Hermida MA, Dinesh Kumar J, Leslie NR. GSK3 and its interactions with the PI3K/AKT/mTOR signalling network. Adv Biol Regul. 2017;65:5-15. Epub 2017/07/18. doi: 10.1016/j.jbior.2017.06.003. PubMed PMID: 28712664.

202. Osteresch C, Bender T, Grond S, von Zezschwitz P, Kunze B, Jansen R, Huss M, Wieczorek H. The binding site of the V-ATPase inhibitor apicularen is in the vicinity of those for bafilomycin and archazolid. J Biol Chem. 2012;287(38):31866-76. Epub 2012/07/21. doi: 10.1074/jbc.M112.372169. PubMed PMID: 22815478; PMCID: PMC3442520.

203. Wang R, Wang J, Hassan A, Lee CH, Xie XS, Li X. Molecular basis of V-ATPase inhibition by bafilomycin A1. Nat Commun. 2021;12(1):1782. Epub 2021/03/21. doi: 10.1038/s41467-021-22111-5. PubMed PMID: 33741963; PMCID: PMC7979754.

204. Duarte-Bueno LM, Álvarez-Pabón Y, Contreras-García GA. De novo 16p13.3-p12.3 duplication in a child with syndromic developmental delay. Gene Reports. 2020;20:100690. doi: https://doi.org/10.1016/j.genrep.2020.100690.

205. Ciaccio C, Tucci A, Scuvera G, Estienne M, Esposito S, Milani D. 16p13 microduplication without CREBBP involvement: Moving toward a phenotype delineation. Eur J Med Genet. 2017;60(3):159-62. Epub 2016/12/23. doi: 10.1016/j.ejmg.2016.12.006. PubMed PMID: 28007608.

206. Tassano E, Alpigiani MG, Calcagno A, Salvati P, De Miglio L, Fiorio P, Cuoco C, Gimelli G. Clinical and molecular delineation of a 16p13.2p13.13 microduplication. Eur J Med Genet. 2015;58(3):194-8. Epub 2015/01/18. doi: 10.1016/j.ejmg.2014.12.016. PubMed PMID: 25596524.

207. Harkin LA, Bowser DN, Dibbens LM, Singh R, Phillips F, Wallace RH, Richards MC, Williams DA, Mulley JC, Berkovic SF, Scheffer IE, Petrou S. Truncation of the GABA(A)-receptor gamma2 subunit in a family with generalized epilepsy with febrile seizures plus. Am J

Hum Genet. 2002;70(2):530-6. Epub 2001/12/19. doi: 10.1086/338710. PubMed PMID: 11748509; PMCID: PMC384926.

208. Huang X, Zhou C, Tian M, Kang JQ, Shen W, Verdier K, Pimenta A, MacDonald RL. Overexpressing wild-type gamma2 subunits rescued the seizure phenotype in Gabrg2(+/Q390X) Dravet syndrome mice. Epilepsia. 2017;58(8):1451-61. Epub 2017/06/07. doi: 10.1111/epi.13810. PubMed PMID: 28586508; PMCID: PMC5554098.

209. HGNC Database, HUGO Gene Nomenclature Committee (HGNC), European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Hinxton, Cambridge, United Kingdom; [cited 2022 February]. Available from: www.genenames.org.

210. Reggiori F, Klionsky DJ. Autophagic processes in yeast: mechanism, machinery and regulation. Genetics. 2013;194(2):341-61. Epub 2013/06/05. doi: 10.1534/genetics.112.149013. PubMed PMID: 23733851; PMCID: PMC3664846.

211. Taki AC, Byrne JJ, Boag PR, Jabbar A, Gasser RB. Practical High-Throughput Method to Screen Compounds for Anthelmintic Activity against Caenorhabditis elegans. Molecules. 2021;26(14). Epub 2021/07/25. doi: 10.3390/molecules26144156. PubMed PMID: 34299431; PMCID: PMC8305057.

212. Kwan P, Sander JW. The natural history of epilepsy: an epidemiological view. J Neurol Neurosurg Psychiatry. 2004;75(10):1376-81. Epub 2004/09/21. doi: 10.1136/jnnp.2004.045690. PubMed PMID: 15377680; PMCID: PMC1738749.

213. Wenger AM, Guturu H, Bernstein JA, Bejerano G. Systematic reanalysis of clinical exome data yields additional diagnoses: implications for providers. Genet Med. 2017;19(2):209-14. Epub 2016/07/22. doi: 10.1038/gim.2016.88. PubMed PMID: 27441994.

214. Deignan JL, Chung WK, Kearney HM, Monaghan KG, Rehder CW, Chao EC, Committee ALQA. Points to consider in the reevaluation and reanalysis of genomic test results: a statement of the American College of Medical Genetics and Genomics (ACMG). Genet Med. 2019;21(6):1267-70. Epub 2019/04/25. doi: 10.1038/s41436-019-0478-1. PubMed PMID: 31015575; PMCID: PMC6559819.

215. Logsdon GA, Vollger MR, Eichler EE. Long-read human genome sequencing and its applications. Nat Rev Genet. 2020;21(10):597-614. Epub 2020/06/07. doi: 10.1038/s41576-020-0236-x. PubMed PMID: 32504078; PMCID: PMC7877196.

216. Schlicher L, Wissler M, Preiss F, Brauns-Schubert P, Jakob C, Dumit V, Borner C, Dengjel J, Maurer U. SPATA2 promotes CYLD activity and regulates TNF-induced NF-kappaB signaling and cell death. EMBO Rep. 2016;17(10):1485-97. Epub 2016/07/28. doi: 10.15252/embr.201642592. PubMed PMID: 27458237; PMCID: PMC5048381.

217. Wagner SA, Satpathy S, Beli P, Choudhary C. SPATA2 links CYLD to the TNF-alpha receptor signaling complex and modulates the receptor signaling outcomes. EMBO J. 2016;35(17):1868-84. Epub 2016/06/17. doi: 10.15252/embj.201694300. PubMed PMID: 27307491; PMCID: PMC5007551.

218. Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson A, Kampf C, Sjostedt E, Asplund A, Olsson I, Edlund K, Lundberg E, Navani S, Szigyarto CA, Odeberg J, Djureinovic D, Takanen JO, Hober S, Alm T, Edqvist PH, Berling H, Tegel H, Mulder J, Rockberg J, Nilsson P, Schwenk JM, Hamsten M, von Feilitzen K, Forsberg M, Persson L, Johansson F, Zwahlen M, von Heijne G, Nielsen J, Ponten F. Proteomics. Tissue-based map of the human proteome. Science. 2015;347(6220):1260419. Epub 2015/01/24. doi: 10.1126/science.1260419. PubMed PMID: 25613900.

219. Patel DC, Wallis G, Dahle EJ, McElroy PB, Thomson KE, Tesi RJ, Szymkowski DE, West PJ, Smeal RM, Patel M, Fujinami RS, White HS, Wilcox KS. Hippocampal TNFalpha Signaling Contributes to Seizure Generation in an Infection-Induced Mouse Model of Limbic Epilepsy. eNeuro. 2017;4(2). Epub 2017/05/13. doi: 10.1523/ENEURO.0105-17.2017. PubMed PMID: 28497109; PMCID: PMC5422919.

220. Scott IC, Blitz IL, Pappano WN, Imamura Y, Clark TG, Steiglitz BM, Thomas CL, Maas SA, Takahara K, Cho KW, Greenspan DS. Mammalian BMP-1/Tolloid-related metalloproteinases, including novel family member mammalian Tolloid-like 2, have differential enzymatic activities and distributions of expression relevant to patterning and skeletogenesis. Dev Biol. 1999;213(2):283-300. Epub 1999/09/10. doi: 10.1006/dbio.1999.9383. PubMed PMID: 10479448.

221. Bachiller D, Klingensmith J, Kemp C, Belo JA, Anderson RM, May SR, McMahon JA, McMahon AP, Harland RM, Rossant J, De Robertis EM. The organizer factors Chordin and Noggin are required for mouse forebrain development. Nature. 2000;403(6770):658-61. Epub 2000/02/25. doi: 10.1038/35001072. PubMed PMID: 10688202.

222. Fukuda T, Sugita S, Inatome R, Yanagi S. CAMDI, a novel disrupted in schizophrenia 1 (DISC1)-binding protein, is required for radial migration. J Biol Chem. 2010;285(52):40554-61. Epub 2010/10/20. doi: 10.1074/jbc.M110.179481. PubMed PMID: 20956536; PMCID: PMC3003354.

223. des Portes V, Pinard JM, Billuart P, Vinet MC, Koulakoff A, Carrie A, Gelot A, Dupuis E, Motte J, Berwald-Netter Y, Catala M, Kahn A, Beldjord C, Chelly J. A novel CNS gene required for neuronal migration and involved in X-linked subcortical laminar heterotopia and lissencephaly syndrome. Cell. 1998;92(1):51-61. Epub 1998/03/07. doi: 10.1016/s0092-8674(00)80898-3. PubMed PMID: 9489699.

224. Willemsen MH, Vissers LE, Willemsen MA, van Bon BW, Kroes T, de Ligt J, de Vries BB, Schoots J, Lugtenberg D, Hamel BC, van Bokhoven H, Brunner HG, Veltman JA, Kleefstra T. Mutations in DYNC1H1 cause severe intellectual disability with neuronal migration defects. J Med Genet. 2012;49(3):179-83. Epub 2012/03/01. doi: 10.1136/jmedgenet-2011-100542. PubMed PMID: 22368300.

225. Hou Q, Wu J, Zhao Y, Wang X, Jiang F, Chen DN, Zheng R, Men M, Li JD. Genotypic and phenotypic spectrum of CCDC141 variants in a Chinese cohort with congenital hypogonadotropic hypogonadism. Eur J Endocrinol. 2020;183(3):245-54. Epub 2020/06/11. doi: 10.1530/EJE-19-1018. PubMed PMID: 32520725.

226. Broer S, Bailey CG, Kowalczuk S, Ng C, Vanslambrouck JM, Rodgers H, Auray-Blais C, Cavanaugh JA, Broer A, Rasko JE. Iminoglycinuria and hyperglycinuria are discrete human phenotypes resulting from complex mutations in proline and glycine transporters. J Clin Invest.

2008;118(12):3881-92. Epub 2008/11/27. doi: 10.1172/JCI36625. PubMed PMID: 19033659; PMCID: PMC2579706.

227. Janssen LME, Averink TV, Blomen VA, Brummelkamp TR, Medema RH, Raaijmakers JA. Loss of Kif18A Results in Spindle Assembly Checkpoint Activation at Microtubule-Attached Kinetochores. Curr Biol. 2018;28(17):2685-96 e4. Epub 2018/08/21. doi: 10.1016/j.cub.2018.06.026. PubMed PMID: 30122526.

228. Stumpff J, von Dassow G, Wagenbach M, Asbury C, Wordeman L. The kinesin-8 motor Kif18A suppresses kinetochore movements to control mitotic chromosome alignment. Dev Cell. 2008;14(2):252-62. Epub 2008/02/13. doi: 10.1016/j.devcel.2007.11.014. PubMed PMID: 18267093; PMCID: PMC2267861.

Appendix A

Supporting Data for Chapter 3

Subunit	Gene	Genomic Region	gnomAD Constraints[a]		Expression Profile[b]					Phenotype and Inheritance	MIM #	Reference
			Missense	Loss of Function	Brain	Kidney	Skin	Bone Marrow	Testis			
V_0a	*ATP6V0A1*	17q21.2	o/e = 0.52	o/e = 0.14	++	+	+	+	+	AD/AR DEE	NA	Aoto *et al.* [101]
	ATP6V0A2	12q24.31	o/e = 0.88	o/e = 0.54	+	+	+	+	+	AR Cutis Laxa (Type IIA)	219200	Kornak *et al.* [91]
	TCIRG1	11q13.2	o/e = 0.92	o/e = 0.63	+	+	+	++	+	AR Osteopetrosis	259700	Frattini *et al.* [96]
	ATP6V0A4	7q34	o/e = 0.89	o/e = 0.76	-	++	+	-	+	AR Distal RTA	602722	Smith *et al.* [92]
V_0c	*ATP6V0C*	16p13.3	o/e = 0.19	o/e = 0.00	+++	+++	+	++	++	AD Epilepsy	NA	Ittiwut *et al* [110]
V_0c"	*ATP6V0B*	1p34.1	o/e = 0.53	o/e = 0.00	++	++	+	+++	+			
V_0d	*ATP6V0D1*	16q22.1	o/e = 0.45	o/e = 0.06	+++	++	+	++	++			
	ATP6V0D2	8q21.3	o/e = 1.16	o/e = 0.94	-	+	-	-	-	AR Skin and Joint Laxity	NA	Alazami *et al.* [90]
V_0e	*ATP6V0E1*	5q35.1	o/e = 0.94	o/e = 0.20	++	+++	++	++	++			
	ATP6V0E2	7q36.1	o/e = 0.85	o/e = 0.25	+++	+	+	+	+			
V_1A	*ATP6V1A*	3q13.31	o/e = 0.49	o/e = 0.18	+++	++	+	+	+	AR Cutis Laxa (Type IID) AD/AR IECEE 3	617403 618012	van Damme *et al.* [89] Fassio *et al.* [100]
V_1B	*ATP6V1B1*	2p13.3	o/e = 0.96	o/e = 0.64	+	++	+	-	+	AR RTA with deafness	267300	Karet *et al.* [93]
	ATP6V1B2	8p21.3	o/e = 0.59	o/e = 0.13	++	++	+	+	+	AD DDOD Syndrome AD Zimmerman-Laband Syndrome 2 AD Epilepsy AD DOORS Syndrome	124480 616455 NA NA	Yuan *et al.* [97] Kortum *et al.* [98] Shaw *et al.* [103] Beauregard-Lacroix *et al.* [104]
V_1C	*ATP6V1C1*	8q22.3	o/e = 0.65	o/e = 0.45	++	+	+	+	+			
	ATP6V1C2	2p25.1	o/e = 0.85	o/e = 0.62	+	+	++	+	+	AR Distal RTA	NA	Jobst-Schwan *et al.* [94]
V_1D	*ATP6V1D*	14q23.3	o/e = 0.72	o/e = 0.53	++	++	+	+	+			
V_1E	*ATP6V1E1*	22q11.21	o/e = 0.59	o/e = 0.20	+++	+++	++	++	++	AR Cutis Laxa (Type IIC)	617402	Alazami *et al.* [90]
	ATP6V1E2	2p21	o/e = 1.14	o/e = 0.59	+	+	+	+	+			
V_1F	*ATP6V1F*	7q32.1	o/e = 0.81	o/e = 0.67	+++	+++	++	++	+++			
V_1G	*ATP6V1G1*	9q32	o/e = 0.77	o/e = 0.40	++	+++	+	++	++			
	ATP6V1G2	6p21.33	o/e = 0.72	o/e = 0.59	+++	+	+	+	+			
	ATP6V1G3	1q31.3	o/e = 0.90	o/e = 0.40	-	+	-	-	-			
V_1H	*ATP6V1H*	8q11.23	o/e = 0.78	o/e = 0.39	++	++	+	+	+	AD Short Stature and Bone Loss	NA	Zhang *et al.* [99]

Table A1 (preceding page). Genes encoding V-ATPase subunits in humans.
[a]gnomAD constraint metrics are a ratio of observed variants (o) over expected number (e) of variants. Values less than 1 indicate less variation is seen than expected, suggestive of evolutionary constraint or intolerance to variation.
[b]Based on human RNA-seq data from normal tissue, accessed through NCBI Gene [155]. Expression was classified into four categories based on mean RPKM values: -, 0.0 ≤ RPKM ≤ 0.99; +, 0.1 ≤ RPKM ≤ 24.99; ++, 25 ≤ RPKM ≤ 49.99; +++, RPKM ≥ 50.
Shaded rows represent genes with no known disease association.
Abbreviations: AD = autosomal dominant; AR = autosomal recessive; DEE = developmental and epileptic encephalopathy; DOOR = deafness, onychodystrophy, osteodystrophy, and mental retardation; IECEE = Infantile or early childhood epileptic encephalopathy; NA = not available; RTA = renal tubular acidosis.

Table A2 (following 6 pages). Clinical and Sequencing Information for each patient.
Detailed clinical information was unavailable for Patients 1 (c.85G>A, p.G29S, *de novo*), 12 (c.294C>A, p.S98R, *de novo*), and 23 (c.448C>T, p.L150F).
[a]Based on reference sequence NM_001694.4.
[b]Classifications using criteria outlined by Richards *et al.* [43]
[c]Previously published as Patient T1911 in Carvill *et al.* [113]
[d]Previously published as DDD4K.04123 in the Deciphering Developmental Disorders Study [111]
[e]Previously published in Ittiwut *et al.* [110]
Abbreviations: BWA = Burrows-Wheeler Aligner, CBD = Cannabidiol, GSW = Generalized spike-wave, GATK = Genome Analysis Toolkit, LVOT = left ventricular outflow tract, MFD = Multifocal discharges, NA = not available

	Patient 2[c]	Patient 3	Patient 4	Patient 5
Variant[a]	c.134_135delCT	c.143G>C	c.158T>G	c.172C>G
Protein Change	p.(S45CfsTer37)	p.R48P	p.M53R	p.P58A
Sex	Male	Female	Female	Male
Seizures?	Yes	Yes	No	Yes
Seizures Controlled?	NA	Yes	NA	NA
Current AED(s)	NA	Levetiracetam	None	NA
Previous AED(s)	NA	NA	None	NA
ID, DD	Severe ID with regression	Walked at 2yr 6mo. Speech between 3-4yrs.	3yr 11mo: no sentences, only 15 words Crawled at 13 months and walked at 2yr	Psychomotor delay
Dysmorphisms	NA	Micrognathia, high palate, low posterior hairline, earlobe creases, prominent nasal bridge, tapering fingers	Long palpebral fissures, long eyelashes, persistent fetal finger pads	NA
EEG	GSW, MFD	NA	14 mo: Normal	NA
MRI	Normal	NA	30 mo: normal	Cerebellar vermis hypoplasia
Other	NA	Hypotonia; Pulmonary valve stenosis; bipolar disorder	Hypotonia; seizure-like episodes from 14mo (freezing and eye-rolling) temporary and self-resolved- Not supported as seizures by EEG.	Dental enamel defects
WES or WGS	WES	Trio WES	Trio WES	NA
Library Capture	SeqCap EX Human Exome Library v2.0	NA	NA	NA
Sequencing	Illumina HiSeq	NA	NA	NA
Alignment	BWA	NA	NA	NA
Variant calling, filtering, and annotation	GATK	NA	NA	NA
ACMG/AMP Classification[b]	Likely pathogenic (PS2, PM2, PM4)	Likely pathogenic (PS2, PM2, PP2, PP3)	Likely pathogenic (PS2, PM2, PP2, PP3)	Likely pathogenic (PM2, PM6, PP2, PP3)

	Patient 6	Patient 7	Patient 8	Patient 9
Variant[a]	c.188G>C	c.214delG	c.220G>T	c.283G>A
	p.G63A			
Protein Change	Male	p.(V72WfsTer9)	p.V74F	p.A95T
Seizures?	Yes	Yes	Yes	Yes
Seizures Controlled?	No	NA	NA	NA
Current AED(s)	Lamotrigine, Levetiracetam	None	Carbmazepine, Zonisimide, Clobazam	Lamotrigine, Topiramate
Previous AED(s)	Vigabatrin, Pyridoxine, ACTH, Topiramate	None	Sodium valproate, Lamotrigine	Phenobarbital, Levetiracetam, Carbamazepine, Clonazepam, Ketogenic diet
ID, DD	Severe ID, non-verbal (at 5.5 yrs, Griffith development scale and Merrill Palmer R show results corresponding to 3-6 months)	Maintain head upright: 4mo, Sitting at 11mo, Babbling starting at 11-12mo	Severe ID (first words at 3-4 yrs, 50 single words at 8yrs) Sat independently at 6-8 months and walked at 13 months	Profound ID, Delayed speech milestones with regression
Dysmorphisms	Microcephaly, small chin	Frontal bossing, low set and floppy ears, hypertelorism, mildly high palate, downslanting palpebral fissures, single transverse palmar crease	No	NA
EEG	At onset: hypsarrhythmia. Primary generalized epileptic abnormality.	Normal	4yr: Bifrontal pattern of low amplitude beta activity	EEG 2002: 4 habitual seizures, ¾ arose in L hemisphere, ¼ arose in R hemisphere EEG 2011- 1 habitual seizure, arose on R hemisphere
MRI	21 mo: Widened liquor spaces, thin corpus callosum and delayed myelination.	<1yr: mild ventricular dilation	Normal	NA
Other	Muscle weakness; scoliosis; feeding difficulties and gastrostomy; cryptorchidism; slightly thickened left ventricular wall; muscle biopsy at 10 mo showed combined complex I and IV defect in Oxidative Phosphorylation pathway	Frequent urinary infections; heart murmur; heterozygous de novo chromosome 20 deletion of 2.3Mb (q11.22q11.23) likely involved in the delayed development and dysmorphism	Loose stool as infant	NA
WES or WGS	WGS	Trio WES	WGS	NA
Library Capture	Illumina TruSeq DNA PCR-free	NA	Illumina TruSeq DNA PCR-free	NA
Sequencing	Illumina HiSeq x Ten	NA	Illumina HiSeq 2500	NA
Alignment	BWA	NA	Isaac	NA
Variant calling, filtering, and annotation	GATK, VEP, Vcfanno, Genmod with MIP 8.2	NA	Platypus	NA
ACMG/AMP Classification[b]	Likely pathogenic (PS2, PM2, PP2, PP3)	Likely pathogenic (PS2, PM2, PM4)	Likely pathogenic (PS2, PM2, PP2, PP3)	Likely pathogenic (PS2, PM2, PP2, PP3)

	Patient 10	Patient 11	Patient 13[d]
Variant[a]	c.283G>C	c.284C>T	c.340_355del16
Protein Change	p.A95P	p.A95V	p.(D115AfsTer12)
Sex	Female	Male	Male
Seizures?	Yes	Yes	Yes
Seizures Controlled?	Has breakthrough seizures	Breakthrough tonic seizures at night	Yes, for 10 years
Current AED(s)	Oxcarbazepine, Topiramate	Oxcarbazepine, Valproic Acid, Levetiracetam	Lamotrigine
Previous AED(s)	Phenobarbital, Levetiracetam	Clobazam, Lamotrigine, Topiramate, Zonisamide, Phenobarbital, Vigabatrin	None
ID, DD	Sat without support at 12 months. At last visit, non-verbal and non-ambulatory (crawls).	Moderate ID, Sat independently at 9 months and walked at 15 months. First words were at 14 months with sentences at 38 months.	Moderate to severe developmental delay. Speaks in short sentences, reads some words and recognizes some shapes.
Dysmorphisms	Round face, widely set eyes, midface hypoplasia, epicanthal folds, somewhat prominent ear lobes	None noted	No
EEG	17hrs at 10 months: Mildly abnormal with mild excessive slowing and probable post-ictal slowing on one occasion. Routine at 10.5 and 13 months: Mildly abnormal due to slowing. ~23hr at 16, 26, and 31 months: normal Prolonged at 37 months: Abnormal due to slowing, 3 electroclinical seizures observed; two consistent with partial seizures.	Background slowing, multifocal epileptiform activity with severe persistence	NA
MRI	10 mo: Normal 34 mo: Bilateral hippocampal sclerosis	Normal	NA
Other	Hypotonia; "Rett-like" episodes of hyperpnea/tachypnea; hand clasping behaviors without apraxia	Hypotonia starting at 15mo with gait ataxia and intentional tremor present; aggressiveness during adulthood that requires medication	Waddles; right side weaker than left with no clear cause; hypertrophic cardiomyopathy with moderate LVOT; mitral valve prolapse; mild-moderate mitral valve regurgitation with moderate mitral stenosis; previous symptoms for cryptorchidism; compound heterozygous for variants in *LZTR1*
WES or WGS	WES	Trio-WES	WES
Library Capture	NA	NA	Agilent SureSelect 55MB Exome Plus
Sequencing	NA	Illumina	Illumina HiSeq
Alignment	NA	BWA	BWA
Variant calling, filtering, and annotation	Sentieon	GATK, VcfAnno. VEP	GATK, SAMtools, Dindel, CoNVex, DeNovoGear, VEP v.2.6
ACMG/AMP Classification[b]	Likely pathogenic (PS2, PM2, PP2, PP3)	Likely pathogenic (PS2, PM2 PP2, PP3)	Likely pathogenic (PS2, PM2, PM4)

	Patient 14	Patient 15	Patient 16	Patient 17
Variant[a]	c. 352_362delinsCATCGGCATCGTGGGGGACGCTGGC	c.395T>A	c.409T>C	c.412G>C
Protein Change	p.(V118HfsTer19)	p.I132N	p.F137L	p.A138P
Sex	Female	Female	Male	Female
Seizures?	Yes	Yes	Yes	Yes
Seizures Controlled?	NA	No	No	NA
Current AED(s)	Clobazam	Levetiracetam, Lamotrigine	CBD, Clobazam, Phenytoin, Levetiracetam	NA
Previous AED(s)	Valproate, Carbamazepine, Levetiracetam, Clonazepam	No	Topiramate, Zonisamide, Depakote, Lorazepam, Carbamazepine, Clonazepam	NA
ID, DD	Mild ID	Crawling at 17mo, walking at 6yr (on toes). Non-verbal, unable to follow simple commands	Profound ID, non-verbal, delayed motor milestones, walks with minimal support	Severe developmental delay
Dysmorphisms	No	Microcephaly, mild synophrys, large mouth, macroglossia, coarsened facial features	Flat occiput, deep set eyes, wide-spaced teeth, full lips, two branchial arch pits on the right side of neck	NA
EEG	NA	NA	Slowing and disorganization of the background, frequent multifocal sharps, focal epileptiform abnormalities	NA
MRI	Small white matter changes in left hemisphere	12 yr: normal	9yr: diffuse supratentorial and infratentorial volume loss	NA
Other	Strange odor, especially before or after GTCS; increased amino acids (glycine, alanine and proline) in plasma.	Noisy breathing; mild ataxia	Broad-based ataxic gait; hypotonia; Ketogenic diet failed; Type I Diabetes diagnosed at 17yrs	Neonatal apneic episodes
WES or WGS	WES	Trio Autism Xpanded Panel	WES	WES
Library Capture	Agilent SureSelect Clinical Research Exome V2	NA	Agilent SureSelect XT Clinical Research Exome Kit	Agilent SureSelect Clinical Research Exome V1
Sequencing	Illumina NovaSeq 6000	NA	Illuminia HiSeq 2000 or 2500	Illumina HiSeq
Alignment	BWA	NA	NA	PEMapper
Variant calling, filtering, and annotation	GATK, Alissa Interpret Clinical Informatics Platform	NA	In-house, custom-built bioinformatics pipeline v.CWES-2.2	PECaller, Bystro
ACMG/AMP Classification[b]	Likely pathogenic (PS2, PM2, PM4)	Pathogenic (PS2, PM1, PM2, PP2, PP3)	Pathogenic (PS2, PM1, PM2, PP2, PP3)	Pathogenic (PS2, PM1, PM2, PP2, PP3)

	Patient 18	Patient 19	Patient 20	Patient 21
Variant[a]	c.412G>C	c.425G>A	c.440T>C	c.445G>A
Protein Change	p.A138P	p.G142D	p.I147T	p.A149T
Sex	Male	Female	Male	Male
Seizures?	Yes	No	Yes	Yes
Seizures Controlled?	NA	NA	No	Yes, for 1 year
Current AED(s)	Topiramate, Clonazepam, Levetiracetam	None	Clobazam, Rufinamide, CBD	None
Previous AED(s)	None	None	Topiramate, Zonisamide, Sodium valproate, Levetiracetam, Lamotrigine	Valproate
ID, DD	Delayed motor and speech milestones	Rolling at 6mo, siting at 14mo. 26mo: non-verbal, no self-feeding, walks with orthotics and assistance	Profound ID (no speech)	Mild ID (IQ 71)
Dysmorphisms	Microcephaly	Microcephaly, periorbital fullness, shortened nasal bridge, triangular nasal tip, retrognathia, long first toes, cortical thumbs	Deep set eyes, prominent brows, deviated nose from breaking. (Many features also seen in unaffected father).	Microcephaly
EEG	NA	NA	14yr: Abnormal due to moderate slow diffuse slowing of background rhythms, rare bifrontal and generalized low amplitude spikes in sleep, and cluster of brief generalized tonic seizures.	Normal EEG background
MRI	7 mo: delayed myelination, hypoplastic corpus callosum, frontal atrophy	10 mo: Pineal region cyst	4.7yr: Diffuse cerebral atrophy with diffuse T2 hyperintense signal involving white matter bilaterally. Bilateral hippocampal sclerosis.	Normal
Other	Hypotonia; dystonic posturing of upper limbs; poor eye contact	Hypotonia	Frequent falling	viral cerebellitis with transient ataxia at 3y; hypernasal speech
WES or WGS	WES	Trio WES	Trio WES	Trio WES
Library Capture	Twist Human Core Exome Kit	NA	Agilent SureSelect Exon V6 plus custom mitochondrial genome capture kit	SeqCap EZ Exome v3; Roche
Sequencing	Illumina NovaSeq 6000	NA	NA	Illumina NovaSeq 6000
Alignment	BWA	NA	NA	BWA
Variant calling, filtering, and annotation	GATK and VarSeq	NA	Alissa	GATK, Alissa Interpret Clinical Informatics Platform
ACMG/AMP Classification[b]	Pathogenic (PS2, PM1, PM2, PP2, PP3)	Pathogenic (PS2, PM1, PM2, PP2, PP3)	Pathogenic (PS2, PM1, PM2, PP2, PP3)	Pathogenic (PS2, PM1, PM2, PP2, PP3)

	Patient 22	Patient 24	Patient 25	Patient 26[e]
Variant[a]	c.445G>A	c.448C>T	c.449T>C	c.467A>T
Protein Change	p.A149T	p.L150F	p.L150P	p.(Ter156LextTer35)
Sex	Male	Male	Female	Female
Seizures?	Yes	Yes	Yes	Yes
Seizures Controlled?	Yes, since 14.5yrs old	3-4yrs with ketogenic diet and topiramate	Yes	No
Current AED(s)	Levetiracetam, Lamotrigine	Topiramate	Lamotrigine, Brivaracetam, Clobazam, CBD	Topiramate, Clonazepam, Levetiracetam, Lamotrigine
Previous AED(s)	Levetiracetam with Topiramate	Sodium valproate, Carbamazepine, Prednisolone, Vigabatrin, Lamotrigine, Clobazam, Gabapentin, Levetiracetam, Rufinamide, Phenytoin, Ketogenic diet, Perampanel	Levetiracetam, Sodium valproate, Topiramate, Rufinamide	NA
ID, DD	Mild Spoke 25 words at 3yrs	Profound ID (no words or sentences). Sat independently at 11 months, but not walking	Mild ID and fine motor delay	ID with regression
Dysmorphisms	Ears a bit low set, beaked nose (shared with mother)	Low columella, beaked nose, downturned mouth, plagiocephaly	Beaked nose and low set ears	NA
EEG	From 6mo to 7yrs, several done. Many normal, four showing spikes, sharp waves (2 Hz), bifronto-temperal-centralt. From 7yrs to current: EEGs have all been normal	Early EEG reported to be normal. Developed left sided epileptiform activity and a left anterior temporal onset to clinical seizures. Interictal EEG shows excess of slow for age.	2-2.5 Hz slow-spike wave discharges, paroxysmal generalized fast activity	7yrs: generalized spike-wave complexes 16yrs: frequent spike-slow waves originating in left frontal region
MRI	4yr: delayed myelination	Delayed myelination	Corpus callosum agenesis/dysplasia	7, 9, and 16 yrs: Unchanged slightly smaller size of left hippocampus
Other	Valgus deformities of the feet	Weaned off ketogenic diet after recurrence of seizures and development of Staghorn calculus and renal calculi thought to be secondary to ketogenic diet; poor weight gain/oral intake; mild scoliosis; no enamel on teeth; soft brittle nails	Lennox-Gastaut Syndrome; s/p laser corpus callostomy; vagal nerve stimulation; frequent drop attacks	NA
WES or WGS	Trio WES	WGS	NA	WES
Library Capture	Twist Human Core Exome Multiplex Enrichment Kit	Illumina TruSeq DNA PCR-free	NA	NA
Sequencing	Illumina NovaSeq 6000	Illumina HiSeq 2500	NA	NA
Alignment	BWA	Isaac	NA	NA
Variant calling, filtering, and annotation	Freebayes	Platypus	NA	NA
ACMG/AMP Classification[b]	Pathogenic (PS2, PM1, PM2, PP2, PP3)	Pathogenic (PS2, PM1, PM2, PM5 PP2, PP3)	Pathogenic (PS2, PM1, PM2, PM5 PP2, PP3)	Likely pathogenic (PS2, PM2, PM4)

Strain/Plasmid	Description	Human Allele	Source
vma3Δ (YEL027W)	*MATa leu2Δ ura3Δ his3Δ TRP1 met15Δ vma3Δ::kanMX*	-	Dharmacon
pRS316	*CEN6, URA3, amp*^{R}	-	Sikorski and Heiter [122]
pKM16	*VMA3* in pRS316, *CEN6, URA3, amp*^{R}	Wild type	This Study
pKM17	*vma3-R46P* in pRS316, *CEN6, URA3, amp*^{R}	p.R48W	This Study
pKM18	*vma3-G61A* in pRS316, *CEN6, URA3, amp*^{R}	p.G63A	This Study
pKM19	*vma3-A93T* in pRS316, *CEN6, URA3, amp*^{R}	p.A95T	This Study
pKM20	*vma3-F135L* in pRS316, *CEN6, URA3, amp*^{R}	p.F137L	This Study
pKM21	*vma3-A136P* in pRS316, *CEN6, URA3, amp*^{R}	p.A138P	This Study
pKM22	*vma3-E137A* in pRS316, *CEN6, URA3, amp*^{R}	p.E139A	This Study
pKM23	*vma3-L148F* in pRS316, *CEN6, URA3, amp*^{R}	p.L150F	This Study
pKM24	*vma3-L148P* in pRS316, *CEN6, URA3, amp*^{R}	p.L150P	This Study
pKM27	*vma3-R46W* in pRS316, *CEN6, URA3, amp*^{R}	p.R48W	This Study
pKM28	*vma3-G101S* in pRS316, *CEN6, URA3, amp*^{R}	p.G103S	This Study
pKM29	*vma3-M129I* in pRS316, *CEN6, URA3, amp*^{R}	p.M131I	This Study
pKM30	*vma3-V72F* in pRS316, *CEN6, URA3, amp*^{R}	p.V74F	This Study
pKM31	*vma3-A93P* in pRS316, *CEN6, URA3, amp*^{R}	p.A95P	This Study
pKM32	*vma3-S96R* in pRS316, *CEN6, URA3, amp*^{R}	p.S98R	This Study
pKM39	*vma3-G27S* in pRS316, *CEN6, URA3, amp*^{R}	p.G29S	This Study
pKM40	*vma3-A147T* in pRS316, *CEN6, URA3, amp*^{R}	p.A149T	This Study

Table A3 – Yeast Strains and Plasmids

Strain	Human Allele	*C. elegans* Gene	*C. elegans* Variant	*Source*
AP08	p.G63A	*vha-3*	p.G69A	This Study
AP09	p.L150F	*vha-3*	p.L156F	This Study
AP10	p.A95T	*vha-2*	p.A101T	This Study
AP11	p.F137L	*vha-2*	p.F143L	This Study

Table A4. *C. elegans* Strains and Variants
Note: The *C. elegans* A101T mutation caused sterility in homozygous *C. elegans* worms and was not used in assays.

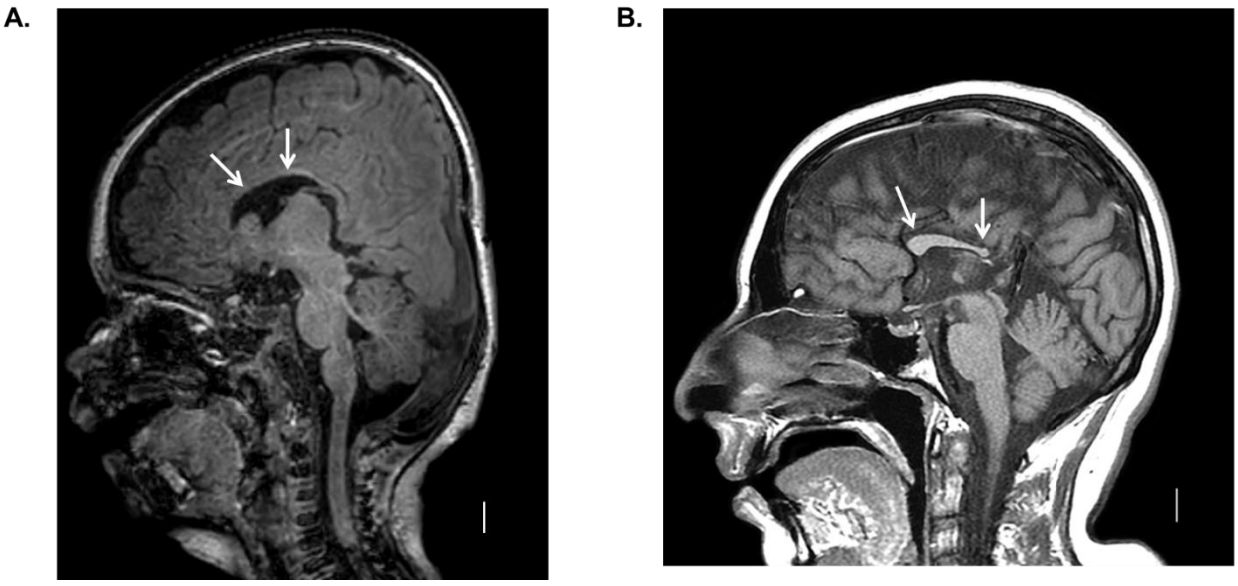


Figure A1. Magnetic Resonance Imaging of Patients.
A) Sagittal MRI from patient 18 demonstrating hypoplasia of the corpus callosum (arrows).
B) Sagittal MRI from patient 25 demonstrating moderately short corpus callosum with absent rostrum, tapered body and very small splenium (arrows).
Horizontal white bars in lower right indicate 1cm.

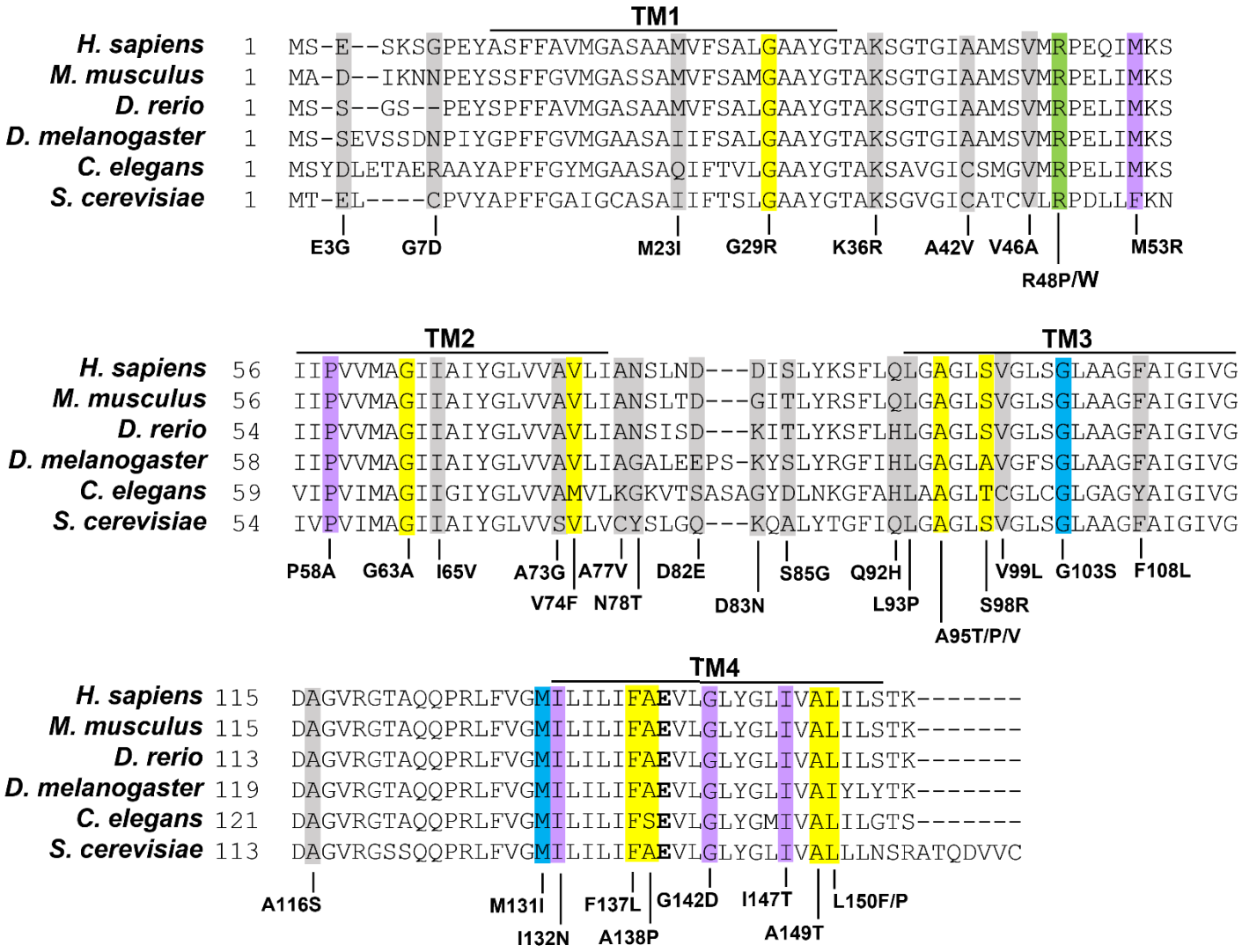


Figure A2. Alignment of human ATP6V0C with selected eukaryotic organisms. Residues affected by missense changes in patients are highlighted in yellow (functionally tested) or pink (not tested). Residues highlighted in blue (functionally tested) and grey (not tested) represent variants from gnomAD (v.2.1.1). R48 (in green) is the location of a patient and gnomAD variant, p.R48P (tested) and p.R48W (tested), respectively. The bolded "E" represents the glutamate residue required for proton transport. Transmembrane regions are denoted by "TM". The following protein sequences were used in the alignments: *H. sapiens*, NP_001685.1; *M. musculus*, NP_001348461.1; *D. rerio*, NP_991117.7; *D. melanogaster*, NP_476801.1; *C. elegans,* NP_499166.1; *S. cerevisiae,* NP_010887.3.

Appendix B

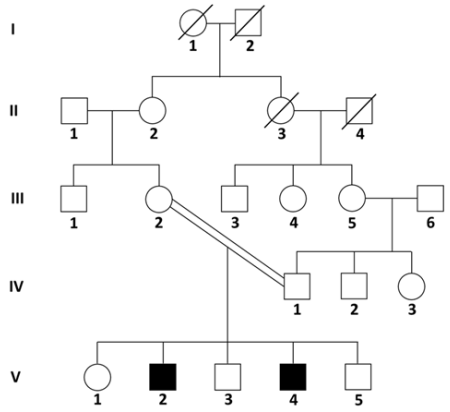
Results from Whole Genome Sequencing in Eight Consanguineous Families with Epilepsy

B.1 Summary

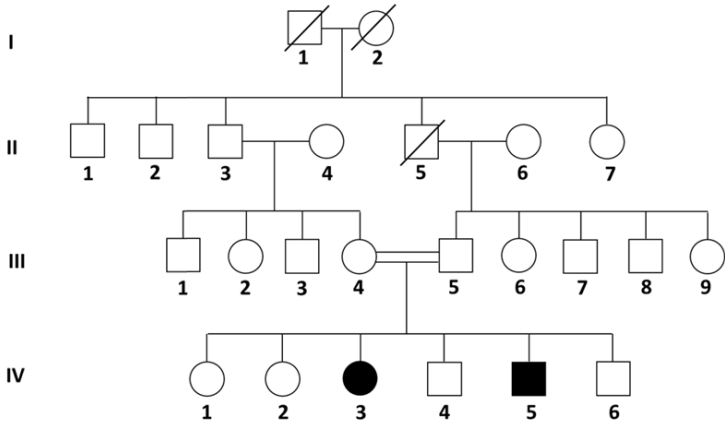
In 2019 we obtained blood samples from 225 individuals from 25 consanguineous families with a history of epilepsy and other neurodevelopmental disorders (provided by visiting PhD student, Noor Badshah). The families belong to the Afridi tribe in Pakistan, and our study was the first to look at the genetic basis of epilepsy within this population. Given the consanguineous nature of the families we hypothesized that the causative variant in each family would be recessive. This is in contrast to the analysis done in **Chapters 2** and **3** where we were looking for *de novo,* heterozygous variants in patients that had already been sequenced.

Two full siblings with epilepsy in ten families were chosen for WGS. After which shared regions of homozygosity were identified and screened for variants. During quality control of the sequence data, two families were excluded from further study. In one family it was determined that the two siblings sequenced were monozygotic twins and in a second family it was determined that it was a nephew-uncle pair that was sequenced instead of full siblings. The identification and analysis of variants in the eight remaining families (**Figure B1)** is described here.

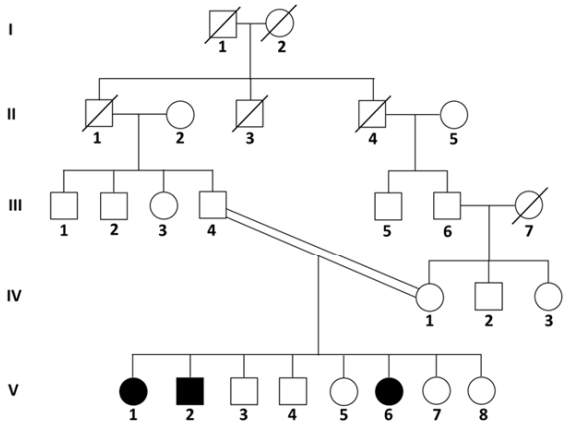
Family 201



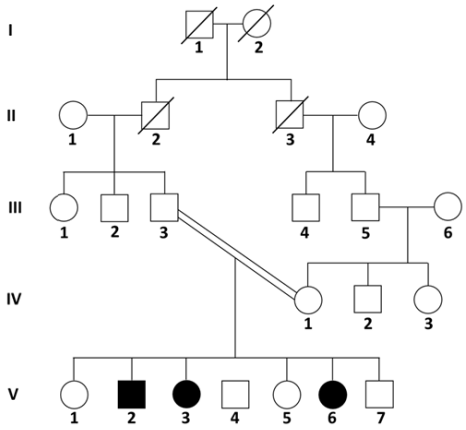
Family 204



Family 205



Family 210



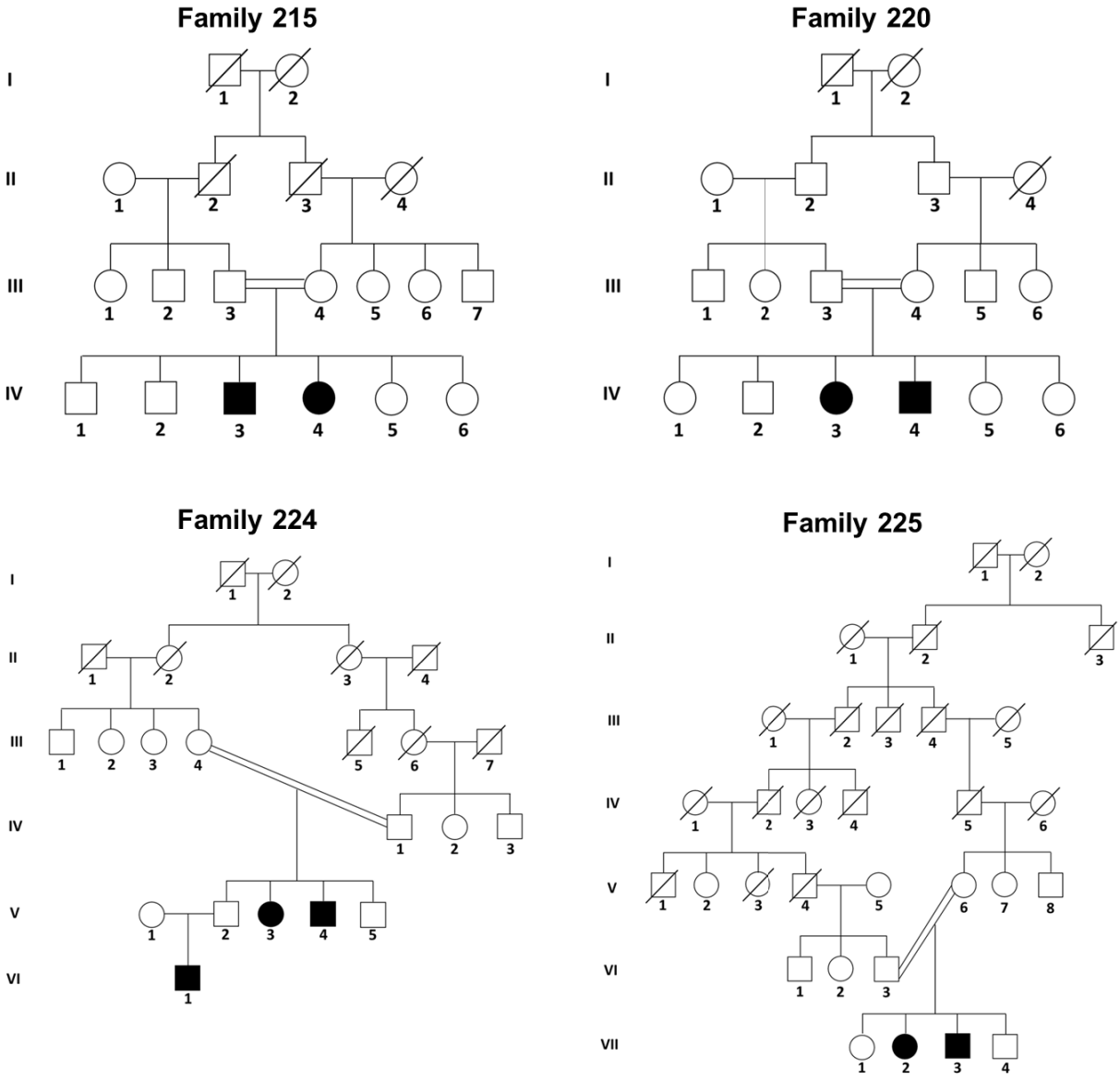


Figure B1. Pedigrees of Eight Pakistani Families with Epilepsy. Two affected (filled symbols) were screened via WGS and shared homozygous variants were identified. Segregation analysis of top variants was performed using gDNA from all available family members.

B.2 Methods

Subjects, Sample Preparation and Sequencing

Clinical information and sample collection, preparation, and sequencing was performed as described for Family 1 in **Chapter 4**. This study was approved by the Ethical Committee at the Institute of Biotechnology and Genetic Engineering, the University of Agriculture, Peshawar, Pakistan.

Variant Annotation and Segregation Analysis

Variants were annotated using Bystro (bystro.io) [162]. Filtering of shared homozygous variants between the sequenced siblings in each family was done using the following command: “CADD > 15, MAF < 0.001, exonic, splice”. MAF values were taken from gnomAD. **Table B1** provides an overview of shared homozygous variants identified in each sibling pair that underwent WGS.

Segregation of all variants that passed the filtering criteria was conducted for each family by sequencing other family members to determine variant segregation within a family. Primers used for PCR and sequencing are provided in **Table B2**.

Population Databases

Three population databases were used to look at variant frequency in a healthy, control population. The count of individuals with an identified variant in a homozygous state was manually curated. In order to ensure that variant counts were not biased by the ancestry of the population dataset, we used three different databases.

gnomAD

Consists of 125,748 exomes and 15,708 genomes. Approximately 65% of the individuals in this database are of European ancestry, and 18% are of Asian ancestry. Version 2.1.1 was used for all analysis (gnomAD.broadinstitute.org) [116].

GMEVariome

The Greater Middle East (GME) Variome consists of 2,497 unrelated individuals. This databased contains individuals of Middle Eastern ancestry, including Pakistan. Data was accessed through: igm.ucsd.edu/gme/index.php [163].

GenomeAsia 100K

GenomeAsia currently consists of 1,739 genomes from across 64 countries in Asia. 598 samples are from India, and 113 samples are from Pakistan (browser.genomeasia100k.org/) [164].

Family	201	204	205	210	215	220	224	225
RoH (Mb)	103.9	87.4	115.4	129.7	114.7	142.1	59.8	31.7
Hom Variants	~1.3 million/individual							~1.25 million/individual
RoH Variants	79,239	81,357	83,654	62,898	58,915	87,334	3,272	14,472
MAF ≤ 0.001	615	449	445	388	254	421	5	66
CADD ≥ 15	21	14	16	15	4	21	0	2
Exonic or Splice	4	4	8	7	0	12	0	1

Table B1. Overview of Homozygous and Shared Homozygous Variants in Sequenced Individuals from Each Family.
RoH (Mb): Total size of shared regions of homozygosity between the two sequenced individuals
Hom Variants: Approximate number of homozygous variants in each individual
RoH Variants: Variants that fall within the shared homozygous regions between the two sequenced individuals
MAF ≤ 0.001: Removed any variants with a minor allele frequency (MAF) > 0.001 based on data from gnomAD
CADD ≥ 15: Removed any variants that have a combined annotation dependent depletion (CADD) score < 15
Exonic or Splice Site: Variants not within exons or splice acceptor or donor sites were excluded

Family	Gene	Forward Primer (5' - 3')	Reverse Primer (5' - 3')
204	*CCDC158*	TTTCATAGACGAAGGGAGAG	CTGAGGCTTCTTCAAAGTCA
	GNB5	CCGGCATAAGCAATAACC	GCTCTCAGATGCTAACTCTC
	LGALS3BP	GTCCTAGCTGCTCTTGTTC	CAACTGTCCAACACAGCC
	SPATA2	GCTGCTCAGAAACGATGC	GGATTTGGAGTGAAGGTGC
205	*EFCAB6*	AGTCTGAAACATCAGCAAAG	CAGAGCCTAATTCATAGAGT
	FAM174B	CAGTTGCTGCTTCACCAC	GAGGGTGGTAAAGGCGAA
	FAM227A	ATCAAAGGACATGAGTGC	GGAACTAAGATGAGCACGA
	FANCC	TGTCAACTGCCATGTGTTC	CATAGTCTGTGCTCTCTGCTG
	ST8SIA2	TCCAGCCCTAATGGAGTG	GCTTATTCTTCAGTGGCG
	TMC3	GAGAGGTCCGAAAGTCAG	CCAGAAAGGGAGATGCCA
	TRIM68	TAGAGCATGATTGGAGTGG	GGATGTAGCCTTAGCAGAC
	ZNF530	GCCAGGGATGGATGTTTG	CGATGTCCCTCAGAATACA
210	*ACTL7A*	CATACGAGTTCAGAGCCAC	CTTCAAACAGCATTTCAGCA
	CHRD	GTAGGTGGGATGGACTAAG	GACTCAGAGACAGCAAGTG
	CRYM	CAACTCGTAGGGTCAATGTC	CTGAGGACTGGAAACAGGTA
	RNF20	TAGTGGTAGTGCCCTCCTG	CATGCCATCAACCAAACA
	RNPEPL1	GGCAAGTGAAGGCTAAGG	CTTGACGACGGCAGAGTA
	RSRC2	TTAGTTACTTGGAAAGCAGG	CATGCCCAGCCAACAAAC
	TNXB	AGCTAACAGTGACAGGATC	GAACAGGGTGTATTACAGGT
220	*ADAMTS19*	GTGACATTATAGGCCAGC	GAGTAGAGTTTGAGGAGCT
	CCDC141	GGTGGATATGCAGAAGATAATG	CTTCAGAAATGCCACGTAA
	CLIP2	ACAGAGTTTGAATGCAGGTC	CTGGCCTCCATACTCTCA
	FZD9	CATCTTCCTCTCCATGTGC	GACCATGAGCTTCTCCAGC
	GLI3	GTCCATATCACCACTCTCC	GCTATTGACTAATGTCTGCA
	HMGXB3	ATTGGCACCTTACGCAAC	GATGGTGTAGTCCCTGTAC
	JAG1	TCTTGGAAGTTTCGCTAGG	GAAACCAGTGGGACACAGA
	SEL12	AAAGTGAAGGTGTTCCAGTG	GGAACTGTAAGTAGCAGGGA
	SEMA3A	AACCATTGAGGCCATGTG	GTAAGGCTTCCAGTCAAC
	SLC36A2	CTGGACCTGGTCATCTCCCT	TGGTGGAGTTGGAAAAGGGG
	SUN3	CTGGAAGAACAACCACACTAG	CTAGCCAACCACCTTCTATC
	UBOX5	TATCTAACTCCTGGGCTAATG	GACTGGTGGCTTAATGAAAT
225	*KIF18A*	AAGTGGTTCATGTTGTGG	GATGTCCGTCCTAATGCA

Table B2. Primers used for PCR and Sequencing.

B.3 Results and Discussion

WGS on two affected individuals with epilepsy was performed in eight consanguineous families. Approximately 1.3 million homozygous variants were seen in each individual sequenced. The average size of the total regions of homozygosity in each family was 98.1 ± 36.9 Mb and 58,892 ± 32,565 variants were found within these regions. Filtering on MAF decreased the number of variants by over 98% in each family. Further filtering using CADD ≥ 15, representing the top 10% of predicted deleterious variants, and selection for variants that fell within exonic sequences or canonical splice sites reduced the average number of variants within each family to 4.5 ± 4.3 (**Table B1**). After filtering, two families, 215 and 224, had zero variants identified for further segregation analysis.

For the eight remaining families, segregation analysis of variants that passed filtering was performed by sequencing additional family members who had provided gDNA. The variants and results of segregation analysis in these families is discussed below.

B.3.1 Family 201

Family 201 had four variants identified for further study. The segregation analysis and follow-up of these variants is provided in detail in **Chapter 4**.

B.3.2 Family 204

Four variants were identified after filtering (**Table B3**). Segregation analysis showed that only one variant, p.Gly387Arg in *SPATA2,* segregated in the expected pattern based on the assumption of AR inheritance. *SPATA2* encodes spermatogenesis associated 2 protein which is required to recruit CYLD to complexes involved in tumor necrosis factor alpha (TNF-α) signaling. Knockdown of *SPATA2* inhibits TNF-α mediated necroptosis and apoptosis [216, 217]. SPATA2 is expressed across all major structures in the brain as well as in the testes at

similar levels [218]. TNF-α is a pro-inflammatory cytokine and increased TNF-α signaling after infection has been linked to hyperexcitability and acute seizures [219]. Therefore, the p.Gly387Arg variant in *SPATA2* might be responsible for the epilepsy in this family, but functional testing will be necessary to determine if this variant has a loss or gain of function effect on TNF-α signaling, and whether the impact on TNF-α signaling might lead to epileptogenesis.

Gene		*CCDC158*	*GNB5*	*LGALS3BP*	*SPATA2*
Chromosome		4	15	17	20
HGVS DNA		NM_001042784.1; c.527G>A	NM_006578.4; c.592G>A	NM_005567.4; c.558G>A	NM_001135773.2; c.1159G>A
HGVS Protein		p.Arg176Gln	p.Val198Ile	p.Trp186Ter	p.Gly387Arg
CADD Score		24	29.5	35	25.5
Population Frequency	**gnomAD**	0	0	0	1
	GMEVariome	0	0	0	0
	GenomeAsia 100K	0	0	0	0
Affected	**IV:3**	Hom	Hom	Hom	Hom
	IV:5	Hom	Hom	Hom	Hom
Unaffected	**III:7**	WT	Het	WT	WT
	IV:2	WT	WT	Hom	Het
	IV:6	Hom	Hom	WT	WT
ACMG Evidence		PM2, PP3, BS2	PM2, PP3, BS2	PM2, PP3, BS2	PM2, PP3
ACMG Classification		VUS	VUS	VUS	VUS

Table B3. Segregation Analysis in Family 204.

Blue shading indicates zygosity in an individual reflects AR inheritance while red indicates zygosity in an individual violates AR inheritance (assuming full penetrance).

B.3.3 Family 205

Eight variants were identified in this family; however, none of them segregated in the expected AR pattern and five would be classified as benign using the ACMG/AMP classifications (**Table B4**). Interestingly the variant in *FAM174B* was called as a 4bp deletion via WGS which would result in premature truncation of the protein. The other family members sequenced over this region had a 6bp deletion that would result in the deletion of two amino acids (p.Ser69_Ser70del). This in-frame deletion is very common in the healthy population being seen in a homozygous state 89,483 times in gnomAD, 699 times in GMEVariome, and 1689 times in the GenomeAsia100K dataset. Sanger sequencing was not done to confirm the 4bp deletion in the two affected siblings; however, given that all other family members are homozygous for the common 6bp deletion it is likely that the variant was incorrectly called from the WGS data and likely does not contribute to epilepsy.

Table B4 (next page). Segregation Analysis in Family 205.
Blue shading indicates zygosity in an individual reflects AR inheritance while red indicates zygosity in an individual violates AR inheritance (assuming full penetrance).
[a]This classification is based on the assumption that affected individuals V:1 and V:2 are homozygous for the 6bp in-frame deletion in *FAM174B.*

Gene		*EFCAB6*	*FAM174B*	*FAM227A*	*FANCC*
Chromosome		22	15	22	9
HGVS DNA		NM_022785.4; c.71G>A	NM_027446.3; c.206_209delGCTC	NM_001013647.2; c.1223C>T	NM_000136.3; c.1264C>A
HGVS Protein		p.Arg24Lys	p.(Ser69ThrfsTer8)	p.Ser408Leu	p.Leu422Met
CADD Score		26.6	23.3	27.2	22.7
Population Frequency	gnomAD	0	0	0	0
	GMEVariome	0	0	0	0
	GenomeAsia 100K	0	0	0	0
Affected	V:1	Hom	Hom	Hom	Hom
	V:2	Hom	Hom	Hom	Hom
	V:6	Het	Δ6bp/Δ6bp	Hom	Hom
Unaffected	V:3	Hom	Δ6bp/Δ6bp	Hom	Hom
	V:4	Hom	Δ6bp/Δ6bp	Hom	Hom
	V:5	Hom	Δ6bp/Δ6bp	Hom	Hom
	V:7	Hom	Δ6bp/Δ6bp	Hom	Hom
	V:8	Hom	Δ6bp/Δ6bp	Hom	Hom
ACMG Evidence		PM2, PP3, BS2, BS4	PM2, PP3, BA1, BS1, BS2	PM2, PP3, BS2	PM2, PP3, BS2
ACMG Classification		Benign	Benign[a]	VUS	VUS

Gene		*ST8SIA2*	*TMC3*	*TRIM68*	*ZNF530*
Chromosome		15	15	11	19
HGVS DNA		NM006011.4; c.382G>A	NM_001080532.3; c.3127T>G	NM_018073.8; c.118C>T	NM_001321981.2; c.47delT
HGVS Protein		p.Asp128Asn	p.Ser1043Ala	p.Leu40Phe	p.(Met17TrpfsTer3)
CADD Score		26.4	24.2	24	18.6
Population Frequency	gnomAD	0	0	0	0
	GMEVariome	0	0	0	0
	GenomeAsia 100K	0	0	0	0
Affected	V:1	Hom	Hom	Hom	Hom
	V:2	Hom	Hom	Hom	Hom
	V:6	WT	WT	Hom	Het
Unaffected	V:3	Het	Hom	Hom	Het
	V:4	WT	WT	Hom	Het
	V:5	Hom	Hom	Hom	Het
	V:7	Het	WT	Hom	Hom
	V:8	Het	Het	Hom	Hom
ACMG Evidence		PM2, PP3, BS2, BS4	PM2, PP3, BS2, BS4	PM2, PP3, BS2	PM2, PP3, BS2, BS4
ACMG Classification		Benign	Benign	VUS	Benign

B.3.4 Family 210

After filtering, seven variants were identified in this family. Of these seven variants, only p.Thr633Ile in *CHRD* segregated in an AR inheritance pattern (**Table B5**). *CHRD* encodes Chordin, the vertebrate ortholog of short gastrulation (SOG), which is required in the dorsalizing of early embryonic tissue [220]. During development *CHRD* is primarily expressed in the anterior primitive streak, node, and axial mesendoderm [221]. Homozygous knockout in mice results in stillborn pups with signs of normal early development [221]. While *CHRD* is expressed throughout the brain, it remains unclear if the p.Thr633Ile variant would be pathogenic and impact neural development in a manner that would lead to epilepsy without major impacts on other vital organs.

Table B5 (next page). Segregation Analysis in Family 210.
Blue shading indicates zygosity in an individual reflects AR inheritance while red indicates zygosity in an individual violates AR inheritance (assuming full penetrance).

Gene		*ACTL7A*	*CHRD*	*CRYM*	*RNF20*
Chromosome		9	3	16	9
HGVS DNA		NM_018226.6; c.564C>T	NM_001304472.1; c.1898C>T	N<_001376256.1; c.421A>G	NM_019592.7; c.1703A>G
HGVS Protein		p.Arg136Leu	p.Thr633Ile	p.Ile141Val	p.Glu568Gly
CADD Score		23.1	26.6	24	22.1
Population Frequency	gnomAD	0	1	0	0
	GMEVariome	0	0	0	0
	GenomeAsia 100K	0	0	0	0
Affected	V:2	Hom	Hom	Hom	Hom
	V:3	Hom	Hom	Het	Het
	V:6	Hom	Hom	Hom	Hom
Unaffected	V:4	Hom	WT	Hom	Het
	V:5	Het	Het	Het	Het
ACMG Evidence		PM2, PP3, BS2, BS4	PM2, PP3	PM2, PP3, BS2, BS4	PM2, PP3, BS4
ACMG Classification		Benign	VUS	Benign	VUS

Gene		*RNPEPL1*	*RSRC2*	*TNXB*
Chromosome		2	12	6
HGVS DNA		NM_018226.6; c.564C>T	NM_023012.6; c.138G>A	NM_001304472.1; c.6542C>T
HGVS Protein		p.Cys188Cys	p.Arg46Arg	p.Thr2181Met
CADD Score		17.1	17.7	24.5
Population Frequency	gnomAD	0	1	0
	GMEVariome	0	0	0
	GenomeAsia 100K	0	0	0
Affected	V:2	Hom	Hom	Hom
	V:3	Het	WT	Het
	V:6	Hom	Hom	Hom
Unaffected	V:4	Het	Het	Het
	V:5	Het	Het	Het
ACMG Evidence		PM2, BS4, BP7	PM2, BS4, BP7	PM2, PP3, BS4
ACMG Classification		L. Benign	L. Benign	VUS

B.3.5 Family 220

Twelve variants were identified in Family 220 and five remained after segregation analysis (**Table B6**). Of the five variants that segregated in an AR inheritance pattern, p.Leu543Val in *CCDC141* is the top candidate. *CCDC141* plays an important role in centrosome positioning and movement during neuronal migration in the developing nervous system [222]. Knockdown of *Ccdc141* in mice results in abnormal cortical migration and accumulation of neurons within the intermediate zone and a decrease in neuronal numbers at that cortical plate [222]. Aberrant neuronal migration has previously been linked to epilepsy and is seen with pathogenic variants in *DYNC1H1* and *DCX* [223, 224]. Heterozygous variants in *CCDC141* have been linked to congenital hypogonadotropic hypogonadism (CHH) in recent years which has no neurologic phenotype. However, nine of the twelve patients reported also had variants in other established CHH genes suggesting that *CCDC141* may only have a minor role, if any, in contributing to the CHH phenotype [225].

Three other variants that segregated in an AR pattern are in, *ADAMTS19, HMGXB3,* and *UBOX5*, which do not have any known pathogenic variants linked to disease, but they also do not have well characterized functional roles. Without more information on the biological pathways these genes are involved in it is difficult to determine whether they may contribute to epilepsy.

The final variant, p.Ala445Thr, in *SLC36A2* is unlikely to contribute to epilepsy as it is almost exclusively expressed in the kidney and has been previously associated with AR iminoglycinuria or hyperglycinuria [226].

Gene		*ADAMTS19*	*CCDC141*	*CLIP2*	*FZD9*	*GLI3*	*HMGXB3*	*JAG1*	*SEL12*	*SEMA3A*	*SLC36A2*	*SUN3*	*UBOX5*
Chromosome		5	2	7	7	7	5	20	20	7	5	7	20
HGVS DNA		NM_133638.6; c.1899C>A	NM_173648.4; c.1627C>G	NM_003388.5; c.685G>A	NM_003508.3; c.920G>A	NM_000168.6; c.1021G>A	NM_014983.3; c.3767A>G	NM_0002144.3; c.1439C>T	NM_001271539.2; c.900+1 G>C	NM_006080.3; c.1263A>G	NM_181776.3; c.1333G>A	NM_001271539.2; c.861+1 G>A	NM_014948.4; c.97G>A
HGVS Protein		p.Ser633Arg	p.Leu543Val	p.Gly229Arg	p.Gly307Asp	p.Ala341Thr	p.Gln1256Arg	p.Ala480Val	NA	p.Ile421Met	p.Ala445Thr	NA	p.Asp33Asn
CADD Score		18.8	23.9	32	25.5	25.3	27.2	22.8	25.2	24.6	23.2	25.3	23.3
Population Frequency	**gnomAD**	2	1	0	0	0	0	7	0	0	0	1	1
	GMEVariome	0	0	0	0	0	0	0	0	0	0	0	0
	GenomeAsia 100K	0	0	0	0	0	0	1	0	0	0	0	0
Affected	**IV:3**	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom
	IV:4	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom
Unaffected	**IV:1**	WT	Het	Hom	Hom	Hom	WT	Het	Het	Hom	WT	Hom	WT
	IV:2	Het	Het	Hom	Unk[a]	Hom	Het	Het	Het	Het	Het	Hom	Het
	IV:5	Het	Het	Hom	Hom	Hom	Het	Hom	Hom	Hom	Het	Hom	Het
ACMG Evidence		PM2, PP3	PM2, PP3	PM2, PP3, BS2	PM2, PP3, BS2	PM2, PP3, BS2	PM2, PP3	PM2, PP3, BS2	PM2, PP3, BS2	PM2, PP3, BS2	PM2, PP3	PM2, PP3, BS2	PM2, PP3
ACMG Classification		VUS	VUS	VUS	VUS	VUS	VUS	VUS	VUS	VUS	VUS	VUS	VUS

Table B6. Segregation Analysis in Family 220.

Blue shading indicates zygosity in an individual reflects AR inheritance while red indicates zygosity in an individual violates AR inheritance (assuming full penetrance).

[a]Sequencing of IV:2 for the *FZD9* variant was inconclusive.

B.3.6 Family 225

Only one variant, p.Thr83Thr in *KIF18A* passed initial filtering in Family 225. This variant did segregate in an AR inheritance pattern (**Table B7**). *KIF18A* is part of the kinesin family which are microtubular-associated molecular motors. KIF18A has been shown to participate in the positioning of mitotic chromosomes and loss leads to mitotic arrest and cell lethality [227, 228]. Despite this function it is unlikely this variant is pathogenic as *KIF18A* has relatively low levels of expression in the brain and given the fact this variant is synonymous and does not alter the protein sequence [218]. While synonymous variants can be disease causing, no functions of KIF18A that involve direct binding of DNA, which are the most likely to be impacted by variants that only affect the DNA sequence, are known.

Gene		*KIF18A*
Chromosome		11
HGVS DNA		NM_031217.4; c.249G>A
HGVS Protein		p.Thr83Thr
CADD Score		16.8
Population Frequency	gnomAD	0
	GMEVariome	0
	GenomeAsia 100K	0
Affected	VII:2	Hom
	VII:3	Hom
Unaffected	VII:1	WT
	VII:4	Het

Table B7. Segregation Analysis in Family 225.
Blue shading indicates zygosity in an individual reflects AR inheritance while red indicates zygosity in an individual violates AR inheritance (assuming full penetrance).

B.4 Conclusions and Future Directions

Overall, we sequenced 10 consanguineous families with epilepsy of Afridi tribal heritage. Two families were excluded during initial quality control of WGS data, and two families did not have any homozygous variants that were shared between the two affected siblings. Of the six remaining families, only Family 201 had a homozygous variant identified in a previously associated epilepsy gene (**Chapter 4**).

For Families 215 and 224, it would be worthwhile to reevaluate the sequencing data. In our initial analysis of variants, we focused on rare, homozygous variants within shared regions of homozygosity. In doing so we ignored any heterozygous variants that may contribute to biallelic (or compound heterozygous) inheritance. Therefore, analysis of shared heterozygous variants between affected individuals within a family might yield biallelic variants that explain their epilepsy.

Families 205, 210, and 224 each have a third affected member. Therefore, performing WGS on these individuals might further refine the list of potentially causative variants. This would be especially useful in Family 205 where no variant identified segregated in an AR inheritance pattern. This strategy would also aid in the analysis of non-coding variants, as our initial analysis pipeline focused on the “low hanging fruit” of coding and canonical splice site variants.

One strength of the approach that we took in analyzing the WGS data from the affected individuals was that it was very unbiased. We did not initially focus on known epilepsy genes, but rather identified rare, homozygous variants. However, it must be noted that the MAF used was derived from gnomAD where the ethnic makeup is not reflective of our study population. Therefore, there is the possibility that we filtered out disease causing variants in our initial review of the data and using population databases such as GMEVariome and GenomeAsia 100K for MAF would be better when performing initial variant filtering. Additionally, we did not remove any synonymous variants that had high CADD scores. Typically, disease variants are

non-synonymous; however, synonymous variants can be pathogenic depending on the effect of the nucleotide change on binding or secondary structures. In Family 225, the only variant identified that subsequently passed segregation analysis was a synonymous change, c.249G>A; p.Thr83Thr, in *KIF18A*. While functional assays are important to validate likely disease-causing variants, it is especially important for synonymous variants as it can be less clear how they would impact protein function and lead to disease.

Lastly, sequencing of additional families should be undertaken. We only sequenced 10 of the 25 families. As no other study has investigated the etiology of epilepsy within the Afridi tribe it would be interesting to determine if any variants are common amongst all families, but not within the general population, and whether these variants may contribute to the epilepsy seen in these consanguineous families. This information would also provide insight towards the diversity of the genetic basis of epilepsy within the Afridi tribe and would be valuable in determining efficacious treatments for affected individuals.

Appendix C

Individual Contributions to Chapters and Appendices

Chapter 1

Kari Mattison wrote this chapter specifically for inclusion in this dissertation, and she generated all figures and tables contained within this chapter. Andrew Escayg assisted with the editing of this chapter.

Chapter 2

Kari Mattison performed RNA extraction and PCR analysis of the splice product, generated Figure 2.2, assisted with Figure 2.1, and wrote the initial draft of the manuscript. Kameryn Butler generated Figure 2.1, collected clinical information from families (Table 2.1) and assisted with editing of this manuscript. George Inglis transfected HEK cells for splice product analysis and mentored Kari Mattison in the generation of the constructs used.

Oshrat Dayan, Hanna Boussidan, and Baruch Kanner performed the GABA uptake assays and provide the associated text for the methods, results and discussion.

Vikas Bhambhani and Bryan Philbrook provided clinical information presented in Table 2.1. Cristina da Silva performed the initial bioinformatic analysis of the 460 patients. John Alexander assisted with the recontact of individuals in this study and with editing of the manuscript.

Andrew Escayg contributed to the experimental design and editing of the manuscript.

Chapter 3

This chapter is the result of collaborations with many researchers, genetic counselors, and clinicians. These individuals providing varying combinations of variant information, clinical information, and critical review of the information contained in this chapter include: Kyrieckos Aleck, Angel Aledo-Serrano, Adnan Alsadah, Lorenzo Botto, Jeroen Breckpot, George Burghel, Kameryn Butler, Matthew Deardorff, William Dobyns, Andrew Douglas, Holly Dubbs, Morten Dunoe, Jakob Ek, Andrew Fazenbaker, Francis Filloux, Kathryn Garber, Antonio Gil-Nagel, Trine Bjoerg Hammer, George Hoganson, Christina Hoei-Hansen, Mira Kharbanda, Eric Marsh,

Halie May, Emma McCann, Rikke Moeller, Karin Naess, Derek Neilson, Catherine Quindipan, Gordana Raca, Allison Schreiber, Arthur Sorlin, James Tao, Katrina Tatton-Brown, Bert van der Zwaag, Michele van Hirtum-Das, Nienke Verbeek, Anna Wedell, Andrea Whitney, and Anna Wrendenberg.

Kari Mattison generated the data in Figures 3.1, 3.2, 3.5, 3.6, 3.7, and 3.8, compiled information for Tables 3.1 and 3.2, and coordinated communication amongst all contributors to this chapter. She also wrote the initial draft of this chapter.

Gilles Tossing with the support of J. Alex Parker, both at the University of Montreal, generated the knock-in *C. elegans* lines and performed all related assays and analyses. This resulted in Figures 3.9, 3.10, and 3.11. They provided initial drafts of the text associated with these assays.

Fred Mulroe and Callum Simmons, with the support of Richard Baines, all at the University of Manchester, performed the seizure induction in *Drosophila* and the associated data analysis that generated Figure 3.3. They provided initial drafts of the text associated with these assays. Jim Warwicker, at the University of Manchester, performed the *in silico* modeling of patient and gnomAD variants and generated Figure 3.4.

Philippe Campeau (University of Montreal), Siddarth Banka (University of Manchester), and Andrew Escayg provided intellectual contribution to experimental design, administrative support and editing of the text in the chapter.

Chapter 4

Kari Mattison performed variant analysis and interpretation, and segregation analysis. She generated all figures and tables presented in this chapter and wrote the initial draft of the text in this chapter.

Noor Badshah (University of Agriculture Peshawar, Pakistan) collected blood samples and clinical information for Family 1. He provided assistance with editing of the text. Micheal Taylor and Gayatri Vadlamani collected samples and clinical information for Family 2. Sequencing analysis of Family 2 was done by the National Health Service in the United Kingdom.

WGS and initial analysis, including variant calling and ROH analysis for Family 1, was performed by the Emory Integrated Genomics Core, H. Richard Johnston, and Pankaj Chopra. GSA and IBS analysis was done by H. Richard Johnson and Dave Cutler, respectively. They provided the text related to this for this chapter.

Michael Zwick, Sohail Ahmad, and Andrew Escayg provided intellectual contribution to experimental design and editing of the text in this chapter.

Chapter 5

Kari Mattison wrote this chapter specifically for inclusion in this dissertation, and she generated all figures and tables contained within this chapter. Andrew Escayg assisted with the editing of this chapter.

Appendix A

Kari Mattison compiled all information presented in this appendix into its final form. She directly generated Tables A1 and A3, and Figure A2. She complied the information provided by numerous collaborators (mentioned in **Chapter 3** contribution statement) to create Table A2 and Figure A1. Gilles Tossing created Table A4.

Appendix B

Kari Mattison wrote the initial draft of this appendix and performed all segregation and other analysis to generate the information in Tables B1-B7. She generated pedigrees as presented in Figure B1 based on information supplied by Noor Badshah. Noor Badshah collected blood samples and clinical histories for all families included in this appendix.

WGS and initial analysis, including variant calling and ROH analysis, was performed by the Emory Integrated Genomics Core, H. Richard Johnston, and Pankaj Chopra.

Andrew Escayg edited the final text presented in this dissertation.