IMPACT OF AUTISM SPECTRUM DISORDER ASSOCIATED *PTEN* POINT MUTATIONS ON PROTEIN FUNCTION AND STABILITY

by

Kathryn L. Post

B.A., Boston University, 2011

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Neuroscience)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

November 2019

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

Impact of Autism Spectrum Disorder associated *PTEN* point mutations on protein function and structure

submitted by Kathryn L. Post in partial fulfillment of the requirements for

the degree of Doctor of Philosophy

in Neuroscience

Examining Committee:

Kurt Haas
Supervisor

Douglas Allan
Supervisory Committee Member

Michael Gordon
University Examiner

Thibault Mayor
University Examiner

Additional Supervisory Committee Members:

Christopher Loewen
Supervisory Committee Member

Ann Marie Craig
Supervisory Committee Member

Abstract

ASD is a prevalent neurodevelopmental disorder characterized by social impairments and repetitive behaviors. Repeated studies indicate a role of genetics in the etiology of the disorder. In the search for genetic causes, whole exome sequencing has identified *de novo* likely gene disrupting and missense mutations in multiple, unrelated probands in many genes. The research community is now faced with the task of determining how dysfunction in genes caused by point mutations contribute to the pathophysiology of the disorder. From the high-confidence genes associated with ASD, we chose phosphatase and tensin homolog, *PTEN*, as our gene of interest. We first utilized an unbiased approach to identify novel genetic interactions of wildtype PTEN using a synthetic dosage lethality (SDL) screen in the *Saccharomyces cerevisiae* model system. Our screen yielded eight strong genetic interactions, 'sentinels' specific for PTEN catalytic function. We then tested for disruptions to these genetic interactions by designing a novel mini-SDL to screen a total of 97 PTEN point mutations. Functional scores obtained from this screen indicated the level of dysfunction caused by each point mutation. Interestingly, we found the dysfunction of variants to lie on a continuum with many variants displaying a range of partial loss of function (LoF). By testing the stability of all PTEN variants in both yeast and HEK293 cells we determined that many variants were dysfunctional due to a partial or complete loss of protein stability. Further, we also tested the ability of PTEN point mutations to negatively regulate the AKT pathway similar to wildtype PTEN. From this assay, we identified variants that were fully functional, partial LoF, complete LoF, and dominant negative. Through the creation of a biallelic PTEN knockout cell line we were able to deduce that the main mechanism of dominant negativity in this assay was likely due to interference of mutant PTEN with endogenous PTEN. Combined, our data indicates that point mutations of PTEN disrupt protein

function via several mechanisms. Understanding how each mutation alters function will hopefully advance our understanding of the disorder and lead to precision therapeutics in the future.

Lay Summary

Autism Spectrum Disorder (ASD) is a developmental disorder associated with repetitive behavior and disruptions to social development. One hypothesis for the origin of ASD is that *de novo* gene mutations may be underlying the disorder. *De novo* mutations are mutations in the genome which occur prior to or during embryo development and are therefore not inherited from either parent. Certain genes have been found to have a greater-than-chance burden of *de novo* mutations in individuals with ASD. However, the identification of these mutations alone does not indicate causality. Therefore, this thesis attempts to characterize the impact of *de novo,* as well as inherited, mutations on a specific ASD-associated gene, *PTEN*. By looking at protein function changes caused by mutations we hope to elucidate potential biological mechanisms underlying ASD.

Preface

The research described in this thesis was designed, performed, and analyzed by K.L. Post except where noted below and with assistance as noted below. All writing was performed by K.L. Post with editing and input from K. Haas, C. Loewen, A.M. Craig, D. Allan, and B. Young.

Chapter 1: Creation of figures and writing was performed by K.L. Post.

Chapter 2: The studies described in this section were designed by K.L. Post, C. Loewen, B. Young, and K. Haas. Identification of *PTEN* variants for this study was performed by K.L. Post, P. Pavlidis, S. Rogic, F. Meili, M. Belmadani, and D.B. Callaghan. Generation of plasmids was performed by K.L. Post as well as A. Niciforovic and F. Meili under K.L. Post's direction. The SDL screens and mini array screens were performed by K.L. Post. The liquid growth assays were performed using strains generated by K.L. Post in the mini array screen by B. Young. The results from the second generation mini array will be included in a publication in submission (Post KL, et al. (2019) Multi-model deep phenotypic profiling of disease-associated PTEN missense mutations identifies multiple molecular mechanisms underlying protein dysfunction. Unpublished manuscript) and the first generation mini array and the liquid growth assays will be included in another publication in progress (Young B, Post KL, Haas K, and Loewen CL. (2019) Sentinel Interaction Mapping (SIM) – A system for the functional analysis of human gene variants in yeast. Unpublished manuscript). Refinement of Balony code for mini array analysis was performed by B. Young. All statistical analysis was done by K.L. Post with input from M. Belmadani.

Chapter 3: The studies described in this section were designed by K.L. Post, K. Haas, and C. Loewen. All yeast western blots and analysis was performed by K.L. Post. The flow cytometry assays were performed under direction of K.L. Post by F. Meili. CRISPR/Cas9 PTEN-KO

HEK293 cell line was generated under direction of K.L. Post by W. Meyers, A. Cau, and S. Safa. Results from this chapter will be included in a paper in submission (Post KL, et al. (2019) Multi-model deep phenotypic profiling of disease-associated PTEN missense mutations identifies multiple molecular mechanisms underlying protein dysfunction. Unpublished manuscript). Statistical analyses for all assays was done by K.L. Post.

Chapter 4: Figure 4.2 contains data presented in Chapter 3 by K.L. Post as well as data collected in the labs of C. Rankin, D. Allan, S. Bamji, and T. O'Connor. Experimentations in these labs were performed by T. McDiarmid, P. Ganguly, R. Dingwall, and C. Harrington. Figure 4.3 contains data from (Mighell et al., 2018) and (Matreyek et al., 2018).

Table of Contents

List of Tables

List of Figures

List of Abbreviations

5-FOA – 5-fluorootic acid

A_{600} – absorbance at wavelength 600

ADHD – Attention Deficit Hyperactive Disorder

AKT – protein kinase B

ASD – autism spectrum disorder

a.u. – arbitrary units

BRRS – Bannayan-Riley-Ruvalcaba syndrome

Ca^{2+} – calcium

CADD – combined annotation dependent depletion

CAG – cytomegalovirus early enhancer/chicken β-actin

CREB – cyclic adenosine monophosphate response element-binding protein

CK2 – casein kinase 2

CMA – chromosomal microarray analysis

CMV – constitutive cytomegalovirus

CNS – central nervous system

CNVs – copy number variants

CT – carboxy-terminal Tail

DD – developmental delay

DMEM – Dulbecco's Modified Eagle medium

DMA – deletion mutation array

DNM – *de novo* mutation

DSM-5 – Diagnostic and Statistical Manual of Mental Disorders

E/I – excitatory to inhibitory synaptic ratio

ERK – extracellular signal-related kinase

EV – empty vector

FBS – fetal bovine serum

FMR1 – Fragile X mental retardation 1

fMRI – functional magnetic resonance imaging

FSC-H – forward scatter-height

G418 – Geneticin

$GABA_A$ – γ-aminobutyric acid receptor type A

GEF – guanine nucleotide exchange factor

gnomAD – genome aggregation database

GoF – gain of function

GSK3b – glycogen synthase kinase 3β

GWAS – genome-wide association study

HA – human influenza hemagglutinin

HDR – homology directed repair

HEK293 – human embryonic kidney 293

ID – intellectual disability

iPSCs – induced pluripotent stem cells

KI – knock-in

KO – knock-out

LGD – likely gene disrupting

LoF – loss of function

MAPK – mitogen activated protein kinase

MECP2 – methyl-CpG binding protein 2

MN^{2+} – manganese

MRI – magnetic resonance imaging

MS – missense

mTor – mammalian target of rapamycin

mTORC1 – mammalian target of rapamycin complex 1

NDDs – neurodevelopmental disorders

NHERF – Na^{+}/H^{+} exchanger regulatory factor

NLGN1 – neuroligin1

NRXN1 – neurexin1

OD – optical density

OE – overexpression

P-loop – phosphate binding loop

pAKT – phosphorylated AKT

PBD – $PI(4,5)P_2$ binding domain

PCR – polymerase chain reactions

PDGFR – platelet-derived growth factor receptor

PDK1 – phosphoinositide-dependent kinase-1

PDZ-BD – post-synaptic density binding domain

pEGH – pEGH-p*GAL*1

PHTS – PTEN hamartoma syndrome

PI3K – phosphoinositide 3-kinase

PI(3)P – phosphatidylinositol 3 monophosphate

$PI(3,4)P_2$ – phosphatidylinositol 3-4 biphosphate

$PI(3,5)P_2$ – phosphatidylinositol 3-5 biphosphate

$PI(3,4,5)P_3$ – phosphatidylinositol 3-4-5 triphosphate

$PI(4,5)P_2$ – phosphatidylinositol 4-5 biphosphate

PIP – phosphatidylinositol

PlK3 – polo-like kinase 3

pLI – probability of being LoF intolerant

PolyPhen – polymorphism phenotyping v2

PTCHD1 – patched domain containing 1

PTEN – phosphatase and tensin homolog

PTK6 – non-receptor tyrosine kinase 6

PTMs – post-translational modifications

RAK – fyn-related kinase

ROCK – RhoA-associated protein kinase

SC – synthetic complete

SDL – synthetic dosage lethality

SDM – site directed mutagenesis

SDS – sodium dodecyl sulfate

SEM – standard error of the mean

SFARI – Simons Foundation for Autism Research Initiative

sfGFP – super fluorescent green fluorescent protein

SGA – synthetic gene array

shRNA – short hairpin RNA

SIFT – sorting intolerant from tolerant

SII – stability impact index

SNAP2 – screening for non-acceptable polymorphisms 2

SNV – single nucleotide variant

SSC-A – side scatter-area

SSD-H – side scatter-height

TBST – tris-buffered saline and Tween-20

TORC1 – target of rapamycin complex 1

UAS – upstream activating sequence

UBC – ubiquitin C

VAMP-seq – variant abundance by massively parallel sequencing

VUS – variant of unknown significance

WES – whole exome sequencing

WGS – whole genome sequencing

WT – wildtype PTEN

WWP2 – WW domain containing E3 ubiquitin ligase

XIAP – X-linked inhibitor of apoptosis protein

Acknowledgements

I would like to first acknowledge my supervisor, Dr. Kurt Haas. You lead by example, and your passion for science and exploration of the unknown remind me why I fell in love with neuroscience. You gave me the opportunity to forge my own path while always being there to provide guidance and support. Thank you for your supervision throughout this journey. A big thank you goes to Dr. Christopher Loewen who kindly allowed me into his lab and who has extended enormous amounts of patience and support to a burgeoning neuroscientist novice to all things yeast. Through your support and inclusion in your lab I have found a love for cell biology and yeast genetics I never thought possible. Without your influence my PhD would look very different and I thank you for taking a chance on me. I would also like to extend my gratitude to my committee members, Dr. Ann Marie Craig, Dr. Douglas Allan, and Dr. Christopher Loewen for guiding me and providing support despite their busy schedules. Thank you also to my funders. My work could not have been completed without my fellowship from Kids Brain Health Network and Brain Canada as well as grant support from the Simons Foundation.

It takes a village, and I am forever grateful for the massive amounts of support from so many that I have received throughout this journey. A huge thank you goes to the investigators in the multi-lab collaboration. Drs. Paul Pavlidis, Sanja Rojic, Christopher Loewen, Douglas Allan, Timothy O'Connor, Catharine Rankin, and Shernaz Bamji initiated this great collaboration and gave me and my fellow graduate students the opportunity to be a part of a community of dedicated ASD researchers. Speaking of which, none of this would be possible without the camaraderie and friendship I gained from my SFARI-grad group. Manuel Belmadani, Benjamin Callaghan, Troy McDiarmid, Riki Dingwall, Matthew Edwards, and Payel Ganguly were brilliant teammates and brought much joy to my PhD. I know the future is bright for all of you.

Haas lab mates Janaina Brusco, Serg Opushnyev, Warren Meyers, Fabian Meili, Alessandro Cau, Sina Safa, Tristan Toth, and Dr. Wun Chey Sin as well as Loewen lab mates Dr. Barry Young, Andrew Wong, Analise Hofmann, Leslie Chan, Seeva Indran, and Jigyasa Verma have lent encouragement and company, oftentimes over a beer, more times than I can count. Additionally, I am very lucky to have the support of friends who have gone above and beyond. They have driven with me across the country, listened to me practice for every science talk, and reminded me that life exists outside of the lab. Thank you for always thinking I could do it.

Finally, I want to give a special thank you to my parents, my sister, and Sam who have been compassionate listeners, shoulders to lean on, and voices of reason throughout this process. I absolutely could not have done this without you.

Dedication

I dedicate this thesis to my family, both near and far, but especially to my cousin Parker Lane.

You lead and inspire me. I do what I do for you.

Chapter 1: Introduction

1.1 Introduction to Autism Spectrum Disorder

Autism spectrum disorder (ASD) was first recognized clinically in 1943 in the seminal paper by Leo Kanner (Kanner, 1943) describing 11 patients who, at the time, were unique both amongst each other as well as amongst all previously reported conditions. Initial diagnosis of this disorder hinged on the patients' inability to connect socially; however, Kanner also recognized the clear heterogeneity present in his patients. Significant progress has been made in our ability to characterize and diagnose ASD since the first descriptions of the disorder and the current iteration of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; American Psychiatric Association, 2013) characterizes ASD as a spectrum disorder with impaired social communication and restricted, repetitive behaviors. In keeping with Kanner's original observations, ASD is still understood to be a highly heterogenous disorder both clinically and etiologically, with probands presenting with a range of intellectual capabilities, verbal skills, and comorbidities, such as: gastrointestinal disorders, attention deficit hyperactive disorder (ADHD), intellectual disability (ID), and epilepsy. Although great efforts have been made to create inclusive and conclusive diagnostic criteria for ASD, the descriptive nature of the definition does not give any insight into the underlying pathophysiology. Additionally, the immense range of clinical presentations, and therefore the potentially diverse range of mechanisms, poses a challenge in determining the etiology of such a complex disorder. It is generally believed that there is not one cause of ASD, but that a combination of environmental and genetic factors likely contribute. So far, no one cause has shown a clear, completely penetrant effect.

1.2 Environmental factors for ASD

The developing brain is highly vulnerable to exposures both in utero and in the early years of childhood (Volpe, 2000). Nocuous exposures or experiences during this time especially can have a profound effect on brain and behavior development (Volpe, 2000; Nelson et al., 2019). Therefore, it is prudent to understand how and which environmental factors contribute to neurodevelopment and neurodevelopmental disorders (NDDs) in order to keep from inadvertently causing harm. Additionally, such insults may also interact with inherited or *de novo* genetic variation to increase risk for or severity of developmental disorders such as ASD. The search for such environmental factors has led researchers to test everything from parenting styles to medical treatments (Modabbernia et al., 2017). While no credible scientific evidence exists that incompetent mothers (Fombonne, 2003) or vaccinations cause ASD (Fombonne, 2003; DeStefano and Shimabukuro, 2019), there are many other credible exposures that may contribute to ASD-risk. For example, cumulating evidence suggests that advanced parental age may impact etiology of psychiatric and neurodevelopmental conditions, including ASD (Janecka et al., 2017). This is especially true for fathers of increased age since a higher percentage of *de novo* mutations (DNM)s in ASD probands are of paternal origin than maternal (Iossifov et al., 2014; Jónsson et al., 2017).

Because there are critical periods during early development that leave a brain particularly susceptible to adverse environmental factors (Hensch, 2005), a focus has been placed on deciphering in-utero exposures which may impart ASD-risk. To this end, maternal conditions during pregnancy such as gestational diabetes (Xiang et al., 2015) or maternal infection (Lee et al., 2015) have both been associated with increased ASD-risk. Data suggests that maternal infection, which has repeatedly been linked with ASD, increases risk due to the presence of

pathogens (Maeyama et al., 2018); however, there is also data which suggests that this association is not due to the specific pathogens but instead due to the maternal immune response (Brown et al., 2015). Additionally, circulating gonadal hormones specifically testosterone, while known to be important for normal brain development (Auyeung et al., 2013), when present at elevated levels has been correlated with increased ASD risk (Baron-Cohen et al., 2015). Furthermore, some evidence exists that environmental exposures such as: vitamin D deficiency, zinc deficiency, maternal use of valproate, and a short interpregnancy interval may also increase ASD risk (Bolte et al., 2019). These and other unknown risk factors could likely be contributing to ASD but may not be sufficient to be causal. Given that environment and genetics do not exist independently, there is additional evidence that environmental influences on gene expression, irrespective of sequence variation, may impact ASD risk. Epigenetic mechanisms such as aberrant DNA methylation have been proposed as a link between environmental and gene factors in ASD (Keil and Lein, 2016).

1.3 ASD genetics

The contribution of genetic factors to ASD pathogenesis has been repeatedly established from family and twin studies (Fombonne, 2005; Persico and Napolioni, 2013). ASD is recognized as having the greatest genetic contribution of all neuropsychiatric disorders due to its high heritability (>90% as estimated by twin studies), monozygotic twin concordance (73-95%), and noticeable sibling recurrence (15-25% for broad ASD) (Geschwind, 2011). However, no single genetic cause has been identified; instead, 100s of genes have been associated with ASD and more continue to be implicated as sequencing techniques improve and the number of probands sequenced increases.

Although there has been a large effort to attain genetic sequencing for many ASD probands, and certain genes are now associated with ASD at high levels of confidence, it is clear that not all probands have an identifiable rare genetic mutation. Therefore, it is believed that while some cases may be the result of single genetic variations, others may be due to the accumulation of many low impact, common variants (Bourgeron, 2015). In the search for genetic causes, early linkage analysis studies identified autism loci, including several regions of interest which contain genes for synaptic proteins such as neuroligin 3 and γ-Aminobutyric Acid Receptor Type A ($GABA_A$) receptors (Klauck, 2006). The search for common variants led to several genome-wide association studies (GWAS) in ASD cohorts. Although some GWAS have shown significant findings, the results are not generally reproducible between studies (Gaugler et al., 2014). However, a recent GWAS which had more than double the sample size of previous GWAS identified several significant loci, including one previously identified in another GWAS (Grove et al., 2019). This gives hope to the hypothesis that increasing sample size will assist in uncovering additional common variants contributing to the pathogenesis of ASD.

Another hypothesis for the lack of significant findings of common variants is that the incredible heterogeneity of the disorder increases the likely genetic variability and hinders identification of significant association findings. To overcome this, some studies employ sub-phenotyping with which they group people with ASD based on secondary factors such as intellectual capabilities, physical attributes, or comorbidities. However, even when sub-phenotyping is applied, so far, there have still been few genome-wide significant associations found (Chaste et al., 2015; Grove et al., 2019). Although sub-phenotyping has so far not proved effective in identifying significant common variants, another hypothesis is that the relative contribution of common variants to the disorder will differ between ASD families who have only

a single affected individual (simplex) versus families with two or more affected individuals (multiplex). One study estimated that 60% of the liability traces to additive effects in the multiplex families studied while only 40% of the liability traces to additive effects in the simplex families (Klei et al., 2012). While there are most likely many common variants, each might have a very small impact on the liability to ASD. Given the range of genetic variability, it is believed that the genetic causes of ASD will include low, medium, and high-risk mutations (Bourgeron, 2015).

1.3.1 Syndromic ASD

While the majority of ASD cases are not thought to be associated with a genetic syndrome, research into syndromic ASD has provided insight into genes and chromosomal regions which are likely critical for neurodevelopment. Dozens of rare medical genetic syndromes have high penetrance for ASD; however, of all the syndromes known to be comorbid with ASD, only Fragile X syndrome accounts for more than 1-2% of ASD cases (de la Torre-Ubieta et al., 2016; Sztainberg and Zoghbi, 2016). Such identifiable Mendelian genetic disorders, including Fragile X syndrome, neurofibromatosis, Rett syndrome, and tuberous sclerosis collectively account for approximately 5-10% of ASD cases (Devlin and Scherer, 2012; Sztainberg and Zoghbi, 2016). Research into the genetic causes of these disorders has identified genes which are now known to be important for normal brain development such as fragile X mental retardation 1 (FMR1) which is important for synaptic plasticity (Mila et al., 2018) and methyl-CpG binding protein 2 (MECP2), a transcriptional repressor with effects on neural stem cell maturation and synaptic plasticity (Qiu, 2018). While these genes may give insight into mechanisms of specific forms of syndromic ASD, because they account for such a small percentage of ASD probands and because people with syndromic ASD often have complex,

multiple comorbidities, they might not be extrapolatable to the larger population of probands. In an attempt to understand non-syndromic ASD, a substantial effort has gone into identifying *de novo* mutations in non-syndromic ASD probands.

1.3.2 *De novo* copy number and single nucleotide variation in ASD

To decipher the contribution of non-inherited genetic variations to ASD, chromosomal microarray analysis (CMA), whole exome sequencing (WES) and whole genome sequencing (WGS) on large samples from multiplex and simplex families have uncovered numerous genes disrupted in many, unrelated probands. CMA, which is capable of detecting submicroscopic chromosomal rearrangements, and other methods for detecting copy number variants (CNVs) such as comparative genomic hybridization, have been used to quantify and identify CNVs present in probands but not their parents or unaffected siblings. Interestingly, spontaneous CNVs are three-to-five-fold more frequent in ASD probands than in unaffected controls (Sebat et al., 2007; Iossifov et al., 2014). ASD cases are known to have a 1.19-fold increased burden of rare CNVs (Pinto et al., 2010) with approximately 13.1% of ASD probands carrying a rare CNV compared to 8.9% in unaffected siblings (Leppa et al., 2016). However, because each individual CNV is rare in the population of studied ASD cases (Sebat et al., 2007), and because many CNVs cover multiple genes, it is often difficult to find a causal link between a particular gene and increased ASD-risk from CNV results alone. Yet, there are several genes which are promising candidates for ASD-risk due to recurrent identification of CNV events including Neurexin 1 (*NRXN1*), Patched domain containing 1 (*PTCHD1*), Neuroligin1 (*NLGN1*), *SHANK1* and *SHANK3* (Woodbury-Smith and Scherer, 2018). The presence of a CNV is not a perfect diagnostic tool, unfortunately. For instance, in *NRXN1*, which is a gene with well characterized association to ASD, CNV deletions have also been identified in 0.02% of population controls

(Dabell et al, 2013). Supplementary support for ASD-risk in genes identified within CNVs also comes from other sequencing methods and discovery of further gene alterations in ASD probands (Leppa et al., 2016).

With the advent WES, the ability to rapidly sequence all protein coding sequences allowed for an unbiased mechanism to discover single-nucleotide variants (SNVs) specific to ASD probands. Through sequencing simplex families, researchers were able to identify *de novo* coding mutations present in offspring diagnosed with ASD. Several large efforts have taken advantage of WES to identify ASD-associated missense (MS) and likely-gene-disrupting (LGD) variants. LGD mutations are mutations which cause an early termination or frameshift leading to early termination. One such study identified 353 genes with DNM in ASD probands (Iossifov et al, 2014). Interestingly, these genes, although heterogenous, often converge in functional networks associated with synaptic development and transcriptional regulation (de Rubeis et al., 2014; Alonso-Gonzalez et al., 2018).

However, DNMs are not present exclusively in probands. Every individual is believed to carry 1-4 *de novo* SNVs within exon sequences, with up to 74 total SNVs (Bourgeron, 2015). Through ongoing genetic sequencing studies, it is also known that being a disease-variant carrier is not a rare event; all individuals are thought to carry from 40 to 110 variants previously reported by the Human Gene Mutation Database as disease-causing (Xue et al., 2012). This added complexity makes it unlikely that all DNMs identified in ASD probands will be causal and it makes identifying single causes of a heterogenous disorder such as ASD highly difficult. However, via WES in simplex families, hundreds of genes have been linked to ASD due to identification of LGD DNMs in probands (O'Roak et al., 2011; O'Roak et al., 2012a; Iossifov et al., 2012; Iossifov et al., 2014; de Rubeis et al., 2014; Codina-Sola et al., 2015). However,

despite these great efforts, there remain few known causal MS mutations identified in sporadic cases. Often, these mutations remain variants of unknown significance (VUS) and their implications on severity of the disorder remain unknown.

In addition to the search for protein coding variation in ASD cohorts, recent progress has been made in identifying potentially causal non-coding risk elements for ASD. WGS, in addition to identifying protein coding mutations, attempts to identify regions of importance in the approximately 98.5% of the genome which is not protein coding. Although the effects of these variations are often less obvious than LGD mutations, they are very likely affecting gene expression and may be potential targets for therapeutics. While it remains unknown to what extent non-coding variations contribute to ASD-risk, it is believed that variation, particularly in promotor regions, likely contribute (An et al., 2018).

1.4 Neurodevelopmental features of ASD

Although ASD is diagnosed clinically based on behavioral phenotypes, evidence for disruption of the central nervous system (CNS) comes from studies on post-mortem neuropathology, magnetic resonance imaging (MRI) and functional magnetic resonance imaging (fMRI). Post-mortem studies find that young ASD probands have a neuron overabundance as well as disruption in neuronal migration and cortical laminar organization (Varghese et al., 2017; Courchesne et al., 2019). This strongly suggests that ASD begins in the womb as these processes begin during gestation (Courchesne et al., 2019). Further evidence from MRI studies have recapitulated these findings and identified that enlargement of cortical surface early in a child's first year of life could be predictive of ASD-risk (Varghese et al., 2017). Clinical evidence also indicates an increase in the prevalence of macrocephaly in children with ASD (Amaral et al., 2008; Barnard-Brak et al., 2011; Varghese et al., 2017). Additional post-mortem studies found

that individuals with ASD have altered neuron size and dendrite morphology as well as an increase in inflammation, microglia density and excitability compared to control brains (Gupta et al., 2014; Rubenstein and Merzenich, 2003; Varghese et al., 2017). CNS regions consistently disrupted in ASD identified via neuropathology studies include the prefrontal cortex, fusiform gyrus, frontoinsular cortex, anterior cingulate cortex, hippocampus, cerebellum, amygdala, and the brain stem (Varghese et al., 2017).

There are several leading theories on what brain dysfunction causes ASD. Identification of detrimental differences that exist in the CNS of ASD probands has given insight into brain regions and mechanisms that are likely contributing. There is an abundance of evidence to support the theory that functional connectivity is disrupted in ASD. However, there continues to be uncertainty on whether the alterations observed are the cause or effect of ASD. One well recognized theory is that there is an increase in local connectivity but a decrease in long range connectivity in ASD. Studies using fMRI find reproducible decreases in cortico-cortical functional connectivity which show a reliable correlation with the severity of social deficits observed (Just et al., 2004; Gotts et al., 2019). Simultaneously, fMRI studies have also observed an increase in functional connectivity between the thalamus and the striatum in ASD (Gotts et al., 2019). Another leading theory on the cause of ASD is an increase in the excitatory to inhibitory (E/I) synaptic ratio (Rubenstein and Merzenich, 2003; Nelson and Valakh, 2015). Identification of this increased E/I ratio found in animal models of ASD and mutations in many genes implicated in this ratio associated with increased ASD-risk also give support to this theory (Nelson and Valakh, 2015; Varghese et al., 2017).

1.5 Animal models of ASD

In an effort to understand the impact of environmental and genetic ASD-risk factors, animal models have been made to address the implications of many of the hypothesized risk factors to development and behavior. These animal models give incredible insight into the consequences of environmental and genetic disruptions on brain and behavior development. Because genetic manipulation tools are most advanced in mice, transgenic mice are often chosen as models of neurodevelopment and NDDs. Mouse models of ASD have been made based on syndromic ASD genes (*FMR1*, *SHANK3*, *MECP2*, etc.) (Zoghbi, 2005; Bozdagi et al., 2010; Kazdoba et al., 2014; Varghese et al., 2017), well-studied risk genes (*CHD8*, *SCN2*, *Syngap1*, etc.) (Aceti et al., 2015; Kayayama et al., 2016; Varghese et al., 2017; Tatsukawa et al., 2019) and copy-number variants (15q11-q13 deletion, 15q13.3 microdeletion etc.) (Nakatani et al., 2009; Kogan et al., 2015; Varghese et al., 2017) as well as environmental risk factors such as: prenatal exposure to valproate, maternal autoantibodies and maternal immune activation (Varghese et al., 2017).

In order to test the face validity of mouse models in recapitulating ASD, mice are tested behaviorally for indications that the animal displays ASD-like symptoms. The ability of the model to phenocopy some aspect of ASD-associated behavior gives indication that the model could be useful in elucidating downstream consequences of the environmental or genetic manipulation causing the behavioral phenotype. For mouse models, researchers define ASD-like behaviors in terms of: (i) social interactions, such as an animal's preference for a novel mouse versus a novel non-social object or a tendency to engage in social interactions, (ii) social communication, such as ultrasonic vocalizations, and (iii) motor behaviors such as marble burying or self-grooming (Crawley, 2012). Given their face and construct validity, transgenic animals which display ASD-like behaviors are good options to test potential therapeutic

treatments. Additionally, not only do many of these models replicate behavioral features of ASD, they also replicate some of the neuropathological features found in post-mortem studies previously reviewed in section 1.4 (Varghese et al., 2017).

Since ASD is such a complex developmental disorder that presents heterogeneously in the population, understanding the mechanism(s) underlying the behavioral phenotype will likely require the use of multiple model systems. While mouse models have proven useful as a mammalian model of human disorder, it is prudent to take advantage of the multitude of sophisticated animal model systems in order to capitalize on unique features of each organism. The mouse as a model of human disease, although frequently used, has had repeated failures in translating results to humans. Reproducibility and reliability are essential when attempting to make conclusions on the pathogenicity of genetic or environmental variation. Also, it is important to be careful when studying behavior in rodents as their behavior is greatly influenced by their housemates, their sex, their position in social hierarchy, their genetic background, their environment, and the sex and proficiency of the researcher (Sukoff Rizzo and Silverman, 2016). In addition to the hundreds of ASD models in mice, many additional model systems including worms, zebrafish, flies, and non-human primates are also utilized in an effort to best understand the implications of variation on cell function and animal development and behavior.

The nematode *Caenorhabditis elegans* (*C. elegans*) is a highly useful tool for studying disruptions in neural circuitry and behavior since they have 302 well characterized neurons and easily tracked behavior (Swierczek et al., 2011). Unlike rodents, however, the evolutionary distance between worms and humans can be a limitation for studying ASD. Despite this distance, their robust behaviors and genetic tractability make them ideal for determining effect of a gene on behavior. The genome-wide estimate of human homologous genes in *C. elegans* is 52% while

up to 80% of the SFARI category 1 and 2 genes have worm orthologs (Kaletta and Hengartner, 2006; Lai et al., 2000; Shaye and Greenwald, 2011). Therefore, worms have been used to study genetic mechanisms via neural activity, behavior, and cell biology experiments (Schmeisser and Parker, 2018). Similarly, zebrafish complement in-vitro studies and mammalian *in-vivo* research due to their transparency, quick developmental timeline, and ease of manipulation. They are a good tool for studying simple social behaviors given their social preference and shoaling behavior (Meshalkina et al., 2018). As sociability is a fundamental diagnostic for ASD, it is beneficial to use a model in which it is possible to identify manipulations which reduce sociability. Additionally, unlike *C. elegans*, zebrafish have brains which are structurally and functionally similar to human brains which allows researchers to study brain development and neural signaling in regions homologous to those known to show perturbation in ASD probands.

Invertebrate models such as *Drosophila melanogaster* offer further opportunities for genetic manipulation. *Drosophila,* which have 75% conservation of sequence to human disease genes (Bier, 2005) also have well-established genetic tools based on temporally inducible upstream activating sequence (UAS)-GAL4 system which allows for precise, reversible genetic manipulation. Human gene functionalization in *Drosophila* can take advantage of the UAS-GAL4 and cell specific promotors to target tissue-specific editing. Given their fecundity and short developmental timeline, it is possible to test a large number of flies quickly.

Taking into consideration that ASD is a uniquely human disorder due to its central diagnostics associated with impaired communication and cognition, it may also benefit from studies conducted on animals with higher evolutionary conservation to humans such as the macaque monkey. With their well-developed brain and complex socialization patterns, macaques may be a good model for determining underlying pathophysiology and developing therapeutics.

Although they more closely match human developmental timelines, this long timeline makes it both time consuming and costly to do long term studies in monkeys. However, research using environmental models of ASD in monkeys suggest that maternal immune activation may act as a 'neurodevelopmental disease primer' that increases risk of a number of NDDs (Bauman and Schumann, 2018).

Animal models present a unique opportunity to test hypotheses on the cause of ASD with each model having unique advantages and disadvantages associated with cost, experimentation time, genetic tractability, and relevance to human pathophysiology and disorder phenotypes. The correlation of results between animal models gives increased credence for identified pathophysiology, genetic, and environmental factors important for ASD. The inability of one animal model to perfectly recapitulate ASD indicates the great potential for the use of multiple animal models in combination to understand the complexity of ASD and ASD-risk factors.

1.6 Multi-model approach to study variants of unknown significance

While this thesis attempts to characterize the impact of MS VUS associated with ASD on protein function and stability using yeast and human cell culture-based approaches, it is part of a broader collaboration using a multi-model approach to profile VUS. In order to take advantage of the benefits of different model systems, an array of assays was used to test VUS in diverse cellular environments, thus allowing for a more comprehensive understanding of the complexity of mutation impact on gene expression and protein function. Such profiling offers an optimized opportunity for distinguishing MS variants that are likely causal from benign variants, and to differentiate variants based on mechanisms and degree of protein dysfunction that may impact disease association, severity, or potential therapeutic approaches.

1.6.1 Multi-model collaboration

Given the increase in the prevalence of genetic sequencing and the uncertainty of the impact of MS genetic variation on protein function, there is an increasing number of VUS associated with diseases and disorders. This creates a huge challenge for clinicians and researchers in determining harmful variants and accurately associating deleterious MS mutations with pathogenicity. In order to address this challenge, a collaboration was initiated at the University of British Columbia which uses bioinformatics and biological models to identify mutations that are disruptive to protein function and/or stability as well as cell/animal development and animal behavior. In order to comprehensively screen proteins for multiple functions in multiple cellular and developmental environments, five model systems were selected for variant profiling. High-throughput assays performed in *Saccharomyces cerevisiae*, *Drosophila melanogaster*, and human embryonic kidney 293 (HEK293) cells are used to identify dysfunctional and unstable variants as well as screen for variants most likely to cause effect in the lower-throughput assays in rat hippocampal cultures and *C. elegans*. Work using *Drosophila melanogaster,* rat hippocampal cultures and *C. elegans* was performed by collaborators in the labs of Dr. Douglas Allan, Dr. Timothy O'Connor, Dr. Shernaz Bamji and Dr. Catharine Rankin while bioinformatic contributions were made by members of the lab of Dr. Paul Pavlidis. A full description of collaborator contribution to this work can be found in the Preface. My work in this thesis covers the yeast and HEK293 cell experiments as well as brief analysis of the multi-model approach using all model systems.

1.7 Yeast to study complex human disorders

Understanding how point mutations in potential disease-associated genes effect protein function is crucial for revealing the mechanisms that underlie complex disorders such as ASD.

The *Saccharomyces cerevisiae* yeast model system is a robust tool often used to parse gene and protein functions (Botstein and Fink, 2011). For decades, researchers have been fine tuning assays as well as yeast themselves such that they are an ideal model organism to study the functions of genes and proteins as well as complex cellular processes extrapolatable to other eukaryotes (Botstein and Fink, 2011; Duina et al., 2014; Giaver and Nislow, 2014; Ferreira et al., 2019). Additionally, yeast and the media used to grow yeast are inexpensive and yeast can be frozen down easily and thawed years later (Amberg et al., 2005; Duina et al., 2014). Yeast are easy to grow in the laboratory and they have been developed such that they grow on top of solid agar in round colonies making it easy to measure colony size, a good proxy of growth rate (Duina et al., 2014).

Huge advancements in the field came via the creation of numerous arrays and collections which make high-throughput experiments possible. The deletion mutant array (DMA), which incorporates a drug resistant gene in place of each individual non-essential yeast gene, allows for genome-scale experiments to be done time and cost effectively (Botstein and Fink, 2011; Giaever and Nislow, 2014). Additional mutant libraries such as the green fluorescent protein (GFP)-fusion library and other fusion libraries facilitate large-scale imaging and protein interaction assays (Botstein and Fink, 2011). Further systematic insertion of temperature-sensitive mutants in the place of essential yeast genes allows for investigation of interactions of essential yeast genes otherwise very hard to study (Li et al, 2011).

Although yeast are single celled organisms, many cellular processes in yeast are highly conserved across eukaryotes. In fact, although the human and *S. cerevisiae* genome diverged from a common ancestor approximately one billion years ago, about one third of the *S. cerevisiae* genome is accounted for by orthologous human genes (Kachroo et al., 2015).

Therefore, the molecular and genetic tools developed for high-throughput assays in *S. cerevisiae* provide a means to understand gene and protein function that can be used to deduce function in higher order organisms. This allows for identification of novel pathways and interactions in *S. cerevisiae* that are relevant for other eukaryotes, including humans.

Not only do yeast have about 30% orthology to humans, nearly 22% of the yeast genome have orthologous gene families associated with human disease (Kryndushkin and Shewmaker, 2011) which increases the ability of yeast experiments to uncover mechanisms of human disease. It also allows for complementation assays in which it is possible to test the ability of the human gene to rescue a yeast deletion. Recently, an analysis was done to assess if human disease variants of homologous yeast genes could functionally complement the yeast gene (Sun et al., 2016). In this way, this research group was able to identify likely dysfunctional human variants from likely benign variants. While this is a powerful model to study human genes with orthologous genes in yeast, the majority of human disease genes do not have orthologous yeast genes and therefore cannot be studied using complementation assays. However, because the yeast system allows for high-throughput screening of many genes or gene variants, it is an ideal model to determine effects specific to sequence variation using other techniques.

As described in previous sections, ASD is a disorder of the nervous system that results in disruption to typical brain, social and behavioral development. In order to investigate human genes implicated in ASD in a model system, the model must be sensitive to the gene's functions. Although some genes associated with ASD encode proteins that are either localized specifically in neurons or which have neuron-specific functions, many of the genes identified to date are likely to have functions in yeast, including genes for proteins involved in chromatin modification, protein translation, Wnt signaling, and translational regulation (de la Torre-Ubieta

et al., 2016; Chahrour et al., 2016). Therefore, despite yeast lacking a nervous system, there is a significant potential that yeast will be able to identify ASD-associated gene and gene variant functions and interactors.

Furthermore, previous research in yeast has been successful in elucidating mechanisms of other human neurologic disorders. For example, knockout of *YFH1*, the yeast homolog of the frataxin gene known to cause Friedreich ataxia, a progressive degenerative disease of the CNS and heart, identified an important mechanism of this disorder. They found that a loss of *YFH1* causes iron build up in mitochondria leading to increased sensitivity to oxidative stress (Pandolfo, 1999) while yeast rescue experiments were able to identify functional variants (Pandolfo, 2008). In Parkinson's disease (PD), a neurodegenerative disorder, yeast assays have identified novel genetic and functional interactions of PD genes as well as regulators of α-synuclein toxicity (Kryndushkin and Shewmaker, 2011; Dhungel et al., 2015; Abeliovich and Gitler, 2016). Additionally, yeast have also proven to be a useful tool to study neurological proteinopothies such as amyotrophic lateral sclerosis and frontotemporal lobar degeneration (Kryndushkin and Shewmaker, 2011). Therefore, while yeast is not the perfect model for studying all human diseases or all human genes, there is overwhelming evidence that high-throughput screening in yeast can elucidate human gene functions, interactions, and potential disease mechanisms of human neurologic conditions.

1.7.1 Assay of genetic interactions in *S. cerevisiae*

One way to identify functional relationships of a gene and its corresponding protein is by exploring its genetic interactions. By taking advantage of model systems, researchers are able to search for genetic interactions through systematic and large-scale screening (Mani et al., 2008). A genetic interaction is defined by a quantitative discrepancy in the fitness of a double mutant

from the fitness of the single mutants (Mani et al., 2008). This difference can indicate a synergistic, or synthetic interaction, or an alleviating interaction. A synergistic interaction occurs when there is a reduction in the fitness of the double mutant greater than expected, which in extreme cases can be lethal (e.g. synthetic lethal), while an alleviating interaction causes an unexpected less severe decrease in fitness (Mani et al., 2008).

In order to study genetic interactions in yeast, an important milestone was the publication of the *S. cerevisiae* genome sequence which allowed researchers to connect genes and function (Botstein and Fink, 2011; Giaever and Nislow, 2014). This also allowed for the creation of the DMA (Giaever et al., 2002). The engineering that went into the creation of the DMA allows researchers to easily and quickly investigate genetic interactions with thousands of genes via mating of yeast from the DMA with query yeast strains (Giaever et al., 2002; Giaever and Nislow, 2014). Briefly, the DMA was created by replacing each non-essential gene in the yeast genome with the geneticin (G418) drug resistance KanMX gene and a synthetic barcode (Giaever et al., 2002). Media containing G418 will select for these strains and the barcoding makes it easy to verify for the correct deletion. Single deletion strains can be combined with additional deletion strains in order to study genetic interactions.

Following the creation of the DMA, the synthetic gene array (SGA) assay was designed to systematically construct double mutants for genetic interaction analysis (Tong et al., 2001; Giaever et al., 2002). SGA allows for the systematic identification of genetic interactions by methodically crossing a strain carrying a query mutation with the DMA using a colony-arraying robot (Tong et al., 2001). These genetic interactions are determined by analyzing the effects of double mutants on yeast fitness compared to the fitness of the single mutants (Tong et al., 2001; Young and Loewen, 2013). Studying double mutants in this way has allowed for the discovery of

nearly a million genetic interactions which either enhance or suppress growth phenotypes and has elucidated biological functions of previously uncharacterized genes (Botstein and Fink, 2011; Costanzo et al., 2016).

A version of the SGA, the synthetic dosage lethality (SDL) screen, uses this technology to examine synergistic, or synthetic, genetic interactions between a mutant and the overexpression of another gene. It identifies genes which when deleted in isolation yield viable yeast, however, when in combination with the expression of other genes, have a severe growth phenotype (Tong et al., 2001). By identifying these interactions, it is possible to uncover redundant pathways, networks of genetic interactions or direct protein interactions previously unknown (Boone et al., 2007). Many SDL studies have given rise to important findings on genes necessary for cellular mechanisms and pathways common to eukaryotes (Tong et al., 2001; Measday et al., 2005; Shin et al., 2016). Additionally, SDL screens have been used to identify genetic interaction of human genes with no yeast homolog (Jo et al., 2017). We plan to use the power of this assay to identify genetic interactions of ASD-associated genes with yeast genes and to exploit these interactions to determine the ability of VUS to recapitulate these genetic interactions.

1.8 HEK293 cells as a model system

HEK293 cells, originally described in 1977 (Graham et al., 1977), are used extensively by researchers to study protein expression and function. They are derived from human epithelial cells and they are often used to study the effect of protein expression on a human cells given that they are likely to appropriately express, fold, and post-translationally modify exogenously expressed human proteins (Thomas and Smart, 2005). However, because they are not neurons, they likely do not express all neuronal proteins that might be important for studying the effects of

exogenous expression of a neuron-specific protein. Many of the ASD-associated genes are not exclusive to the CNS but rather encode proteins such as those involved in signaling pathways, transcriptional regulation, and chromatin modification which are found in cells outside the CNS. Therefore, while not all ASD-associated genes will be ideal for study in HEK293 cells, many will be.

While cultured neurons are an ideal tool for evaluating the effect of aberrant protein expression on neuron function and development, because cultured neurons are post-mitotic, they require frequent collection and are more difficult to grow and transfect than immortalized cell lines. Additionally, they are generally collected from a rodent model, which inhibits them from identifying important human-specific targets of exogenous gene and protein expression. Meanwhile, HEK293 cells, which are of human origin, are known to be easy to grow and transfect and have a high rate of division (Thomas and Smart, 2005). Although HEK293 cells are generally thought of as renal-derived, they are known to express the markers of renal progenitor cells, neuronal cells and adrenal gland cells (Stepanenko and Dmitrenko, 2015). When HEK293 cells are induced to express neuronal proteins in mixed cultures with neurons, HEK293 cells can be used to study synaptogenesis and the synaptogenic properties of neuronal proteins (Stepanenko and Dmitrenko, 2015). Therefore, HEK293 cells were selected as a good model to specifically investigate effects of VUS on protein stability and functions.

1.8.1 Flow cytometry

Flow cytometry-based techniques are ideal for high-throughput analysis of multiple protein targets at once because they allow for single-cell imaging of multiple emission wavelengths. Because flow cytometry can quantitatively measure various fluorescence signals at single-cell resolution in large numbers of cells, it allows for the possibility of using multiple,

highly specific antibodies as well as genetically encoded fluorophores to quantitatively measure multiple factors, such as enzyme activity or protein stability, at a single-cell level (Davies et al., 2016; Matreyek et al., 2018). This approach has previously been used measure how ASD-associated synaptic proteins are impacted by environmental alterations (Kaushik et al., 2017). It is also used to identify and test for potential biomarkers indicative of disorder or disease state (Liang et al., 2016; Ahman et al., 2018). It is an ideal technology to analyze how expression of a gene or gene variant alters the expression or phosphorylation state of multiple proteins at once.

1.9 Selection of ASD genes for study

Given the complicated landscape of ASD genetics, scientists and clinicians have made efforts to categorize and rate genes for their likely involvement in ASD pathogenesis. As more data from genetic sequencing studies becomes available, databank resources are tasked with updating and prioritizing meticulously so researchers are kept up to date with the current state of genetic research. Determining the strength of association is made more complicated by the need to integrate results from studies with vastly different experimental conditions, rigor of positive results, and scope of genetic testing. Therefore, this is a fluid endeavor and genes often fluctuate in their 'rank' as more information becomes available. The Simons Foundation for Autism Research Initiative (SFARI) is a leader in the field of autism research and one of their initiatives is to make a reference for researchers to benefit from collated genetic studies. Their Human Gene Module considers the strength of evidence for each gene and scores genes to quantify their likely involvement in ASD (Abrahams et al., 2013). Genes are ranked from high confidence (Category 1) to not supported (Category 6) with an additional category (Syndromic) if they are linked to additional characteristics not associated with ASD. The gene studied in this thesis, PTEN, which is further described in sections 1.10 through 1.10.4, scores as Syndromic and Category 1 on the

SFARI Human Gene Module. Another gene scoring database is AutismKB (Yang C et al, 2018), which also finds significant evidence linking PTEN to ASD. In addition to these resources, several additional efforts exist which assist in collating all known variants of genes found in ASD probands which includes denovodb, VariCarta, AutDB, MSSNG, and others. Additional support for a gene's association with ASD can also come from enrichment analysis (deRubeis et al., 2014) and bioinformaticians assessment of strength of forecasting via machine learning (Brueggeman et al., 2018).

1.10 Association of the gene *PTEN* to ASD

The dual-specificity phosphatase and tensin homolog (*PTEN*) gene was originally identified as a tumor suppressor gene due to the detection of *PTEN* mutations in multiple sporadic cancers. Following this discovery, additional *PTEN* mutations have been found in the germline of patients with several cancer-predisposing syndromes including Cowden syndrome, Bannayan-Riley-Ruvalcaba syndrome (BRRS), and Proteus syndrome which are collectively referred to as PTEN hamartoma syndrome (PHTS). People with PHTS have a broad range of disruption to their neurodevelopment as well as benign tumors ranging from subtle skin papules to vascular anomalies (Mester and Eng, 2015). Additionally, people with PHTS are at increased risk for developing malignant tumors (Mester and Eng, 2015). The first indications of *PTEN* involvement in neurodevelopment and NDDs came through directed sequencing of *PTEN* in patients with PHTS comorbid for macrocephaly and ASD and characterized by germline *PTEN* mutations (Goffin et al., 2001; Butler, 2005; Blumenthal and Dennis, 2008). Given this association, an effort was made to complete mutational analysis of *PTEN* in ASD probands comorbid for macrocephaly but not diagnosed with PHTS (Butler, 2005). The identification of *PTEN* mutations in these probands caused an increase in *PTEN* mutational screening. Since then,

many additional germline and *de novo PTEN* mutations have been identified in patients with ASD, as well as in patients with both ASD and macrocephaly (Buxbaum et al, 2007; Herman et al, 2007; Orrico et al, 2009; Varga et al, 2009; McBride et al, 2010; Hobert et al, 2014; Klein et al, 2013; Vanderver et al, 2014). PTEN was chosen as an ideal candidate for study in our multi-model approach given that it exhibits non-neuron specific functions and has many identified ASD-associated VUS.

1.10.1 PTEN functions, binding partners, and localization

PTEN is a dual-specificity protein and lipid phosphatase with a preference to dephosphorylate the phosphate at the third position of the inositol ring of phosphatidylinositol 3-4-5 triphosphate ($PI(3,4,5)P_3$). In this mechanism, PTEN is known to act as a negative regulator of protein kinase B (AKT) by opposing phosphoinositide 3-kinase (PI3K) function (Figure 1.1). PI3K phosphorylates phosphatidylinositol 4-5 biphosphate $PI(4,5)P_2$, to increase available $PI(3,4,5)P_3$. The translocation of AKT to the plasma membrane via binding of its plekstrin homology domain to $PI(3,4,5)P_3$ causes AKT to be in close proximity to phosphoinositide-dependent kinase-1 (PDK1) which phosphorylates AKT, thus activating AKT. Once phosphorylated, AKT is then able to activate mammalian target of rapamycin complex 1 (mTORC1) which has many down-stream effects including increased protein synthesis, cell growth, proliferation, metabolism, angiogenesis, metastasis and cell survival (Sansal and Sellers, 2004; Martini et al., 2014; Ersahin et al., 2015). Therefore, reduction in PTEN function can lead to an increase in $PI(3,4,5)P_3$ and a downstream increase in activated AKT leading to dysregulated cell survival which may result in tumor phenotypes. While $PI(3,4,5)P_3$ is known to be PTEN's main catalytic substrate, it has also been found to dephosphorylate the third

phosphate in phosphatidylinositol 3-4 biphosphate PI(3,4)P_2 (Malek et al., 2017) and is known to bind phosphatidylinositol 3 monophosphate (PI(3)P) (Naguib et al., 2015).

In addition to its lipid phosphatase function, PTEN is known to act as a protein phosphatase which has also been linked to tumor suppression (Pulido, 2015). While PTEN's protein phosphatase activity is less well characterized than its lipid phosphatase activity, there are several known PTEN protein phosphatase substrates including PTEN itself (Tibarewal et al., 2012). Another example is the non-receptor tyrosine kinase 6 (PTK6) whose activation at the plasma membrane has been observed in many types of cancer and has been linked to survival and metastases of cancer cells (Wozniak et al., 2017). Additionally, the dephosphorylation of the actin-binding protein cofilin 1 by PTEN is believed to enhance actin depolymerization (Serezani et al., 2012) while the colocalization and dephosphorylation of cyclic adenosine monophosphate response element-binding protein (CREB) by PTEN alters transcriptional activation of CREB (Gu et al., 2011). Interestingly, an actin cytoskeleton-organizing protein, Drebrin, which plays a role in dendritic spine morphogenesis and organization is also known to be dephosphorylated by PTEN in a neuronal activity dependent manner (Kreis et al., 2013; Garcia-Junco-Clemente and Golshani, 2014). There is also evidence that the mitotic kinesin motor EG5, which is required for proper spindle organization during mitosis, is dephosphorylated by PTEN. Loss of PTEN could therefore cause elevated levels of phosphorylated EG5 which impairs its proper association with microtubules and centromeres thus hindering faithful chromosome transmission during mitosis (He et al., 2016).

Protein-protein interactions, independent of PTEN phosphatase activity, are also known to be important for PTEN localization, function, and stability. PTEN shuttles from the cytoplasm to the membrane and is also known to translocate to the nucleus where it has additional binding

partners and functions. PTEN's ability to interact with binding partners is important for proper localization and organelle-specific functions as well as regulation of function. Importantly, PTEN homodimerizes and is known to be catalytically active in its dimerized state (Papa et al., 2014). This homodimerized PTEN has also been found to exist in the nuclear and cytoplasmic compartments (Papa et al., 2014). The C2 domain is believed to bind PI(3)P located in the membranes of endosomes residing in the cytoplasm (Naguib et al., 2015) as well as $PI(4,5)P_2$ (Maehama et al., 2001).

Membrane localization of PTEN is thought to mediate its regulation of PI3K. PTEN achieves membrane localization via interaction of basic residues in its phosphatase and C2 domains with acidic lipids as well as its post-synaptic density binding domain (PDZ-BD), with which it has been ascertained interacts with several proteins at the membrane (Leslie et al., 2016). One such protein is the Na^+/H^+ exchanger regulatory factor (NHERF), which recruits PTEN to a complex with platelet-derived growth factor receptor (PDGFR) at the membrane. Since PDGFR is known to activate PI3K, it is believed that this interaction is important for regulating PI3K (Takahashi et al., 2006). Another protein, a neutral endonuclease Neprilysin, has been shown to interact with PTEN's C-terminal domain, recruit PTEN to the plasma membrane and increase PTEN stability and lipid phosphatase function and further inhibit the PI3K pathway (Sumitomo et al., 2004). Additional proteins with documented interactions with PTEN are also involved in cell migration, invasion, and proliferation and are associated with various cancers such as the scaffolding proteins MAGI-1b (Kotelevets et al., 2005) and MAGI-2 (Hu et al., 2007) as well as p85, the regulatory subunit of PI3K (Rabinovsky et al., 2009).

In addition to PTEN's established function at the plasma membrane, it has been identified translocating to the nucleus where it has additional functions and interacting partners (Figure

1.1). While PTEN does not have a nuclear localizing sequence, it enters the nucleus via diffusion, active shuttling and monoubiquitylation-dependent import (Planchon et al., 2008). While polyubiquitylation of PTEN targets it for degradation, monoubiquitylation by ubiquitin ligases such as NEDD4-1 (Wang et al., 2007), X-linked inhibitor of apoptosis protein (XIAP) (Van Themsche et al., 2009), WW domain containing E3 ubiquitin ligase (WWP2) (Maddika et al., 2011) and Ret-finger protein (Lee et al., 2013) have been shown to inhibit PTEN's phosphatase activity as well as cause its translocation to the nucleus (Bermudez Brito et al., 2015). A dominant nuclear localization of PTEN has been found in certain cell types (Leslie et al., 2016) while nuclear exclusion of PTEN has been found in sporadic cancer (Planchon et al., 2008) and certain ASD-associated PTEN variants (Fricano-Kugler et al., 2018). Once in the nucleus, PTEN is thought to regulate gene transcription and genomic stability (Planchon et al., 2008; Gong et al., 2015). Through its C2-domain, PTEN is believed to assist in maintaining heterochromatin structure independent of its phosphatase activity (Gong et al., 2015). PTEN has also been found to bind the centromere-specific binding protein C (CENP-C), which is required for centromere stability (Planchon et al., 2008; Gong et al., 2015). Additional evidence supports the understanding that PTEN can induce expression of the DNA-repair protein RAD-51 which is involved in double-stranded break repair (Shen et al., 2007). Nuclear PTEN is also believed to slow growth rates by causing cells to arrest in G0-G1 phase of growth by downregulating cyclin D (Planchon et al., 2008) as well as regulate nuclear PIP levels (Fricano-Kugler et al., 2018).

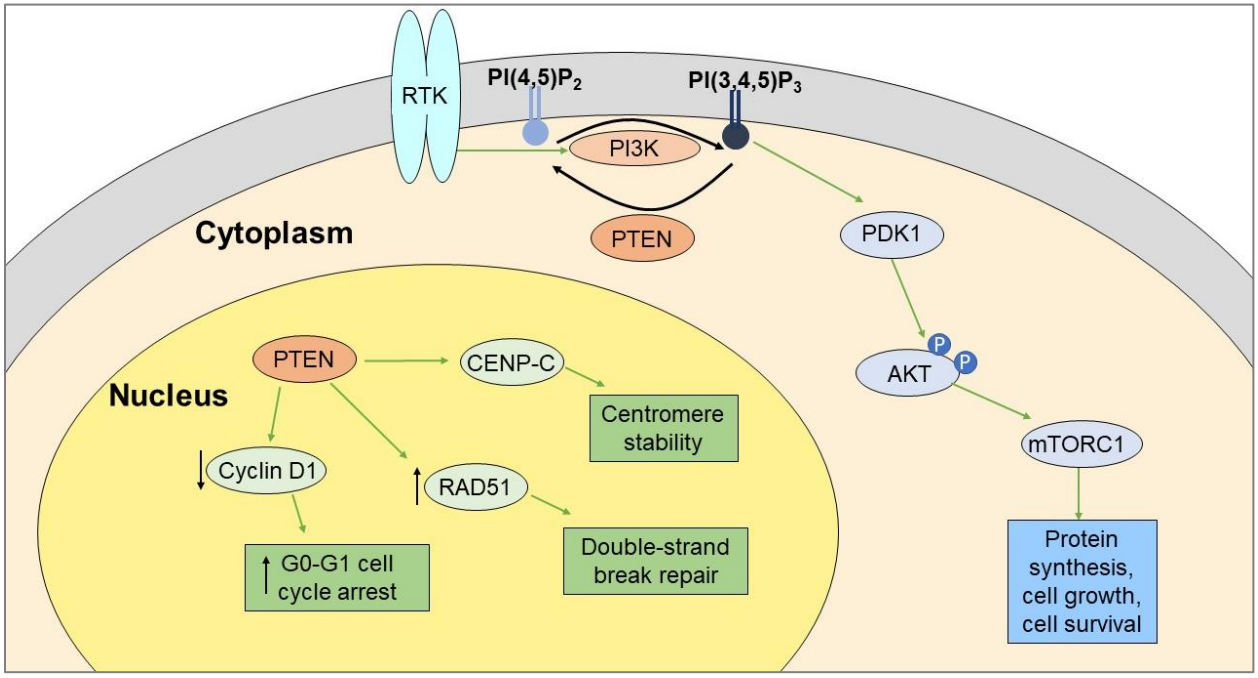

Figure 1.1: PTEN functions in the cytoplasm and nucleus

PTEN exists in the cytoplasm where it dephosphorylates $PI(3,4,5)P_3$ in antagonism of the kinase PI3K. PTEN also translocates to the nucleus where it has been found to increase centromere stability and double-strand break repair as well as arrest the cell cycle.

1.10.2 PTEN post-translational modifications

Canonical PTEN is composed of 403 amino acids and contains several important functional domains (Figure 1.2). It has a $PI(4,5)P_2$ binding domain (PBD) (aa 1-13), a catalytic phosphatase domain (aa 14-185), a C2 tensin-type domain (aa 190-350), a carboxy-terminal tail (CT) (aa 350-400) and a PDZ-BD (aa 401-403) (Bermudez-Brito et al., 2015). The crystal structure of PTEN was solved in 1999 (Lee et al., 1999). Within the phosphatase domain, the wall of the active site is made up of three loops: a phosphate binding loop (P-loop) (aa 123-130), a WPD-loop (aa 88-98), and a TI-loop (aa 165-171 (Smith and Briggs, 2016). PTEN is known to

undergo several post-translational modifications (PTMs) including phosphorylation, acetylation, ubiquitination, and oxidation. These PTMs are known to alter PTEN function, stability, localization, and protein-protein interactions.

Within the C tail, there are several known phosphorylation sites which are phosphorylated by casein kinase 2 (CK2) and S6K at S370, S380, T382, T383, and S385, by glycogen synthase kinase 3β (GSK3b) at S362 and T366, and by polo-like kinase 3 (PlK3) at T366 and S370 (Wang and Jiang, 2008; Bermudez-Brito et al., 2015). Phosphorylation of PTEN causes a decrease in phosphatase activity, a reduction in localization to the plasma membrane and in increase in stability (Ross and Gericke, 2009). Within the C2 domain, PTEN is phosphorylated at T366 by fyn-related kinase (RAK) and RhoA-associated protein kinase (ROCK) at S229, T232, T319 and T321 (Shi et al., 2012; Bermudez-Brito et al., 2015). While this is understood to also effect stability, it is believed that RAK activity prevents PTEN ubiquitination by NEDD4-1 thus inhibiting localization to the nucleus and initiation into degradation pathways (Bermudez-Brito et al., 2015). Additional details on PTEN ubiquitination can be found in section 1.10.1.

PTEN is thought to auto-dephosphorylate as well as be dephosphorylated by protein phosphatase 2A (Bermudez-Brito et al., 2015). In addition to phosphorylation, PTEN is also regulated by acetylation. Acetylation by p300/calcium-binding PTEN association factor at K125, K128 downregulates phosphatase activity while acetylation by CREB-binding protein at K402 modulates PTEN protein-protein interactions (Shi et al., 2012; Bermudez-Brito et al., 2015). Additionally, oxidation can reduce the catalytic activity of PTEN by forming a disulfide bond between C124 and C71 (Shi et al., 2012; Bermudez-Brito et al., 2015).

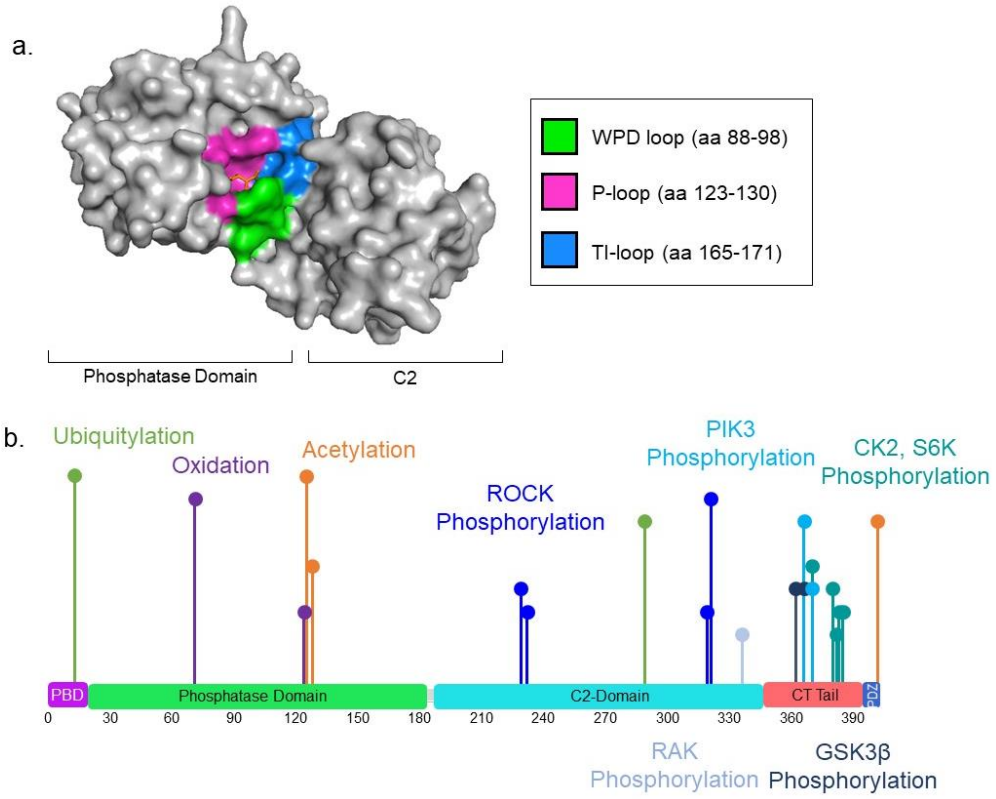

Figure 1.2: PTEN functional domains and post-translational modifications

The structure of PTEN, displayed using PyMol, (a) includes three loops which make up the catalytic pocket including the WPD-loop (aa 88-98), the P-loop (aa 123-130) and the TI-loop (aa 165-171). The functional domains of PTEN (b) include it's PBD (aa 1-13), phosphatase domain (aa14-185), C2-domain (aa190-350), CT-tail (350-400) and PDZ binding domain (401-403). PTEN post translational modifications include phosphorylation by ROCK (S229, T232, T319, T321), RAK (T366), PlK3 (T366, S370), GSK3β (S362, T366), and CK2 and S6K (S370, S380, T382, T383, S385) as well as ubiquitylation (K13, K289), oxidation (C71, C124), and acetylation (K125, K128, K402).

1.10.3 PTEN function in the CNS

The presence of a connection between ASD, macrocephaly and mutations in *PTEN* likely indicates a role for *PTEN* in CNS development. Murine *PTEN* knock-out (KO) and *PTEN* variant knock-in (KI) models have begun to elucidate the anatomical and functional

repercussions on CNS development due to decreased and mutated *PTEN* expression. Complete deletion of *PTEN* in mice is embryonically lethal and mice which are heterozygous for a *PTEN* deletion have a high incidence of tumors (Backman et al., 2001). Therefore, most animal models use either conditional or region-specific KO. The reduction of PTEN in conditional KO mice has repeatedly been correlated with a general increase in brain and neuronal soma size (Backman et al., 2001; Kwon et al., 2001; van Diepen and Eickholt, 2008).

Mice lacking *PTEN* specifically in the cerebral cortex and hippocampus exhibit macrocephaly, hyperactivity, increased anxiety-like behavior and deficits in social interactions in striking similarity to some ASD cases (Kwon et al., 2006). Conditional KO of *PTEN* in the hippocampus and cerebral cortex of mice also causes an increase in cell size and dendritic spine density accompanied with an increase in excitatory synaptic function (Fraser et al., 2008; Luikart et al., 2011). Additional studies found that *PTEN* knockout in neurons in the dentate gyrus led to somal hypertrophy, increased dendrite length, increased expression of the activity markers p-S6 and c-Fos, increased excitability, altered electrophysiological properties and increased mushroom spine and filipodia density (Williams et al., 2015). *PTEN* conditional KI mouse models are beginning to parse the effect of mutations on PTEN function in the brain. These alterations in behavior and development are thought to be at least partially mediated by downstream PTEN AKT signaling via the mammalian target of rapamycin (mTor) protein kinase as treatment with mTor has been shown to ameliorate the effects of *PTEN* knockdown in mice (Kwon et al., 2003).

To investigate the impact of specific *PTEN* variants on neuron development, one group took advantage of $PTEN^{flox/flox}$ mice which they co-injected with mCherry-T2A-Cre to knockdown endogenous *PTEN* expression and GFP-*PTEN* to express human *PTEN* or *PTEN*

variants (Fricano-Kugler et al., 2018). They found hypertrophy and increased pS6 signal in *PTEN* knockout animals and separate KI of six ASD associated variants was unable to rescue the observed hypertrophy and increased pS6 signal caused by *PTEN* deletion (Fricano-Kugler et al., 2018). Additionally, because mutations in PTEN are generally found in only one allele, some groups have attempted to study animals heterozygous for *PTEN* mutations. This is an ideal method for determining if the presence of mutated PTEN impacts wildtype PTEN. Indeed, several variants have been identified which are more detrimental in terms of AKT activation and tumorigenesis when expressed in combination with wildtype PTEN (Papa et al., 2014).

ASD affects brain regions in addition to the hippocampus and the cerebral cortex. Post-mortem analysis has identified a loss of Purkinje cells in the cerebellum as well as a reduction of neural density in the lateral nucleus of the amygdala (Wegiel et al., 2014). A Purkinje cell conditional *PTEN* KO mouse model identified an increase in Purkinje soma and dendrite size along with a decrease in axon terminal number, total Purkinje cell number, excitability, and frequency of spontaneous firing associated with *PTEN* loss (Cupolillo et al., 2016). In the amygdala, conditional *PTEN* KO mice have increased soma size, dendrite and mushroom spine diameter and neuronal excitability with no significant increase in spine density, dendrite length, or social and anxiety-related behaviors (Haws et al., 2014). Adult neural stem cells in the subventricular zone are also regulated by *PTEN* and KO in these cells causes an increase in terminal astrocyte differentiation (Gregorian et al., 2009; Bonaguidi et al., 2011).

Presynaptically, in conditional hippocampal KO mice, there is an enhanced synaptic vesicle number and release probability (Takeuchi et al., 2013; Garcia-Junco-Clemente and Golshani, 2014) which also likely contributes to synaptic strength changes following *PTEN* KO. Additionally, $PI(3,4,5)P_3$ has been found to be a chemo-attractive cue for growth cones in the

developing brain (Henle et al., 2013) and a reduction of PTEN function has been found to increase CNS axon regeneration via AKT-dependent and -independent pathways (Huang et al., 2019). In addition to KO models, overexpression (OE) of PTEN also gives insight into PTEN's role in neuronal function. For example, PTEN OE in organotypic hippocampal slices has been linked to a decrease in spine density which is sensitive to the protein phosphatase activity of PTEN (Zhang et al., 2012). Taken together, these results indicate that PTEN plays a role in modulating behavior, neuron and neurite size as well as synapse number, maturity, and excitability which is brain region and cell type specific.

1.10.4 VUS of PTEN

A proposed hypothesis stemming from studies done in hypomorphic mouse models of PTEN believes that PTEN function is best described on a continuum. In this model, PTEN protein function is best described using a non-discrete range instead of stepwise due to changes in copy number (Song et al., 2012). The second and third chapters of my thesis investigate the impact of 97 point mutations on PTEN protein function and stability. Importantly, the identification of *PTEN* point mutations in ASD cases does not prove causality of the disorder nor indicate that the mutation causes protein disruption. However, given the prevalence of variation in *PTEN* within the ASD population, it's imperative to experimentally test these variants to identify those which are affecting protein function and therefore potentially impacting ASD risk or severity.

Research groups have recently undertaken the task of testing VUS of *PTEN* in various assays. Several groups have studied a few PTEN mutations at depth (Rodríguez-Escudero et al., 2011; Rodríguez-Escudero et al., 2015; Papa et al., 2014; Fricano-Kugler et al., 2018) while a few have investigated thousands of mutations in *PTEN* using a singular bioassay (Mighell et al.,

2018; Matreyek et al., 2018). In 2018, two important studies were published which addressed the ramifications of *PTEN* VUS associated with human conditions such as ASD, PHTS, and somatic cancer on lipid phosphatase activity and protein stability.

In an article by Mighell et al., the researchers utilized a previously validated assay where yeast are humanized via overexpression of p110α, the catalytic subunit of human PI3K (Rodriguez-Escudero et al., 2005). When p110α is overexpressed in yeast, it phosphorylates $PI(4,5)P_2$ at the third position of the inositol ring which depletes the essential pool of $PI(4,5)P_2$. This depletion of $PI(4,5)P_2$ and subsequent increase of $PI(3,4,5)P_3$ is lethal to yeast without concurrent overexpression of PTEN to reduce $PI(3,4,5)P_3$ levels and increases the level of $PI(4,5)P_2$ (Rodriguez-Escudero et al., 2005; Rodriguez-Escudero et al., 2011; Mighell et al., 2018). Using saturation mutagenesis, this group was able to look at the ability of 86% of all possible *PTEN* MS mutations to antagonize the function of p110α. This double overexpression assay is ideal for specifically studying the lipid phosphatase activity of PTEN. However, it misses the impact of *PTEN* VUS on other mechanisms of PTEN dysfunction such as protein phosphatase activity, protein interactions, localization, stability, or dominant negativity.

Another paper by Matreyek et al., used HEK293 cells to determine the protein stability of *PTEN* MS mutations using saturation mutagenesis and variant abundance by massively parallel sequencing (VAMP-seq) techniques (Matreyek et al., 2018). Using these techniques they were able to identify 1,138 *PTEN* variants which caused low protein abundance suggesting a potential mechanism for *PTEN* mutations in disease or disorder (Matreyek et al., 2018). In combination, these studies add depth to our knowledge of how point mutations effect some aspects of protein function. Importantly, these and other studies give ample evidence to suggest that point mutations in *PTEN* can cause dysfunction. Determining the effect of VUS on PTEN lipid

phosphatase function and protein stability are the first steps in developing a comprehensive understanding of the impact of VUS on the totality of PTEN functions. In order to more fully assess a range of PTEN functions, this thesis as well as the collaboration detailed in section 1.6.1 hopes to take advantage of multiple model system to increase validation of variant impact.

1.11 Thesis overview and objectives

This thesis covers the selection and creation of *PTEN* variants as well as experiments designed to understand how point mutations disrupt PTEN function and stability. At the initiation of this study, no labs had published work looking at a large number of ASD-associated *PTEN* variants in multiple model systems. Due to the large number of ASD-associated variants identified in *PTEN*, it was an ideal candidate to investigate our hypothesis that ASD-associated variants are less functional than population or predicted low-impact variants. To examine the impact of 97 PTEN variants, we designed several experiments irrespective of previously determined protein functions which had the potential to test how variants could impact different aspects of PTEN. By taking advantage of variants with previously validated loss of function, we were able to identify variants with varying degrees of dysfunction. Importantly, we wanted to compare ASD-associated *PTEN* variants with somatic cancer associated variants, variants identified in people with PHTS, and variants identified in the general population. We used assays in yeast and HEK293 cells to test the hypothesis that *PTEN* variants associated with ASD, especially *de novo* variants, would be more deleterious to protein function than *PTEN* variants identified in the general population.

In Chapter 2, we designed initial unbiased yeast experiments to identify potential functions of PTEN in yeast with the objective to study the impact of VUS. Because there is no direct PTEN homolog nor $PI(3,4,5)P_3$ in yeast, we chose to take advantage of the yeast DMA to

identify novel genetic interactions specific to PTEN function. We hypothesized that expressing an exogenous human protein without a direct yeast homolog would identify genetic interactions pertinent to PTEN protein function. By doing SDL screens with WT PTEN, empty vector (EV), and the catalytically inactive variant C124S, we identified 44 potential genetic interactions of PTEN. In each of these deletion strains, overexpression of *PTEN* resulted in a reduction of colony growth. After these initial screens, we created a new mini SDL screen to test the effect of individual expression of each of 41 *PTEN* variants with each of the 44 deletion strains. Through this testing, it was clear that only 8 of the gene deletion mutants showed strong enough, significant negative genetic interactions with *PTEN* that we could use to see a range of function from the variants. Therefore, we created a new redesigned SDL in order to test all *PTEN* variants with the eight significant deletion strains in a high-throughput manner. These assays allowed us to use genetic interaction as a proxy for function given that the loss of function (LoF) variant C124S had no genetic interaction with these strains. We found that we were able to distinguish between WT-like levels of function, hypomorphic function, and LoF thus giving further credence to the hypothesis that PTEN protein function is on a continuum and not all-or-nothing. We were then also able to validate our results using a liquid growth based approach. By measuring the optic density of liquid cultures of *PTEN* variants expressed in one of the deletion strains we were able to get a more precise measure of the growth effects due to the genetic interaction.

In Chapter 3 we endeavored to investigate two potential mechanisms that underlie the PTEN dysfunction characterized via the genetic interaction experiments from Chapter 2. In this way we could test our hypothesis that the genetic interactions we identified are indicating levels of PTEN lipid phosphatase activity. Using yeast and HEK293 cell assays we determine which variants alter protein stability. Since KO animal studies have proven that loss of PTEN affects

brain development, knowing which variants have a LoF phenotype due to disrupted protein abundance could potentially impact therapeutic options. Additionally, we used flow cytometry metrics to determine the levels of phosphorylated AKT (pAKT) following expression of *PTEN* variants in HEK293 cells. Because increased levels of catalytically active PTEN reduces p AKT levels, a lack of, or decreased level of this reduction indicates a variant that has lost some level of function compared to wildtype PTEN. Interestingly, we identified several variants which indicate not just a LoF phenotype but that they may be acting in a dominant negative manner. To test the hypothesis that the mechanism of dominant negativity of these variants is due to interaction with endogenous PTEN, we created a PTEN-KO HEK293 cell line using the CRISPR/Cas9 technique. We then expressed several PTEN variants in this PTEN-KO HEK293 line and measured pAKT levels. We found that KO of *PTEN* eliminated or reduced the observed dominant negativity, giving credence to the hypothesis that these variants are interacting with endogenous PTEN to create a dominant negative effect. Finally, we compared the results from variants associated with ASD to those associated with PHTS and somatic cancers. We found no difference between ASD- and PHTS-associated, however, those variants found many times in somatic cancers were significantly more deleterious.

Chapter 2: *PTEN* variant selection and functional analysis

2.1 Introduction

Repeated identification of LGD mutations in *PTEN* from human genetic sequencing studies has strengthened confidence for mutations in *PTEN* causing an increase in ASD risk (Varga et al., 2009; Klein et al., 2013; Vanderver et al., 2014; Frazier et al., 2015). However, more *PTEN* VUS have been detected than *PTEN* LGD mutations (Abrahams et al., 2013). Advancing our knowledge of how specific VUS impact protein function is necessary for progress to be made in understanding if these VUS are deleterious to PTEN and how variation may impact PTEN protein function and the pathogenesis of ASD. While there has been progress made on how a lack of PTEN affects cell, circuit, brain, and animal development, until recently there have been few studies which undertake investigation of how a large number of the ASD-associated *PTEN* VUS effect protein function.

In addition to previous studies mentioned in section 1.10.4, further evidence for the likely deleterious nature of point mutations to PTEN function comes from WGS and WES of control populations. In the analysis of protein-coding variation, genes are assigned a Z score which indicates the difference between the number of expected versus observed missense mutations found in control populations (Lek et al., 2015). Positive Z scores indicate genes which are intolerant to variation, and PTEN has a Z score of 3.49 according to the gnomAD database (Lek et al., 2015). This indicates that *PTEN* mutations might be disruptive to normal development. However, in the yeast lipid phosphatase specific assay performed by Mighell et al., 67% of nonsynonymous mutations have wild-type levels of function and about half of all amino acid positions of *PTEN* are tolerant to missense mutation (Mighell et al., 2018). This incongruity between PTEN's high Z score and apparent tolerance to mutation in this particular yeast assay

suggest that performing additional assays may elucidate additional functional deficits caused by these and other *PTEN* VUS which may be relevant to disease.

In the work presented in this thesis, our goal is to characterize how specific VUS associated with ASD change the protein's ability to function in a similar fashion to the wild-type protein. In our initial experiments, instead of selecting a PTEN-specific assay, we chose to use a non-biased screen to determine genetic interactions of PTEN. From the results of this screen we endeavored to test the ability of *PTEN* VUS to maintain these genetic interactions. By taking advantage of the wealth of knowledge already available on yeast genes, we hypothesized that expressing an exogenous human protein even without a direct yeast homolog would yield genetic interactions pertinent to PTEN protein function.

2.2 Methods

2.2.1 Plasmids

For this study and the collaboration, a plasmid containing wildtype *PTEN* was acquired from Addgene (#28298) (Takahashi et al., 2006) and cloned into a pENTR (Thermo Fisher Scientific, catalog# K250020) vector using pENTR/D-TOPO Cloning Kit (Thermo Fisher Scientific, catalog# K240020). Creation of all 97 variants was done in the pENTR clone using site directed mutagenesis (SDM). In brief, SDM was performed using forward and reverse primer sequences (up to 60 nucleic acids long) which were designed to overlap the desired mutation. Polymerase chain reactions (PCRs) were completed using QuikChange Site-Directed Mutagenesis and Pfu polymerase (Agilent, catalog# 600390). Each variant was sequence validated. Once validated, variants were moved into assay-specific vectors using Gateway cloning. For yeast assays, PTEN was amplified from the pENTR vector with overhangs for the pEGH1 (Mitchell et al., 1993) yeast vector. The HEK293 destination vector (pCAG-mtagRFPt-

P2A-sfGFP) was used for HEK293 cell experiments (mtagRFPt and super fluorescent GFP(sfGFP) were gifts from Dr. Brian Chen, McGill University). Additional information about the creation of plasmids for the other model systems as well as a list of all plasmids acquired and created for this project can be found in the Appendix.

2.2.2 Yeast media and growth conditions

Yeast liquid and agar media was prepared using standard methods (Amberg et al., 2005). Transformations were performed with standard methods using homologous recombination (Amberg et al., 2005). Briefly, the expression of wild-type human *PTEN* or *PTEN* variants was induced in the Y7093 yeast strain through gap-repair. The Y7093 yeast strain was created through modification of the Y7092 yeast strain (genotype: *can1delta::STE2pr-Sp_his5 lyp1Δ his3Δ1 leu2Δ0 ura3Δ0 met15Δ0*) (a gift from Dr. Charles Boone, Toronto, CA) and has the *natMX6* marker integrated at the *TRP1* locus. To do gap-repair, a PCR product containing PTEN was amplified from the pENTR vector with overhangs for the pEGH1 yeast vector. The pEGH1 vector was cut with the Sac1 enzyme (New England Biolabs, Ipswich, MA) and phosphatase treated with calf-intestinal alkaline phosphatase (New England Biolabs, Ipswich, MA). The product was run on an agarose gel and gel purified using GeneJET Gel Extraction and DNA Cleanup Micro Kit (Thermo Fisher Scientific, cat #K0831). Transformations were performed using a 1:10 ratio of PTEN insert to the cut and phosphatase treated pEGH vector, thus taking advantage of yeast's affinity for homologous recombination. Successful homologous recombination was verified via plasmid recovery and transformation into *Escherichia coli* followed by sequence validation.

2.2.3 Synthetic dosage lethality screens

Synthetic dosage lethality (SDL) screens were performed by mating the Y7093 *MATa* starting strain transformed with pEGH1 (pEGH), pEGH1-*PTEN*, or pEGH1-*C124S* to the yeast haploid DMA (Open Biosystems) at 1536-array density using a ROTOR High Density Array/ Colony Picking Robot (Singer Instruments, Somerset, United Kingdom). The DMA is MATα and has G418 resistance as well as a *MFA1pr-HIS3* reporter. *MFA1pr-HIS3* allows yeast to synthesize histidine but is only expressed in MATa cells. The Y7093 query yeast strain (created through modification of the Y7092 strain, a gift from Dr. Charles Boone, Toronto, CA) has the *CAN1* and the *LYP1* genes knocked out which renders them canavanine- and thialysine-resistant respectively (Tong and Boone, 2005; Singh et al., 2009). Following mating in media containing agar (2g/100mL), yeast extract (1g/100mL) (Difco) and peptone (2g/100mL) (Bioshop), the diploids were copied in triplicate onto enriched sporulation media and incubated at 25°C for 14 days. Solid sporulation media contains agar (2g/100mL), potassium acetate (10g/L), yeast extract (1g/L), dextrose (0.5g/L), sporulation amino-acid mix (0.1g/L), and G418 (200mg/ml). After sporulation, *MATa* haploid cells were generated by germination on synthetic complete (SC) agar plates without histidine, arginine and lysine with canavanine and thialysine. SC media includes agar (2g/100mL), yeast nitrogen base without ammonium sulfate (1.5g/L), amino acid mix (2g/L), glutamate (1g/L), canavanine (100mg/mL), thialysine (100mg/mL), clonNAT (100mg/mL), and G418 (200mg/mL). Amino acid mix contains adenine (3g), uracil (2g), para-Aminobenzoic acid (0.2g), alanine (2g), aspartic acid (2g), asparagine (2g), cysteine (2g), glutamic acid (2g), glutamine (2g), glycine (2g), isoleucine (2g), methionine (2g), leucine (10g), phenylalanine (2g), proline (2g), serine (2g), threonine (2g), tryptophan (2g), tyrosine (2g), valine (2g), and inositol (2g). Amino acids were supplied from Bioshop. Treatment with

canavanine (100mg/mL) and thialysine (100mg/mL) selects for cells without the *CAN1* and the *LYP1* genes while the removal of histidine from the media selects for MATa cells expressing the *MFA1pr-HIS3* reporter (Tong et al., 2001; Tong and Boone 2005).

Cells expressing the plasmid and sentinel deletion were selected by two rounds of incubation on SC agar plates without histidine arginine, lysine, and uracil with canavanine, thialysine, and G418. Removal of uracil from the media selects for yeast expressing the query plasmid which contains the *URA3* gene which allows yeast to make uracil (Mitchell et al., 1993). A control set of single mutants was generated by two rounds of incubation on SC agar plates without histidine, arginine, and lysine with canavanine, thialysine, G418, and 5-Fluoroorotic acid. Yeast with the *URA3* gene have a sensitivity to 5-fluorootic acid (5-FOA) (Singh et al., 2009), so by adding 5-FOA to media, cells with the query plasmid cannot survive (Singh et al., 2009). This allows for the selection of control MATa haploids expressing only the deletion strains which came from the same sporulation plates as the experimental MATa haploids expressing the deletion strains and the query plasmid.

Growth of colonies was ascertained under inducing conditions (2% galactose, 2% raffinose) on SC media. Ratios of colony size were determined by comparing experimental colonies of haploids bearing a single gene deletion and expressing the query strain to control colonies of haploids bearing a single gene deletion without query overexpression using the analysis software Balony (Young and Loewen, 2013). Experimental colonies are grown on plates without uracil while control colonies are grown on plates with uracil. Hits for individual screens are automatically determined by Balony cut-offs dependent upon the range of colony sizes and the estimated noise within the experiment.

2.2.4 Liquid growth assay

To determine the growth rate of yeast we measured the optical density (OD) via absorption at wavelength of 600nm (A_{600}). As A_{600} is only directly proportional to cell number over a limited range, we took multiple measurements of serially diluted yeast cultures to calibrate our equipment by fitting the measurements to a polynomial equation of the form $y = a(A_{600})^3 + b(A_{600})^2 + c(A_{600}) + d$ (Jung et al., 2015), where y is the relative number of cells. When fitted to an exponential function of the form $a.e^{kx}$, the rate constant k is a measure of the rate of growth of that strain. High-throughput liquid growth assays were performed to acquire the growth rate constant of each variant. An automated plate reader Epoch2 (BioTek, Winooski, VT, USA) was used to determine the change in absorbance at A_{600} over time. The maximum phase of logarithmic growth was determined to occur for the first 6 hours of growth after dilution and plating in 96-well plates. *PTEN* variants expressed in the sentinel strain *Δvac14* were grown in liquid cultures overnight under inducing conditions (2% galactose, 2% raffinose). Each sample was then diluted to approximately 0.1 OD then allowed to continue growing for 4 hours until yeast were in log-phase growth. Then 96-well plates were seeded with one variant per column to give eight technical replicates and the Epoch2 was used to measure the OD of the liquid samples every 15 minutes for up to 24 hours. Therefore, each plate contained ten variants, PTEN and pEGH. The growth rate (k) for 87 total variants was measured in the *Δvac14* sentinel via this high-throughput liquid growth assay.

2.2.5 Generation of functional scores

Functional scores were determined for *PTEN* variants in both the first and second generation mini array as well as the liquid growth assay. For the mini arrays, colony size was first determined using the analysis software Balony (Young and Loewen, 2013). For the first

generation mini array, a difference score was determined via calculating the median of the difference between the size of the variant colony size and the size of the *PTEN* colony. To determine the functional scores for the second generation mini array, the difference between the median of the variant colony size and the median of the associated *PTEN* colony was calculated for each sentinel as was the difference between the median pEGH colony size and the median of the associated *PTEN* colony for each sentinel. The functional score is one minus the ratio of these two differences:

$$\text{Functional score} = 1 - \frac{\mathrm{M}(\mathrm{A}_{\mathrm{var}} - \mathrm{A}_{\mathrm{WT}})}{\mathrm{M}(\mathrm{A}_{\mathrm{pEGH}} - \mathrm{A}_{\mathrm{WT}})}$$

In this formula, the numerator is the median difference in spot size between the variant and the associated *PTEN* control and the denominator is the median difference in spot size between pEGH and its associated *PTEN* control. For the liquid growth assay, the difference between the k of the variant (k_{var}) and the k of *PTEN* (k_{WT}) was generated from strains grown on the same day on the same plate as was the difference between the pEGH expressing yeast k (k_{pEGH}) and the *PTEN* yeast k (k_{WT}), also grown the same day on the same plate. The functional score was calculated by generating a ratio of this difference:

$$\text{Functional score} = 1 - \frac{k_{var} - k_{WT}}{k_{pEGH} - k_{WT}}$$

2.2.6 Statistical analysis

Values are expressed as means ± the standard error of the mean (SEM). Spearman correlations were calculated using Microsoft Excel. *PTEN* variant colony sizes were compared to paired wildtype *PTEN* colony sizes and significance was determined via multiple t-tests in Balony. Growth rate constants in the liquid growth assays were calculated using GraphPad

prism. Kruskal-Wallis multiple comparison tests were performed in GraphPad Prism. Statistical significance was defined as $p<0.05$ and p-values are indicated in figures as $p<0.05$=*.

2.3 Results

2.3.1 Identification of PTEN Variants

We selected a total of 97 PTEN point mutations, including all known ASD- and developmental delay (DD)-associated MS variants identified from the literature for high-throughput assays in yeast and HEK293 cells. Additionally, because PTEN is a well-studied lipid and protein phosphatase, several variants had been previously characterized with known disrupted functions. These, referred hereafter as 'biochemical variants' were used as controls of known PTEN dysfunction. Additional VUS identified in PHTS and/or somatic cancer were selected for comparison to the ASD VUS. Furthermore, in order to better understand the full range of PTEN function in the overall human population, several MS VUS identified in control populations were selected for inclusion. Finally, given the plethora of bioinformatic tools available for predicting dysfunction, several variants not yet identified in any human population were selected based on several bioinformatic prediction algorithms and are referred to as 'predicted high-impact' and 'predicted low-impact' variants. All together, these selected variants are located across the entirety of the PTEN protein and in all known functional domains (Figure 2.1). The full list of variants can be found in Tables 2.1, 2.2, and 2.3.

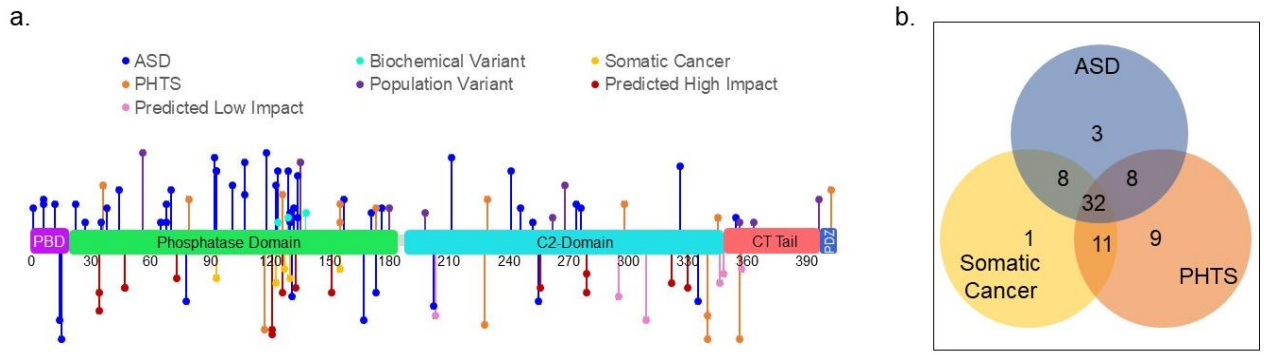

Figure 2.1: Summary of *PTEN* variants

PTEN is a 403 amino acid long protein with a PBD (aa 1-13), phosphatase domain (aa14-185), C2-domain (aa190-350), CT-tail (350-400) and PDZ binding domain (401-403). The variants selected for study in in this this are depicted in the lollipop plot (a). A Venn diagram depicting the overlap of classifications of variants associated with ASD, somatic cancer, and/or PHTS (b). Colored circles indicate the classification each variant belongs to with most variants present in more than just one condition. Variants not found in ASD, somatic cancer, or PHTS are not in this Venn diagram.

2.3.1.1 Biochemical variants

A literature search identified five variants with biochemically validated PTEN dysfunction: C124S, 4A, C124S-4A, G129E and Y138L (Table 2.1). The MS variant C124S, which is catalytically inactive, was originally identified in a glioma cell line and has been well established as a variant which has no lipid or protein phosphatase activity (Maehama and Dixon, 1998; Ishii et al., 1999; Rodriguez-Escudero et al., 2015). Since its discovery, C124S has been used extensively as a negative control for PTEN phosphatase function (Davidson et al., 2010; Rodriguez-Escudero et al., 2011; Rodriguez-Escudero et al., 2015; Fricano-Kugler et al., 2018). Additionally, this variant has been identified 16 times in cancer via the Catalogue of Somatic Mutations In Cancer (COSMIC) databank (Tate et al., 2019). Another well-studied variant, 4A contains four point mutations which disrupt the phosphorylation of PTEN. These four mutations

substitute an alanine at four of the phosphorylation sites on PTEN: S380, T382, T383, and S385 (Vazquez et al., 2001). The phosphorylation of PTEN at these sites causes a conformational shift in the protein leading to an increased interaction of the C-terminal domain with the phosphate binding domain (Rahdar et al., 2008; Song et al., 2012; Pulido, 2015). This decreases the ability of PTEN to bind substrates and reduces phosphatase activity. Therefore, by eliminating the potential for PTEN phosphorylation, the 4A variant eliminates the probability of this conformational change and renders the protein constitutively active (Rahdar et al., 2008; Song et al., 2012; Pulido, 2015). The C124S-4A variant combines the C124S mutation and the 4A substitutions. Despite being catalytically inactive, when expressed in hippocampal slices, this variant has been reported to significantly reduce spine density similar to overexpression of catalytically active PTEN (Zhang et al., 2012), suggesting that aspects of the protein, in addition to canonical phosphatase activity, may impact neuron development.

The G129E variant was initially discovered in the germline of two families with Cowden syndrome (Liaw et al., 1997). Further analysis found that this variant is incapable of dephosphorylating phospholipids, however, it retains the protein phosphatase activity of PTEN (Myers et al., 1998). G129E has also been identified in 10 cancer samples submitted to the COSMIC databank (Tate et al., 2019). Additionally, in a retrospective study, this mutation was found in two ASD probands. One proband had an inherited mutation and macrocephaly while the other had a *de novo* mutation and DD, macrocephaly, infantile seizures and PHTS (Hansen-Kiss et al., 2017). Finally, the Y138L variant was chosen as it was identified as a variant which selectively abolishes protein phosphatase activity while maintaining lipid phosphatase activity (Davidson et al., 2010). In a subsequent report by another group, this variant was verified to lack phosphatase activity towards proteins, however, unlike in the initial report, it maintained

considerably less lipid phosphatase activity than wildtype PTEN (Shi et al., 2014). Because these variants were selected due to their known impact on PTEN function, for this study, they will be denoted as biochemical variants in analysis and figures despite some also being found in other conditions.

Variant	SNAP2	CADD phred	ASD-associated	Somatic Cancer	#times in COSMIC	PHTS	gnomAD Freq	Known dysfunction	Clinical Pathology
4A	--	--	N	N	0	N	0	Constitutively active	
C124S	83	27.5	N	Y	16	N	0	Catalytically inactive	
C124S-4A	--	--	N	N	0	N	0	Catalytically inactive	
G129E	89	29.7	Y	Y	10	Y	0	No lipid phosphorylation	ASD, DD, macrocephaly, Cowden disease, BRRS (Hansen-Kiss et al., 2017); Cowden disease (Smith et al., 2019)
Y138L	44	--	N	N	0	N	0	No protein phosphorylation	

Table 2.1: Classification of *PTEN* biochemical variants

The five biochemical variants selected for this study with their bioinformatic scores, clinical associations and known dysfunction.

2.3.1.2 ASD-associated variants

All variants associated with ASD or other NDDs were selected for inclusion in high-throughput assays, regardless of the sequencing technique used to identify them or whether the inheritance status was known. This led to the inclusion of a total of 52 variants categorized as ASD-associated. This broad criteria allowed for a large number of variants with varying degrees of clinical annotation to be incorporated into this study. For all graphical analysis, any variant

(except biochemical variants) with an annotated connection to ASD are grouped in the ASD classification regardless of additional phenotypes noted. Information on inheritance, bioinformatic scores, phenotypes, additional classifications and references for all ASD-associated *PTEN* variants can be found in Table 2.2.

2.3.1.3 PHTS variants

PHTS encompasses several cancer-predisposing syndromes and is described in section 1.10. While each have slightly different clinical manifestations, they are all characterized by germline *PTEN* mutations and overgrowth including lipomas, macrocephaly, and hamartomas (Yehia et al., 2019). In order to graphically depict variants, any variant with an annotated link to PHTS but with no ASD-association nor greater than 8 somatic cancer links on the COSMIC databank, with the exception of biochemical variants, is displayed as a PHTS variant in analysis. Variants identified in Cowden syndrome, BRRS, Proteus syndrome, and hereditary-cancer predisposing syndrome are annotated as PHTS. All variants studied with PHTS association can be found in Table 2.2.

2.3.1.4 Somatic cancer variants

Because *PTEN* mutations are so often found in somatic cancers, many of the variants associated with ASD and PHTS have also been identified in cancers. While there are many resources that denote variants found in somatic cancer, the COSMIC databank was used for determining the frequency of somatic cancer association. The COSMIC databank lists variants which have been identified in somatic cancer as well as the frequency of identification and types of cancer in which they have been found in. For this dataset, a variant is classified primarily as a somatic cancer variant if it is found in at least eight cancer samples as annotated by COSMIC

while never found in any ASD probands. Full COSMIC annotations for all variants in our study and a list of all variants classified as somatic cancer variants can be found in Table 2.2.

2.3.1.5 Population variants

The wildtype sequence of *PTEN* used in this study is the NCBI reference sequence (NM_000314.7). While this is the canonical sequence, there are synonymous and protein coding mutations found in the human population. The genome aggregation database (gnomAD) combines data from the sequencing of 125,748 exomes and 15,708 whole genomes and measures the intolerance of genes to variation based on identification, or lack thereof, of predicted LoF variants (Karczewski et al., 2019). According to gnomAD, *PTEN* is expected to have a total of 20.8 LoF variants and 216.406 MS mutations. Expected variant counts are predicted by gnomAD using a depth corrected probability of mutation for each gene (Karczewski et al., 2019). However, *PTEN* has only 5 observed LoF variants and only 72 observed MS mutations found in control populations (Karczewski et al., 2019). This discrepancy indicates that *PTEN* is not tolerant to LoF and MS mutations and that mutations in *PTEN* may correlate with disease or disorder. From this list of 72 observed MS mutations in the purportedly healthy population, several of the variants with the highest frequency were selected for inclusion in the high-throughput assays. However, some variants associated with ASD, PHTS, and somatic cancer have also been identified in gnomAD. Annotation of variants included in this study identified in the population can be found in Table 2.2.

Variant	Primary Categorization	SNAP2	CADD phred	De Novo	ASD-associated	DD	Somatic Cancer	#times in COSMIC	PHTS	gnomAD Freq	Clinical Pathology
C124R	ASD	98	25.9	NA	Y		Y	2	Y	0	ASD, macrocephaly, leukodystrophy, lipoma (Smith et al, 2019); sporadic Cowden, macrocephaly, thyroid adenoma on thyroid (Nelen et al, 1997); 2 unrelated families with Cowden, germline (Marsh et al, 1998)
C135fs	ASD	--	--	Y	Y		Y	5	Y	0	ASD (O'Roak et al, 2012)
C211W	ASD	73	25.5	N	Y	Y	N	0	Y	0	ASD, DD (Vanderver et al, 2014)
D107G	ASD	39	26.1	NA	N	Y	Y	2	Y	0	DD, macrocephaly (Vanderver et al, 2014)
D107V	ASD	25	25.2	Y	Y		Y	2	Y	0	ASD (O'Roak et al, 2014; Hansen-Kiss et al, 2017); ID, Macrocephaly, Cowden Syndrome (Hansen-Kiss et al, 2017)
D22E	ASD	18	16.96	N	Y		Y	1	N	0	ASD, Found in father and one sib with ASD (Buxbaum et al, 2007)
D252G	ASD	80	26.5	NA	Y	Y	Y	6	Y	0	ASD, DD, macrocephaly (Butler, 2005; Buxbaum et al, 2007); speech apraxia, sensory integration problems, a short attention span (Butler, 2005)
D326N	ASD	74	35	Y	Y	Y	N	0	N	0	ASD, DD, macrocephaly (Buxbaum et al, 2007)
D92N	ASD	90	33	Y	Y		Y	4	N	0	ASD (Iossifov et al, 2014; O'Roak et al, 2014)
E157G	ASD	8	25.3	Y	Y		N	0	Y	4.06E-06	ASD, macrocephaly, identical twin also ASD (Varga et al, 2009)
F241S	ASD	21	17.1	NA	Y	Y	Y	5	Y	0	ASD, DD, macrocephaly (Butler, 2005; Buxbaum et al, 2007); overgrowth (Butler, 2005)
G129R	ASD	92	31	NA	Y	Y	Y	17	Y	0	2 ASD, DD, macrocephaly, epilepsy, leukodystrophy, benign thyroid growth, colon polyps, mucocutaneous, lipoma, generalized overgrowth (Smith et al, 2019); Cowden (Elia et al, 2012; Spinelli et al, 2015)

Variant	Primary Categorization	SNAP2	CADD phred	De Novo	ASD-associated	DD	Somatic Cancer	#times in COSMIC	PHTS	gnomAD Freq	Clinical Pathology
G132D	ASD	92	31	NA	Y	Y	Y	21	Y	0	ASD (Frazier et al, 2015; Smith et al, 2019), DD (Smith et al, 2019), macrocephaly (Frazier et al, 2015; Smith et al, 2019); leukodystrophy, lipoma, mucocutaneous lesions (Smith et al, 2019); patient2: PHTS patient: macrocephaly, skin cancer, kidney cancer (Smith et al, 2019)
G44D	ASD	65	31	NA	N	Y	Y	4	Y	0	DD, macrocephaly (Varga et al, 2009; Vanderver et al, 2014); speech delay (Varga et al, 2009); abnormal white matter (Vanderver et al, 2014); Hansen-Kiss says that this individual, also reported in Varga, has ASD (Hansen-Kiss et al, 2017)
H118P	ASD	85	23.6	Y	Y	Y	N	0	N	0	ASD, DD, macrocephaly, (Orrico et al, 2009) early diagnosis of ASD later questioned
H123Q	ASD	82	22	NA	Y	Y	N	0	Y	0	ASD, DD, macrocephaly (McBride et al, 2010)
H93R	ASD	64	23.2	Y	Y		Y	11	Y	0	ASD, macrocephaly, macrosomia, speech delay (Butler, 2005); PHTS, eosinohpilic gastrointestinal disorder, eczema, allergic rhinitis (Henderson et al, 2014)
H93Y	ASD	67	26.3	NA	Y	Y	Y	8	Y	0	ASD (Frazier et al, 2015; Smith et al, 2019), DD (Smith et al, 2019), macrocephaly (Frazier et al, 2015; Smith et al, 2019); lipoma, mucocutaneous lesions (Smith et al, 2019); patient2: PHTS, macrocephaly, skin cancer (Smith et al, 2019)
I101T	ASD	44	23.4	Y	Y	Y	Y	13	N	0	ASD (Smith et al, 2019), DD (Smith et al, 2019; Vanderver et al, 2014), macrocephaly (O'Roak et al, 2014; Smith et al, 2019; Vanderver et al, 2014); generalized overgrowth, dysmorphic features (Smith et al, 2019); Notes:also found in 1 unrelated

Variant	Primary Categorization	SNAP2	CADD phred	De Novo	ASD-associated	DD	Somatic Cancer	#times in COSMIC	PHTS	gnomAD Freq	Clinical Pathology
											patient with macrocephaly but no ASD (O'Roak et al, 2014)
I135fs	ASD	--	--	Y	Y		Y	5	Y	0	ASD (O'Roak et al, 2012)
K6E	ASD	67	19.93	NA	N	Y	Y	3	Y	0	DD, macrocephaly, hypotonia (Vanderver et al, 2014)
K6I	ASD	64	18.24	NA	Y	Y	Y	1	Y	0	ASD, DD (Vanderver et al, 2014)
L70V L320X	ASD	49	21.6	NA	Y	Y	N	0	Y	0	ASD (Frazier et al, 2015; Hobert et al, 2014; Smith et al, 2019), DD (Smith et al, 2019), macrocephaly (Frazier et al, 2015; Smith et al, 2019); epilepsy (Hobert et al, 2014; Smith et al, 2019); Oral mucosal papillomatosis (Hobert et al, 2014); lipoma, Hashimoto disease (Hobert et al, 2014; Smith et al, 2019); mucocutaneous lesions (Smith et al, 2019)
M134I	ASD	-15	31	N	N	Y	Y	6	N	0	Macrocephaly (Busa et al, 2013; Busa et al, 2015); Notes: 2 siblings with macrocephaly, one with speech delay, one with overgrowth
M134T	ASD	-22	22	N	N	Y	Y	1	Y	0	DD (McBride et al, 2010) mild MR, obesity, conductive hearing loss, affective disorder, and behavior problems (oppositional and anger issues) (McBride et al, 2010); PHTS (Smith et al, 2019); Notes: Multiple family members were affected with MR/DD, macrocephaly and/or cancer, with all tested affected individuals possessing the same PTEN mutation (McBride et al, 2010)
M1I	ASD	-92	19.05	NA	Y		Y	1	Y	0	ASD, macrocephaly (Frazier et al, 2015; Smith et al, 2019; Hobert et al, 2014); epilepsy, mucocutaneous lesions (Smith et al, 2019); Notes: 2

Variant	Primary Categorization	SNAP2	CADD phred	De Novo	ASD-associated	DD	Somatic Cancer	#times in COSMIC	PHTS	gnomAD Freq	Clinical Pathology
											ASD patients, also found in 2 PHTS patients (Smith et al, 2019)
M35V	ASD	44	23.6	NA	Y		Y	11	Y		ASD, PHTS (Smith et al, 2019)
N12T	ASD	-39	16.03	NA	Y		N	0	Y	0	ASD, macrocephaly (Frazier et al, 2015; Smith et al, 2019); mucocutaneous lesions (Smith et al, 2019)
N276S	ASD	87	24.5	Y	Y		Y	1	N	0	ASD, ID, macrocephaly, recurrent seizures (Orrico et al, 2009)
P246L	ASD	35	19.43	N	Y	Y	Y	13	Y	0	ASD, DD (Vanderver et al, 2014; Smith et al, 2019); macrocephaly, dysmorphic features, mucocutaneous lesions (Smith et al, 2019)
P354Q	ASD	-52	14.55	NA	Y		N	0	Y	8.34E-05	ASD, ovarian cancer (Smith et al, 2019); 2 Cowden (Mester et al, 2011)
P38H	ASD	44	28.1	NA	Y		Y	2	Y	0	ASD, macrocephaly, recurrent otitis media, lipomas, and pigmented glans of the penis (Klein et al, 2013)
Q171E	ASD	45	27.2	N	Y	Y	Y	3	Y	0	ASD, DD (Vanderver et al, 2014)
R130L	ASD	92	34	NA	Y		Y	27	Y	0	ASD, macrocephaly (Klein et al, 2013); PHTS, Colon polyps, mucocutaneous lesions, lipomas (Smith et al, 2019)
R130Q	ASD	81	36	NA	N	Y	Y	205	Y	0	ASD, DD, macrocephaly, hypotonia (Ciaccio et al, 2018); 5 PHTS, macrocephaly (Smith et al, 2019)
R130X	ASD	--	48	N	N	Y	Y	161	Y	1.22E-05	ASD (Klein et al, 2013), DD (Varga et al, 2009; Vanderver et al, 2014), macrocephaly (Vanderver et al, 2014; Klein et al, 2013); sensory processing disorder (Varga et al, 2009) ; Notes: same mutation found in normal father (Varga et al, 2009); has also been found in CS, BRRS, and CS/BRRS (Zori et al 1998; Parisi et al, 2001)

Variant	Primary Categorization	SNAP2	CADD phred	De Novo	ASD-associated	DD	Somatic Cancer	#times in COSMIC	PHTS	gnomAD Freq	Clinical Pathology
R14G	ASD	52	16.13	N	Y		Y	5	Y	0	ASD, macrocephaly (Frazier et al, 2015; Hansen-Kiss, et al 2017); 2 PHTS, macrocephaly (Smith et al, 2019)
R15S	ASD	65	20.9	Y	N	Y	Y	4	Y	0	DD, macrocephaly (Vanderver et al, 2014)
R173H	ASD	54	35	N	Y	Y	Y	60	Y	8.13E-06	ASD (McBride et al, 2010); DD (Varga et al, 2009); macrocephaly (Butler, 2005; Varga et al, 2009; Hansen-Kiss et al, 2017); ADD (Hansen-Kiss, et al 2017); Notes: found in two patients (McBride et al, 2010)
R335X	ASD	--	48	NA	Y		Y	43	Y	0	ASD, macrocephaly (Frazier et al, 2015; Hobert et al, 2014); seizure, macrocephaly, lipoma, overgrowth (Busa et al, 2015)
T131I	ASD	80	27.8	Y	Y		Y	2	N	0	ASD (O'Roak et al, 2012)
T167N	ASD	45	28.6	Y	Y		N	0	N	0	ASD (O'Roak et al, 2012)
T202I	ASD	-14	28.5	Y	N	Y	Y	1	Y	0	DD (Varga et al, 2009)
T78A	ASD	-77	12.12	Y	Y		Y	1	N	0	ASD (high functioning) without PHTS and head in 25% (Schaaf et al, 2011)
V255A	ASD	36	22	NA	Y		Y	2	N	0	ASD, macrocephaly (Klein et al, 2013)
W274L	ASD	81	26.1	NA	Y	Y	N	0	Y	0	DD, macrocephaly (McBride et al, 2010); Notes: patients has family history of DD, BRRS
Y176C	ASD	-47	13.94	Y	Y		N	0	Y	1.63E-05	ASD, macrocephaly, speech delay, absence of social interaction (Orrico et al, 2009)
Y27C	ASD	17	16.92	Y	N	Y	Y	8	Y	0	DD, macrocephaly, hypotonia (Vanderver et al, 2014)
Y65C	ASD	33	20.7	N	Y	Y	Y	1	Y	0	ASD, DD (Vanderver et al, 2014); ASD, DD, leukodystrophy, macrocephaly (Smith et al, 2019)
Y68H	ASD	86	23.3	NA	Y		Y	22	Y	0	ASD, macrocephaly (Frazier et al, 2015); ASD, DD, leukodystrophy, macrocephaly, dysmorphic features (Smith et al, 2019)

Variant	Primary Categorization	SNAP2	CADD phred	De Novo	ASD-associated	DD	Somatic Cancer	#times in COSMIC	PHTS	gnomAD Freq	Clinical Pathology
Y68N	ASD	87	23.7	NA	Y		Y	4	Y	0	ASD, flattened nasal bridge and plagiocephaly (Klein et al, 2013); PHTS, macrocephaly epilepsy (Smith et al, 2019)
A126P	PHTS	79	32	NA	N		Y	3	Y	0	PHTS, macrocephaly (Smith et al, 2019)
A79T	PHTS	-95	13.23	NA	N		Y	4	Y	1.00E-04	3 PHTS patients, Hashimoto, ovarian cancer (Smith et al, 2019); PHTS (Mester et al, 2011)
G36E	PHTS	82	31	NA	N		Y	4	Y	0	NA
I135V	PHTS	-72	20.7	NA	N		Y	7	Y	0	Cowden (Marsh et al 1998; Caux et al, 2007)
K402N	PHTS	41	12.21	NA	N		N	0	Y	0	NA
L345V	PHTS	36	11.92	NA	N		N	0	Y	0	PHTS, cancer (Smith et al, 2019); Cowden; (Heald et al, 2010; Mester et al, 2011)
N117S	PHTS	-92	10.49	NA	N		Y	1	Y	8.13E-06	NA
N228S	PHTS	-78	7.946	NA	N		N	0	Y	4.06E-06	NA
N340D	PHTS	-14	11.81	NA	N		N	0	Y	0	NA
N340H	PHTS	77	24.9	NA	N		N	0	Y	4.07E-06	NA
N356D	PHTS	-59	13.42	NA	N		N	0	Y	4.16E-06	PHTS (Smith et al, 2019)
Q298E	PHTS	-75	10.98	NA	N		N	0	Y	1.22E-05	NA
R173P	PHTS	31	31	NA	N		Y	1	Y	0	NA
S229T	PHTS	-84	11.24	NA	N		N	0	Y	0	NA
Y155H	PHTS	62	26	NA	N		Y	6	Y	0	NA
Y155N	PHTS	78	26.4	NA	N		Y	1	Y	0	NA
Y180H	PHTS	-78	11.78	NA	N		N	0	Y	4.07E-06	NA
G127R	Somatic Cancer	91	32		N		Y	7	Y	0	
H123Y	Somatic Cancer	80	28.2		N		Y	11	Y	0	
H93Q	Somatic Cancer	87	20.03		N		Y	5	N	0	
R130P	Somatic Cancer	93	28.7	Y	N		Y	15	Y	0	Macrocephaly, cutaneous lesions (Ciaccio et al, 2018)

Variant	Primary Categorization	SNAP2	CADD phred	De Novo	ASD-associated	DD	Somatic Cancer	#times in COSMIC	PHTS	gnomAD Freq	Clinical Pathology
Y155C	Cancer	22	23		N		Y	15	Y	0	
D268E	Population Variant	-89	1.668		N		Y	3	N	9.00E-04	
F56C	Population Variant	12	22		N		Y	2	N	4.29E-06	
I135T	Population Variant	33	24.2		N		N	0	N	8.24E-06	
M198I	Population Variant	-85	8.539		N		N	0	N	8.26E-06	
N262S	Population Variant	-63	11.48		N		N	0	N	4.06E-06	
N356H	Population Variant	-82	12.87		N		N	0	N	4.16E-06	
Q396R	Population Variant	-51	13.5		N		N	0	N	8.31E-06	
T363N	Population Variant	-53	12.86		N		N	0	N	4.22E-06	

Table 2.2: Classifications of *PTEN* ASD-associated, PHTS, somatic cancer, and population variants

This table includes all variants studied which were identified in a human population and denotes their primary classification, Combined Annotation Dependent Depletion (CADD) scores, Screening for Non-Acceptable Polymorphisms 2 (SNAP2) scores, *de-novo* status, ASD-association, developmental delay (DD) association, somatic cancer association, number of cancer samples identified with variant from COSMIC, PHTS-association, frequency in the population from gnomAD, and clinical pathology.

2.3.1.6 Bioinformatic variants

To add breadth to the variants selected based on association with human conditions, extra variants were identified for inclusion based on their bioinformatically predicted pathogenicity scores. This inclusion can aid in assessment of the quality of predictive algorithms. Although these variants have not yet been identified in any human population, they are scored by variant annotation tools as either likely or unlikely to cause disruption to protein function giving added dimension to the VUS identified in humans to date. A plethora of bioinformatic tools exist to score gene variants, including Polymorphism Phenotyping v2 (PolyPhen-2), Sorting Intolerant from Tolerant (SIFT), Combined Annotation Dependent Depletion (CADD), and Screening for Non-Acceptable Polymorphisms 2 (SNAP2). Tools such as SIFT and PolyPhen-2 have binary predictions of damage based on the effects of amino acid substitutions by related sequence alignments, predicted structure and the physical properties of the amino acids (Adzhubei et al., 2010; Sim et al., 2012). Because our biological assays are not binary, CADD and SNAP-2 were selected as more appropriate tools for variant selection and later comparison to the biological data.

CADD is a bioinformatic tool that estimates the relative pathogenicity of a single nucleotide substitution compared to all other single nucleotide substitutions. Annotation scores for CADD were generated using data which spanned a range of metrics including conservation, transcription factor binding, exon-intron boundaries, etc. as well as protein level scores from other predictive models such as SIFT and PolyPhen-2 (Kircher et al., 2014). In total, 63 annotations were made for each mutation which contribute to the raw CADD score. Following annotation, the model was trained via comparing differences in observed variants identified in human sequencing studies as well as the inferred human-chimpanzee genome to simulated

variants (Kircher et al., 2014). The raw scores of each substitution are mapped to a Phred-like scale to generate the final CADD phred score. This logarithmic score is defined by how damaging a particular mutation is compared to all other mutations in protein coding sequences (Kircher et al., 2014). Therefore, a mutation with a CADD score of 10 is more pathogenic than 90% of all other mutations in the genome and a score 30 is more pathogenic than 99.9% of all other mutations in the genome. This is a highly useful bioinformatic tool as it takes advantage of a number of known attributes of genome structure as well as protein-level annotations. Some variants studied do not have associated CADD phred scores because they either require more than one nucleic acid change or result in a frameshift mutation.

The other bioinformatic tool selected for inclusion was SNAP2, which generates predictive scores based on likelihood that a particular polymorphism will change the molecular function of a gene product (Bromberg and Rost, 2007; Hecht et al., 2014; Reeb et al., 2016). Unlike CADD, which seeks to measure pathogenicity, the goal of SNAP2 is to distinguish neutral variations from those that have the potential to alter protein function. SNAP2 provides a reliability index which gives confidence on how likely a particular amino acid substitution is to be either neutral (-100) or non-neutral (100) based on training using the protein mutant database, experimentally annotated enzymes and variants associated with disease (Bromberg and Rost, 2007; Hecht et al., 2014). Some variants do not have SNAP2 scores because they are either multiple amino acid changes (4A) or result in early protein termination. While both CADD and SNAP2 give annotation scores, because they are trained differently and because CADD focuses on nucleotide difference while SNAP2 focuses on amino acid difference they may each provide slightly different predictions for *PTEN* VUS. A list of all ‘predicted high impact’ and predicted low impact’ variants can be found in Table 2.3. CADD phred and SNAP2 scores can be found

for all additional variants in Tables 2.1 and 2.2. A graphical representation of CADD and SNAP2 scores for all variants tested in this study can be found in Figure 2.2. Because they were trained using slightly different methods, they do not have a perfect correlation. However, for PTEN they do correlate relatively well with a Spearman correlation coefficient of 0.7911.

Variants	Primary Catergoization	SNAP2	CADD phred	ASD	DD	Somatic Cancer	#times in COSMIC	PHTS	ExAC Allele Freq	ExAC Allele Counts	ExAC Allele #
A121E	Predicted High Impact	79	32	N	N	Y	2	N	0	0	0
A121P	Predicted High Impact	83	33	N	N	Y	3	N	0	0	0
A126D	Predicted High Impact	76	32	N	N	Y	5	N	0	0	0
A151P	Predicted High Impact	33	33	N	N	N	0	N	0	0	0
A34D	Predicted High Impact	86	32	N	N	Y	1	N	0	0	0
A34P	Predicted High Impact	82	33	N	N	N	0	N	0	0	0
E256K	Predicted High Impact	35	35	N	N	Y	2	N	0	0	0
E73K	Predicted High Impact	76	36	N	N	Y	1	N	0	0	0
F279I	Predicted High Impact	31	31	N	N	N	0	N	0	0	0
F279L	Predicted High Impact	31	31	N	N	Y	1	N	0	0	0
K322E	Predicted High Impact	31	31	N	N	Y	1	N	0	0	0
K330E	Predicted High Impact	31	31	N	N	N	0	N	0	0	0
R47K	Predicted High Impact	82	35	N	N	Y	1	N	0	0	0
R47W	Predicted High Impact	90	31	N	N	N	0	N	0	0	0
V133I	Predicted High Impact	87		N	N	Y	3	N	0	0	0
Y346F	Predicted Low Impact	-55	0.328	N	N	N	0	N	0	0	0
L295V	Predicted Low Impact	-67	4.363	N	N	Y	1	N	0	0	0
A309S	Predicted Low Impact	-80	4.925	N	N	N	0	N	0	0	0
P357S	Predicted Low Impact	-81	3.123	N	N	Y	2	N	0	0	0
T348S	Predicted Low Impact	-83	3.372	N	N	N	0	N	0	0	0
I203V	Predicted Low Impact	-84	4.851	N	N	Y	1	Y	0	0	0

Table 2.3: *PTEN* bioinformatic variants

Variants not yet identified in a human population or in fewer than eight cancer samples are classified as bioinformatic variants. The SNAP2 and CADD phred scores for all such variants as well as any identification in cancers are denoted.

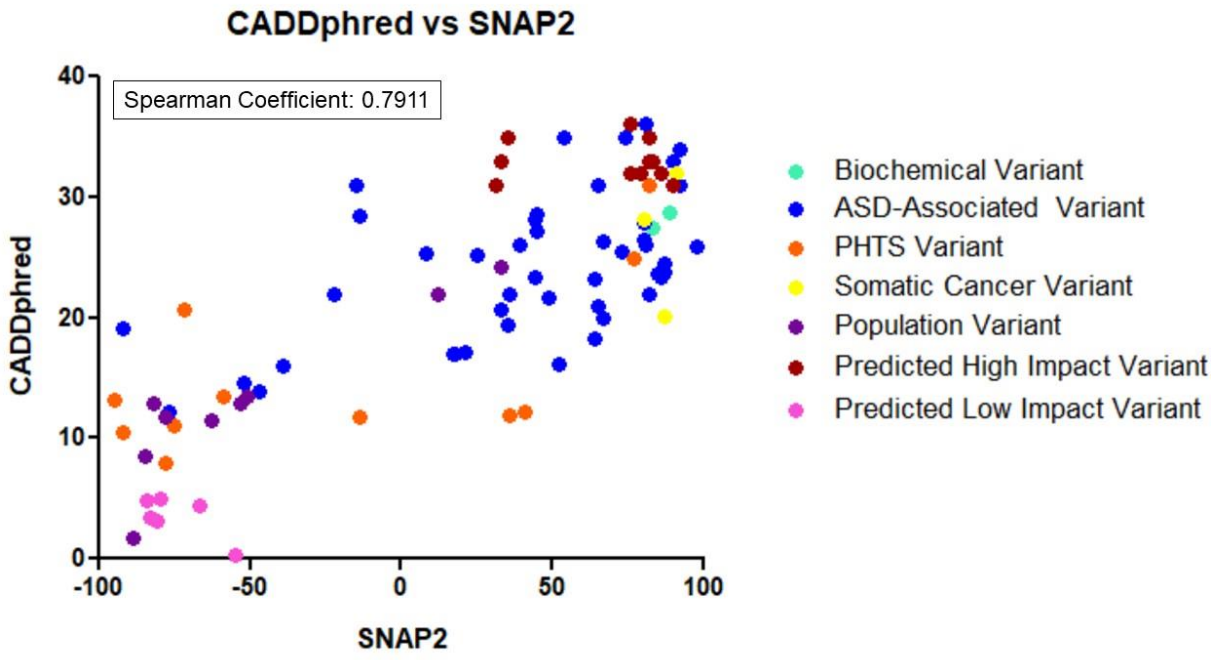

Figure 2.2: Correlation of bioinformatic scores for *PTEN* variants

CADD phred and SNAP2 scores collected for each variant tested in our assays. There is significant positive correlation between the two scores for the *PTEN* variants studied as indicated with a Spearman correlation coefficient of 0.7911 with a $p<0.05$ as determined by two-tailed t-test.

2.3.2 Identification of *PTEN*-specific genetic interactions

We aimed to identify genetic interactions specific to *PTEN* expression which could then be used to screen *PTEN* VUS for retention of function. To do this, we performed SDL screens in *Saccharomyces cerevisiae* (Figure 2.3). SDL screens allow us to take advantage of the incredible amount of engineering and foresight used to create the DMA yeast deletion strains. This type of

screen is ideal for identifying genetic interactions. By combining two different genotypes, in this case a yeast gene deletion and PTEN overexpression, we are able to identify instances where the combination of genotypes causes a phenotype not expected by analyzing the effect of the two genotypes independently. For our screens we are expressing a gene in all of the deletion strains to see if the expression of our gene of interest causes synthetic lethality. Not only do we express the wildtype version of *PTEN*, but we are also interested in identifying genetic interactions that are specific to PTEN protein function. Therefore, we also did an SDL screen of the empty vector, pEGH, and the catalytically inactive biochemical variant *C124S*. By comparing the results of these screens we were able to identify DMA strains which have a genetic interaction with *PTEN* but not with pEGH or *C124S*.

The screens we performed were versions of the SDL screen in which the query plasmids (*PTEN*, *C124S*, or pEGH) were introduced into each of the approximately 5000 nonessential strains of the yeast DMA. The plasmids were designed with a GAL promoter which allows for galactose induction, or query gene expression when galactose is used as the main carbon source in the media in place of dextrose. Briefly, to complete the SDL screen, yeast containing the query plasmids were mated with the DMA. These diploids underwent sporulation then haploids were selected for and allowed to germinate via incubation on media without histidine, arginine, and lysine with canavanine and thialysine. To identify genetic interactions in each screen, experimental and control plates were generated. Experimental plates contained media selecting for haploid colonies containing the deletion strain and the query plasmid while the control plates contained media selecting for haploid colonies containing only the yeast gene deletions. In this way, both the experimental and control plates were generated from the same sporulation source plate. Comparison of colony size between the experimental and control plates for all strains gave

an indication of cellular fitness and allowed for the identification of genetic interactions for each query. To increase confidence, two independent SDL screens were performed for *PTEN*.

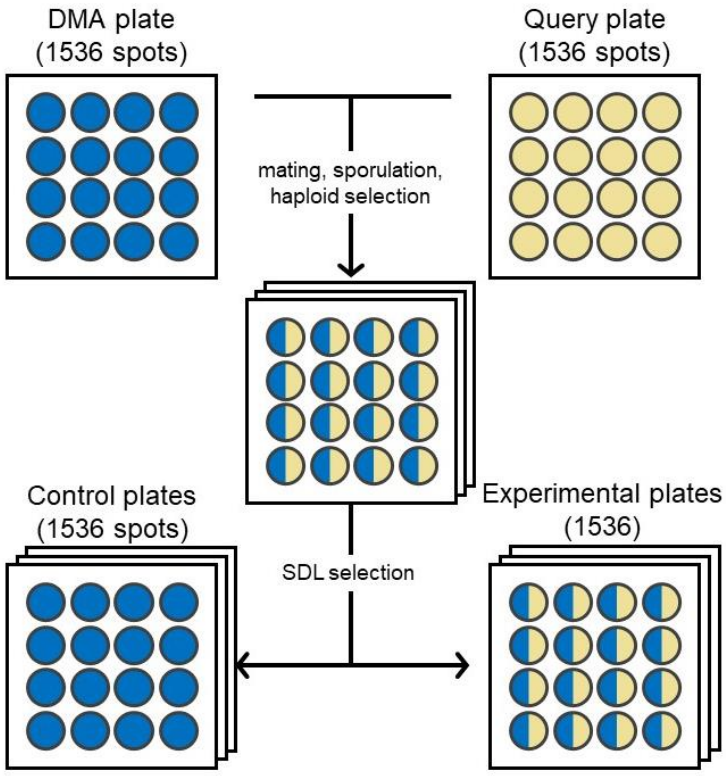

Figure 2.3: Cartoon of the synthetic dosage lethality screen (SDL)

The query plate is mated with each of the 4 DMA plates which combined contain colonies for the greater than 5000 yeast deletion strains. Technical triplicate plates undergo sporulation and haploid selection. Control plates are generated via incubation on SC-Histidine/Arginine/Lysine +Canavanine/Thialysine/G418+ 5-Fluoroorotic acid plates and experimental plates are generated via incubation on SC-Histidine/Arginine/Lysine/Uracil +Canavanine/Thialysine/G418 plates.

Comparison of the results from the *PTEN* screens with the results from the screens for *C124S* and pEGH yielded several strains which indicated a genetic interaction with PTEN but not with *C124S* or pEGH. We termed these strains ‘sentinels’. From these initial screens, seven high confidence hits were identified as well as 30 low confidence hits (Figure 2.4). High

confidence hits were yeast strains which had a genetic interaction in both PTEN SDL replicates but in neither the empty vector nor C124S SDL screens. Low confidence hits were strains which either show genetic interaction in only one of the PTEN SDL replicates or which show a genetic interaction in the empty vector and/or C124S SDL screens in addition to the PTEN SDL screens. While some of the low confidence hits showed very little difference in genetic interaction between wildtype *PTEN* and pEGH or *C124S*, by including a broad range of potential sentinels we hoped to identify sentinels reporting on different aspects of protein function in addition to those sensitive to mutation at C124. Due to PTEN's known function in phosphatidylinositol (PIP) dephosphorylation, an additional seven strains were selected which have known phosphoinositide functions despite not originally showing a strong genetic interaction with *PTEN*. Since we were naive to the repercussions of many of the VUS, we did not know if any of them might uncover a new genetic interaction with these strains. Although the purpose of this experiment was not to learn what human *PTEN* does in yeast, the identification of specific genetic interactions via this overexpression assay gave some insight into potential actions of exogenous PTEN in yeast. Because yeast do not express a direct homolog of PTEN nor the main catalytic substrate of PTEN, $PI(3,4,5)P_3$, prior to our experiment, it was unclear if PTEN would have a function in yeast.

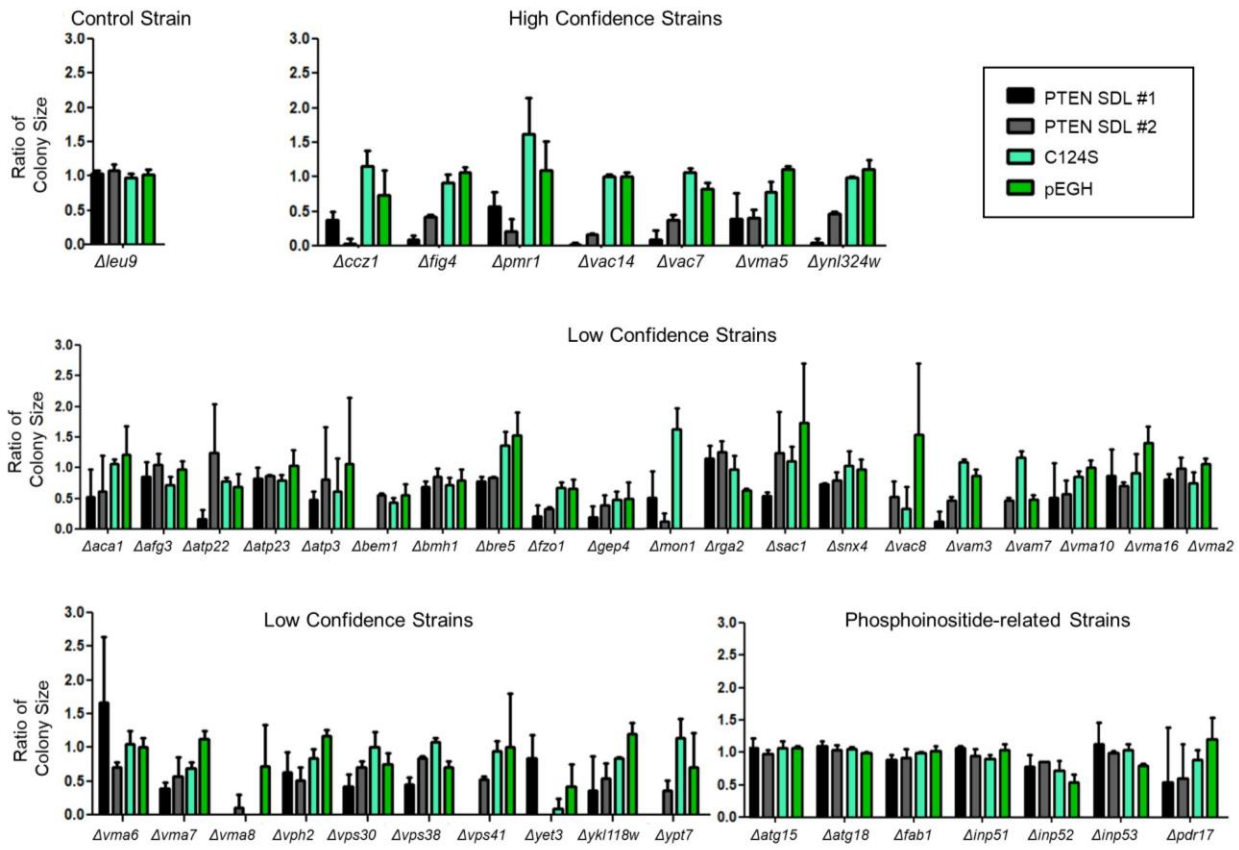

Figure 2.4: Deletion strains selected for inclusion in the first generation mini-array

The ratio of colony size for each sentinel expressing either *PTEN*, *C124S* or pEGH was generated from experimental results from the initial SDL screens. The ratios shown were generated by taking the ratio of median colony size of the experimental condition over the median colony size of the control condition. Strains which show a genetic interaction in 3/3 replicates from both *PTEN* screens are considered 'high confidence strains' while strains with genetic interaction with *PTEN* anywhere from 3/6 to 5/6 replicates are considered 'low confidence strains'. Phosphoinositide-related strains have known phosphoinositide-related functions in yeast. The control strain contains a deletion for alpha-isopropylmalate synthase II and shows no genetic interaction with PTEN, C124S or pEGH.

2.3.3 First generation mini array

2.3.3.1 Selection of deletion strains which report PTEN function

While SDL screens are often done to identify genetic interactions, our novel approach takes this one step further to look for a change in previously identified genetic interactions in variants of the gene of interest. Once a genetic interaction is identified, VUS can be tested by observing if sequence variants can replicate the interaction. Following completion of the original SDL screens, it was clear that most of the deletion strains in the DMA did not give insight into PTEN function and therefore were nonessential for further study with the *PTEN* variants. Therefore, in order to best screen all *PTEN* variants in the most efficient way possible, we designed a screen in which we could more rapidly test all variants with all relevant sentinels. To do this, we generated 'query plates' which contained wildtype PTEN and PTEN variants strategically arrayed and a novel 'sentinel plate' which contained yeast strains from the DMA selected for inclusion.

Strains best suited to screen for genetic interaction changes identified in the initial SDL screens were selected as sentinels and included in the first generation mini array. In this first iteration of the mini-array screen, we choose to include a broader selection of potential sentinel strains to make sure all relevant strains were included. Therefore, 45 sentinel strains were selected for inclusion (Figure 2.5). As previously described, many of the genes were selected based on experimental results and therefore likely to give insight into PTEN function. Strains were considered high confidence if they were hits in 3/3 replicates from both PTEN SDL screens. Low confidence strains were selected which were identified as hits from 3/6 to 5/6 total replicates from the PTEN SDLs. Additional deletion strains were included which have known phosphoinositide-related functions in order to better understand if PTEN variants have altered

effects on these strains. Finally, a control strain, *Δleu9*, was selected which showed no genetic interaction with *PTEN, C124S* or pEGH.

Because the sentinels for the first generation mini array were selected based mainly on experimental results a deeper look at proteins deleted in the high confidence strains may give insight into PTEN function. A full list of all sentinels selected for inclusion in the first generation mini array and their functions, cellular localizations, and effects of their knockout were compiled from the Saccharomyces Genome Database (Cherry et al., 2012) and can be found in Table 2.4. A more detailed look specifically at the high confidence hits gives interesting insight into what pathways and cellular functions *PTEN* is likely affecting. Gene deletion in many of the high confidence hits causes a disruption to vacuole morphology which may make yeast vulnerable to PTEN expression. Additionally, many of these proteins are localized at membranes and are known to interact with each other via physical or genetic interactions. For example, Ccz1 is a membrane bound guanine nucleotide exchange factor (GEF) which resides mainly in the late endosome and loss of this protein causes fragmented vacuoles, mis-localized vacuolar proteins and sensitivity to calcium (Ca^{2+}) (Kucharczyk et al., 2000). Interestingly, two proteins whose deletion strains were identified as potential sentinels in our initial screens are known to supress the calcium sensitivity of this strain, Ypt7 and Pmr1. Ypt7 is a small GTPase involved in the transport between late endosomes and vacuoles. Additionally, loss of this protein has the same phenotypic effects as the deletion of *CCZ1* which can be supressed with the overexpression of Pmr1. However while expression of *YPT7* supresses *Δccz1* effects, overexpression of *CCZ1* does not supress *Δypt7* effects (Kucharczyk et al., 2000). Pmr1, another protein which suppresses *Δccz1* effects, is a high affinity Ca^{2+}/manganese (Mn^{2+}) P-type ATPase required for Ca^{2+} and Mn^{2+} transport into the Golgi.

Another sentinel gene, *VMA5*, is one of many *VMA* genes identified and encodes a protein which is a subunit of the membrane domain of the vacuolar-ATPase which is required for the multi-subunit enzyme to properly assemble onto the vacuolar membrane (Ho et al., 1993). One study found that the *Δvma5* strain had defects in vacuole fragmentation as did several other sentinels identified in our screens, *Δypt7*, *Δvma2*, *Δvps38*, *Δfab1*, *Δfig4*, and *Δvac14* (Michaillat and Mayer, 2013). *FIG4* and *YNL324W*, which is an open reading frame that overlaps with *FIG4*, encode the PI(3,5)P_2 phosphatase. PI(3,5)P_2 is synthesized in yeast by Fab1 and decreased levels of PI(3,5)P_2 cause an enlarged vacuolar morphology while increases in PI(3,5)P_2 causes the creation of multiple, small vacuoles (Rudge et al., 2004). The PI(3,5)P_2 kinase, *FAB1* was chosen to be included in the first generation mini-array because it is positively regulated by both *VAC7* and *FIG4* which were identified as having strong genetic interactions with *PTEN*. The vacuolar membrane protein Vac7 is known to increase levels of PI(3,5)P_2 and genetic studies indicate a role for Vac7 as an upstream regulator of Fab1 (Gary et al., 2002). Finally, *Δvac14* which encodes an enzyme regulator involved in PI(3,5)P_2 synthesis is also known to cause a reduction in PI(3,5)P_2 levels (Dove et al., 2002). Vac14 is known to physically interact with Fig4 which increases the proper localization and activation of Fig4 (Rudge et al., 2004). The identification of all these genetic interactions point towards PTEN having a role in PI metabolism and/or the formation of normal vacuoles.

	Gene	Protein Function	Cellular Localization	Consequences of Null Mutation
Control	*LEU9*	Alpha-isopropylmalate synthase II	Mitochondria	no identifyable phenotype
High Confidence Strains	*CCZ1*	Subunit of a membrane-bound GEF	Late endosome	fragmented vacuoles, mis-localized vacuolar proteins, absence of autophagy, sensitivity to calcium and caffeine
	FIG4	Phosphatidylinositol 3,5-bisphosphate (PtdIns[3,5]P) phosphatase	Vacuole membrane	reduced mating efficiency, abnormal vacuolar morphology

	Gene	Protein Function	Cellular Localization	Consequences of Null Mutation
	PMR1	High affinity Ca2+/Mn2+ P-type ATPase	Golgi membrane	fragmented vacuoles, sensitivity to calcium and magnesium, reduced mating efficiency
	VAC14	Enzyme regulator	External component of vacuolar membrane	abnormal vacuolar morphology, increased cell size, decreased lifespan
	VAC7	Vacuolar segregation protein	Vacuole membrane	abnormal vacuolar morphology, increased cell size, decreased lifespan
	VMA5	Subunit C of the V1 peripheral membrane domain of V-ATPase	Vacuole membrane	abnormal vacuolar morphology, abnormal cell shape, abnormal budding, decreased lifespan
	YNL324W	Open reading frame overlapping Fig4		
Low Confidence Strains	*ACA1*	ATF/CREB family basic leucine zipper (bZIP) transcription factor	Nucleus	decreased competitive fitness
	AFG3	Mitochondrial inner membrane m-AAA protease component	Vacuole membrane, mitochondrion	reduced competitive fitness, increased resistance to oxidative and heat stress
	ATP22	Subunit of F1F0 ATP synthase	Mitochondrial inner membrane	slow growth, increased frequency of mitochondrial genome loss
	ATP23	Putative metalloprotease	Mitochondrial intermembrane space	slow growth, respiratory deficiency
	ATP3	Gamma subunit of the F1 sector of mitochondrial F1F0 ATP synthase	Mitochondrial inner membrane	slow growth, respiratory deficiency
	BEM1	Scaffold protein containing SH3-domains	Bud neck	slow growth, reduced mating efficiency, abnormal vacuolar morphology
	BMH1	14-3-3 protein	Nucleus, cytoplasm	increased gross chromosomal rearrangements, lengthened lifespan
	BRE5	Ubiquitin protease cofactor	P-body, cytoplasm	defect in retrograde transport from the Golgi to the ER, decreased competitive fitness
	FZO1	Transmembrane GTPase	Mitochondrial membranes	fragmented mitochondria, shortened lifespan
	GEP4	Mitochondrial phosphatidylglycerophosphatase	Mitochondrial matrix	abnormal mitochondrial morphology
	MON1	Subunit of a heterodimeric GEF	Cytosol, endosome, vacuole membrane	abnormal vacuolar morphology, absent autophagy
	RGA2	GTPase-activating protein	Bud neck septin rings	abnormal cellular morphology
	SAC1	Phosphatidylinositol phosphate phosphatase	Endoplasmic reticulum, Gogi	defects in protein secretion, endocytosis and vacuolar transport, abnormal actin cytoskeleton
	SNX4	Sorting nexin	Cytosol, early endosome	increased meta resistance, decreased mitophagy
	VAC8	Vacuolar membrane protein	Vacuole membrane, cytosol	decreased autophagy, abnormal vacuolar morphology
	VAM3	Syntaxin-like vacuolar t-SNARE	Vacuole membrane	shortened lifespan, vacuolar fragmentation defects
	VAM7	Vacuolar SNARE protein	Vacuole membrane	abnormal vacuolar morphology, absent autophagy
	VMA10	Subunit G of the V1 peripheral membrane domain of V-ATPase	Vacuole membrane	abnormal vacuolar morphology, decreased metal resistance
	VMA16	Subunit C of the vacuolar ATPase	Vacuole membrane	decreased resistance to chemicals, decreased metal resistance
	VMA2	Subunit B of V1 peripheral membrane domain of vacuolar H+-ATPase	Cytoplasm, vacuole membrane	sensitive to metal ions such as Zn2+, Cd2+, and Ni2+, shortened lifespan

	Gene	Protein Function	Cellular Localization	Consequences of Null Mutation
	VMA6	Subunit d of the V0 integral membrane domain of V-ATPase	Vacuole membrane	decreased metal resistance
	VMA7	Subunit F of the V1 peripheral membrane domain of V-ATPase	Vacuole membrane	decreased metal resistance
	VMA8	Subunit D of the V1 peripheral membrane domain of V-ATPase	Vacuole membrane	decreased metal resistance
	VPH2	Integral membrane protein required for V-ATPase function	Endoplasmic reticulum	decreased metal resistance
	VPS30	Subunit of PI3K complexes I and II	Cytosol, vacuole membrane	mitochondrial genome instability, reduced autophagy, shortened lifespan
	VPS41	Rab GTPase binding protein; Subunit of the HOPS endocytic tethering complex	Vacuole membrane, endosome	fragmented, enlarged, multilobed vacuoles, decreased growth rate, absent autophagy
	YET3	Endoplasmic reticulum transmembrane protein 3	Endoplasmic reticulum	reduced protein secretion
	YKL118W	Open reading frame partially overlapping Vph2 gene, an integral membrane protein required for V-ATPase function	Endoplasmic reticulum	decreased metal resistance
	YPT7	Rab family GTPase	Vacuole membrane	abnormal vacuolar morphology, decreased fitness
Phosphoinositide-related Strains	*ATG15*	Phospholipase	Vacuolar lumen	decreased lifespan, absent autophagy
	FAB1	1-phosphatidylinositol-3-phosphate 5-kinase	Vacuole membrane, endosome	defective vacuolar transport and morphology, sensitivity to calcium and metal ions, defective chromosome segregation
	INP51	Phosphatidylinositol 4,5-bisphosphate 5-phosphatase	Cytoplasm	abnormal vacuolar morphology, decreased competitive fitness
	INP52	Polyphosphatidylinositol phosphatase	Cytoplasm	abnormal vacuolar morphology, decreased endocytosis
	INP53	Polyphosphatidylinositol phosphatase	Cytoplasm	abnormal vacuolar morphology, abnormal cell wall morphology
	PDR17	Phosphatidylinositol transfer protein	Cytosol, cytoplasm, endosome	abnormal vacuolar morphology, reduced endocytosis, decreased lifespan

Table 2.4: Functions of the sentinel genes used in the first generation mini array

All sentinels used in the first generation mini array are listed with a brief description of their functions, cellular localizations, and dysfunctions associated with null mutations of each sentinel strain.

2.3.3.2 Technical design of first generation mini-array

After sentinel strains were selected for inclusion in the first generation of the mini-array, we designed a novel deletion plate which included only these strains. In order to reduce effects based on location on the plate, each sentinel was expressed in two distinct positions on the novel deletion plate. Additionally, because we were primarily interested in the difference in genetic interaction profile between wildtype *PTEN* and *PTEN* variants, we wanted to directly compare

each variant to a paired wildtype *PTEN* located in close proximity. In this way, we would be able to more easily account for inherent non-uniformity of growth on plates. For this experiment, unique deletion and query plates were designed such that, when mated, the effect of each *PTEN* variant on all sentinel strains could be analyzed against wildtype *PTEN* effects on all sentinel strains on the same plate with a sample size of eight. A cartoon depiction of the design can be found in Figure 2.5. Sentinel plates were designed such that all sentinels are represented in a 4x4 block twice on a 1536 colony-dense plate. Query plates were arrayed at 1536 density and designed with alternating rows of *PTEN* and *PTEN* variant. In similarity with the original SDL screens, the query and sentinel plates were mated then underwent sporulation under enriched sporulation conditions at 25C for 14 days. Haploid cells expressing the query plasmids and deletion strains were selected for as in the traditional SDL. For this experiment, the control is the *PTEN* expressing colony on the same plate as the *PTEN* variant, so no separate control plates were generated. Colony sizes were determined using the analysis software Balony (Young and Loewen, 2013).

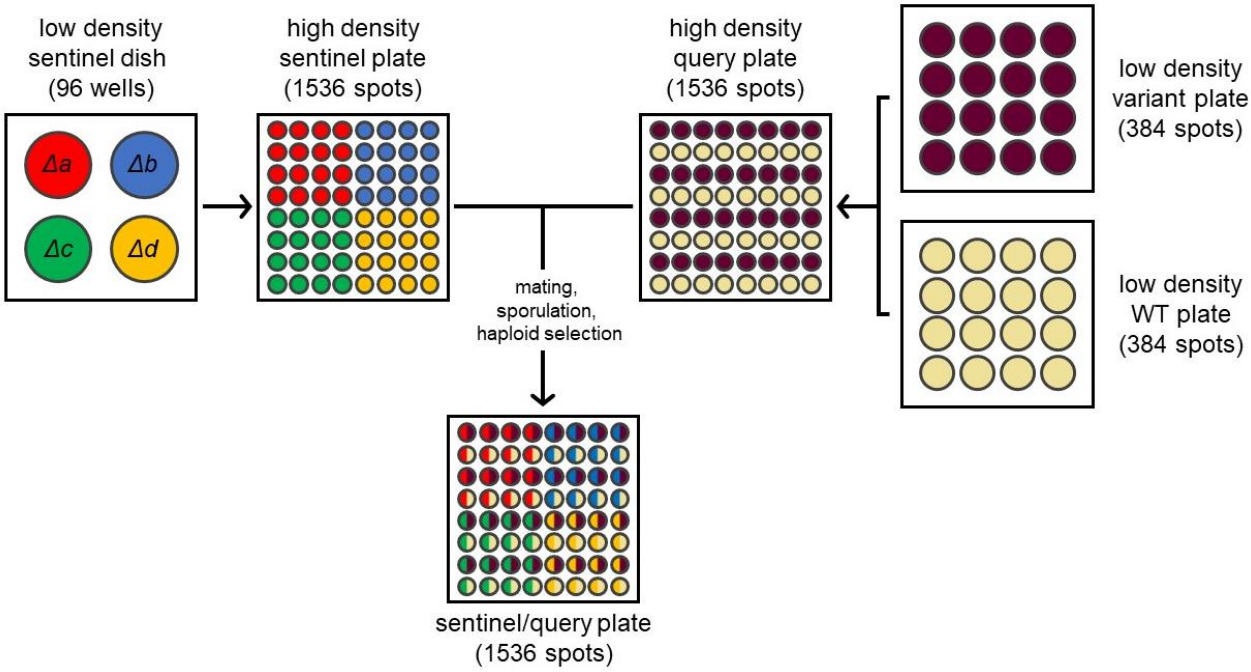

Figure 2.5: Technical design of the first generation mini array

A sentinel plate was created at 1536 density in which each of the 44 sentinels were expressed in a two separate 4x4 blocks on the plate. A query plate was created for each variant tested with alternating rows of variant and PTEN. The sentinel plate was mated with each of the query plates. Following sporulation and haploid selection, gene expression was induced via incubation on SC-Histidine/Arginine/Lysine/Uracil +Canavanine/Thialysine/G418/Galactose/Raffinose plates.

2.3.3.3 Select sentinel strains indicate reduced PTEN function in ASD-associated variants

Following the completion of the first generation of the mini-array screen, we reanalyzed the value of the inclusion of all the sentinel strains. The benefit of this initial screen was that it gave insight into which variants induce a change in genetic interaction as well as which sentinel strains reflected changes in interactions consistently. Many of the sentinels which were initially chosen due to meeting the 'low-hit' criteria or involvement in phosphoinositide-related functions

were found to not have a strong enough difference between genetic interaction with wildtype *PTEN* and any of the variants to be able to identify variants with loss of function. These were initially included because we were unsure if particular variants would have novel genetic interactions compared to wildtype PTEN. This proved not to be the case. Therefore, it was deemed unnecessary to include all of them in the succeeding iteration of the mini-array experiment. Unfortunately, some sentinel strains such as *Δvma8* and *Δmon1* had too much of a growth deficit on their own making colonies of these strains less likely to make it to the end of the experiment. Therefore, these strains had to be eliminated from consideration for further mini-array designs. There were ten sentinel strains which were reliably reporting on changes in PTEN function (Figure 2.6). Other sentinels, while initially promising, did not reliably detect differences between variants and wildtype PTEN and were therefore not useful for characterization.

To analyze the ability of the sentinels to indicate a change in PTEN function, a difference score representing PTEN genetic interaction was calculated for each PTEN variant for each sentinel in this screen. For difference score, the difference between the variant colony size and the wildtype PTEN colony size immediately next to it was calculated. This gave a difference value where a positive number indicates a variant which reduces the genetic interaction (Figure 2.6). The higher the value the greater the difference in genetic interaction between wildtype *PTEN* and the variant. A two-tailed t-test was performed on the colony sizes of each variant in each sentinel compared to the PTEN colony sizes in each sentinel found on the corresponding variant plate. The $\log_{10}$(p-value) (Figure 2.6) indicates the magnitude of the significance of the difference between the variant and *PTEN* per sentinel. From the sentinels which showed the

greatest changes in genetic interaction between *PTEN* and the variants, the eight sentinels with the highest significance were selected for the next generation mini array.

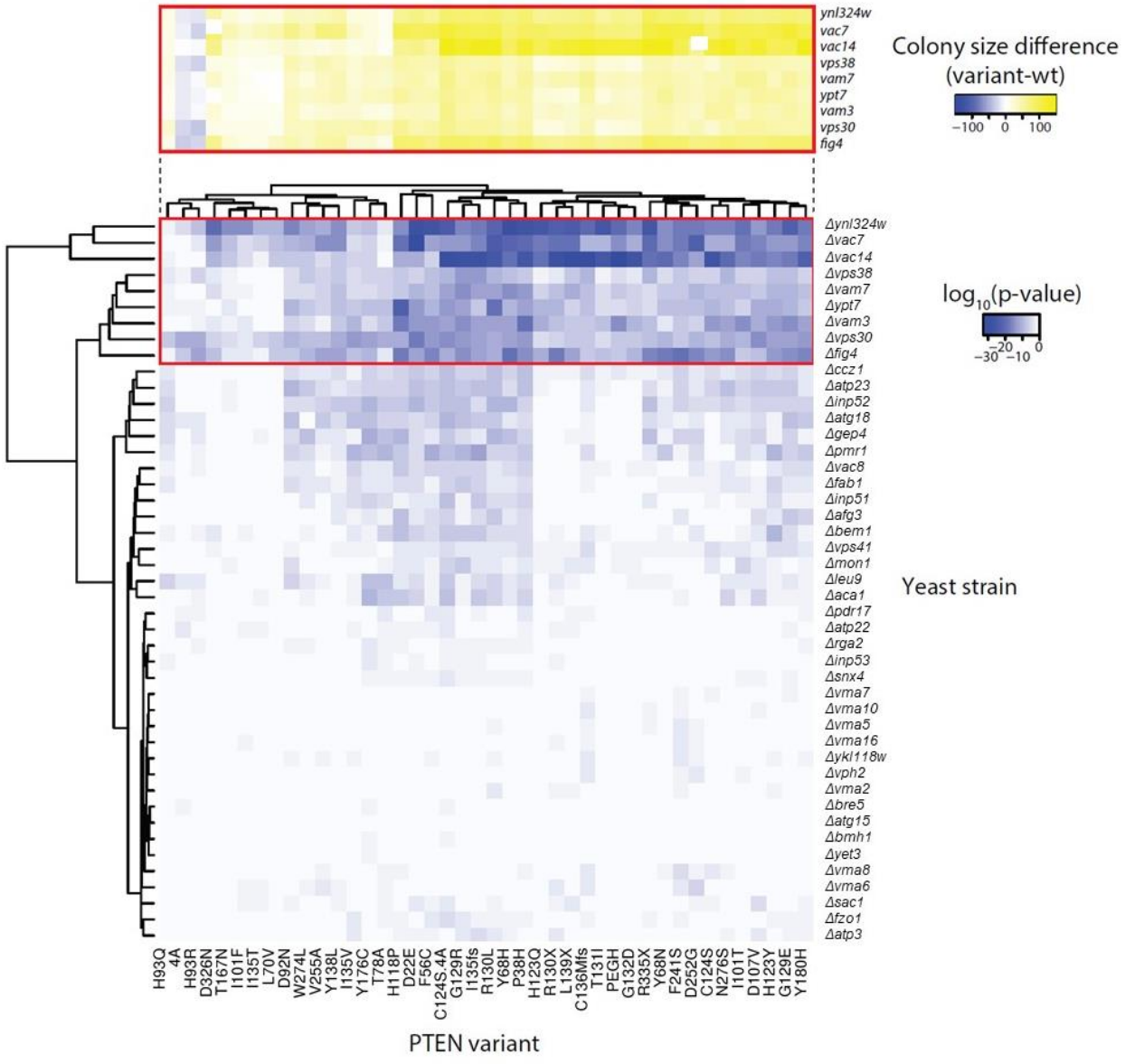

Figure 2.6: Results from the first generation mini array

Results from the first generation mini array in heatmaps generated via hierarchical clustering using average linkage and the Euclidean distance measurement method on heatmapper.ca. The difference between variant and wildtype

PTEN colony sizes are displayed in the upper heatmap while the significance is determined via two-way t-tests and $\log_{10}$(p-values) are displayed below.

2.3.4 Second generation mini array

2.3.4.1 Refined sentinel selection

Insights from the first generation mini array allowed for prioritization and refinement of sentinel selection. Including a smaller subset of sentinels in the mini array allows for an increase in sample size as well as an ability to test many more variants at a time. Because many of the sentinels from the initial mini array design yielded little difference between wildtype *PTEN* and variant expression, it was unnecessary to include these sentinels in the next generation of the mini array. Therefore, after careful consideration, eight high confidence sentinel hits were selected for inclusion. These eight sentinels (Figure 2.7) were selected because of their replicable, significant genetic interaction with *PTEN* and a reduction or complete loss of this interaction in pEGH or *C124S*. Some of these hits were high confidence hits from the initial SDL screens, however a few were previously identified as low confidence hits.

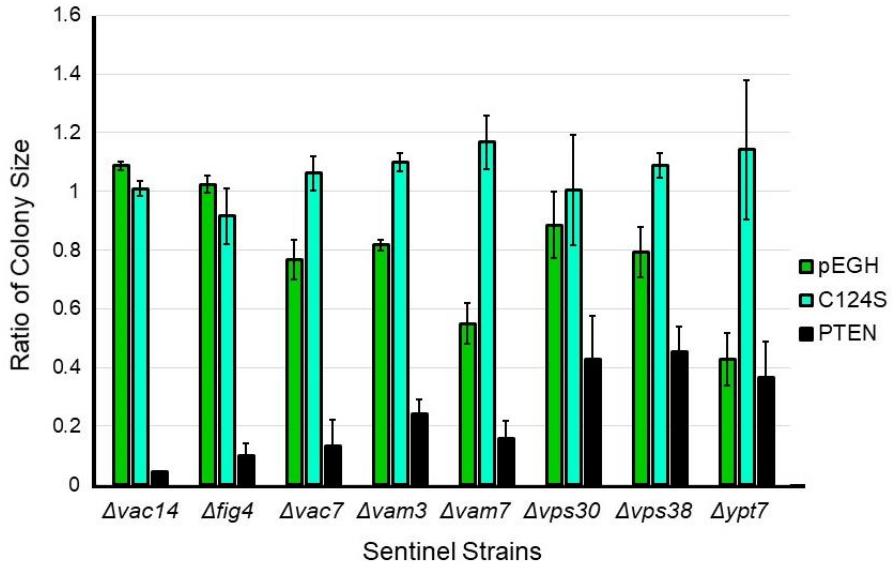

Figure 2.7: Colony growth of sentinels selected for second generation mini array

The ratio of colony size for each sentinel selected for the second generation mini array expressing either pEGH, C124S, or PTEN was generated from the initial SDL screens. The ratios shown were generated by taking the ratio of median colony size of the experimental condition (expressing the query plasmid) over the median colony size of the control condition (not expressing the query).

Interestingly, five of the eight sentinel genes (*VPS38*, *VPS30*, *FIG4*, *VAC7* and *VAC14*) are directly involved in phosphatidylinositide metabolism (Figure 2.8) indicating a potential function of PTEN in yeast. As previously described, Vac7 and Vac14 physically interact and regulate the synthesis of $PI(3,5)P_2$ by activating FAB1, a phosphatidylinositol 3-phosphate 5-kinase (Gary et al., 2002; Rudge et al., 2004). Fig4 is localized to the vacuolar membrane by Vac14 and dephosphorylates $PI(3,5)P_2$ to form PI3P (Gary et al., 2002; Rudge et al., 2004). Given these functions, the lack of these proteins in the deletion strains likely results in altered metabolism of $PI(3,5)P_2$ and PI(3)P. Because we identified *FIG4* and not *FAB1*, and since the VAC7/VAC14 complex plays roles in regulation of both enzymes, we hypothesize that PTEN

antagonizes already compromised levels of PI(3)P in these mutants by its lipid phosphatase activity thereby further reducing levels of PI(3)P.

Two of the other sentinel strains further support this hypothesis. Vps30 and Vps38 exist in the PI3K complex which synthesizes PI(3)P and is essential for autophagy and vacuolar protein sorting (Kihara et al., 2001). The kinase of this complex, Vps34, is an essential protein and may be important for transport vesicles maintaining appropriate membrane curvature (Daum et al., 1998). Of the last three sentinel genes selected for the second generation mini array, *VAM3* and *VAM7* function together as t-SNARE proteins known to directly interact and bind PI(3)P and be involved in vacuole fusion (Sato et al., 1998; Boeddinghaus et al., 2002). The last sentinel strain, *YPT7*, encodes a Rab GTPase also involved in vacuole fusion (Schimmoller and Riezman, 1993). Ypt7 interacts with the Mon1 and Ccz1 members of the class C Vps HOPS complex for tethering of vesicles prior to the formation of SNARE protein complex (Klionsky, 2005; Yang and Klinosky, 2009). Interestingly, Vam7 directly interacts with Ypt7 to mediate docking of vesicles (Ungermann et al., 2000). Ypt7 also negatively regulates Ivy1 (Numrich et al., 2015), a phospholipid-binding protein, which in turn inhibits Fab1 kinase activity (Malia et al, 2018).

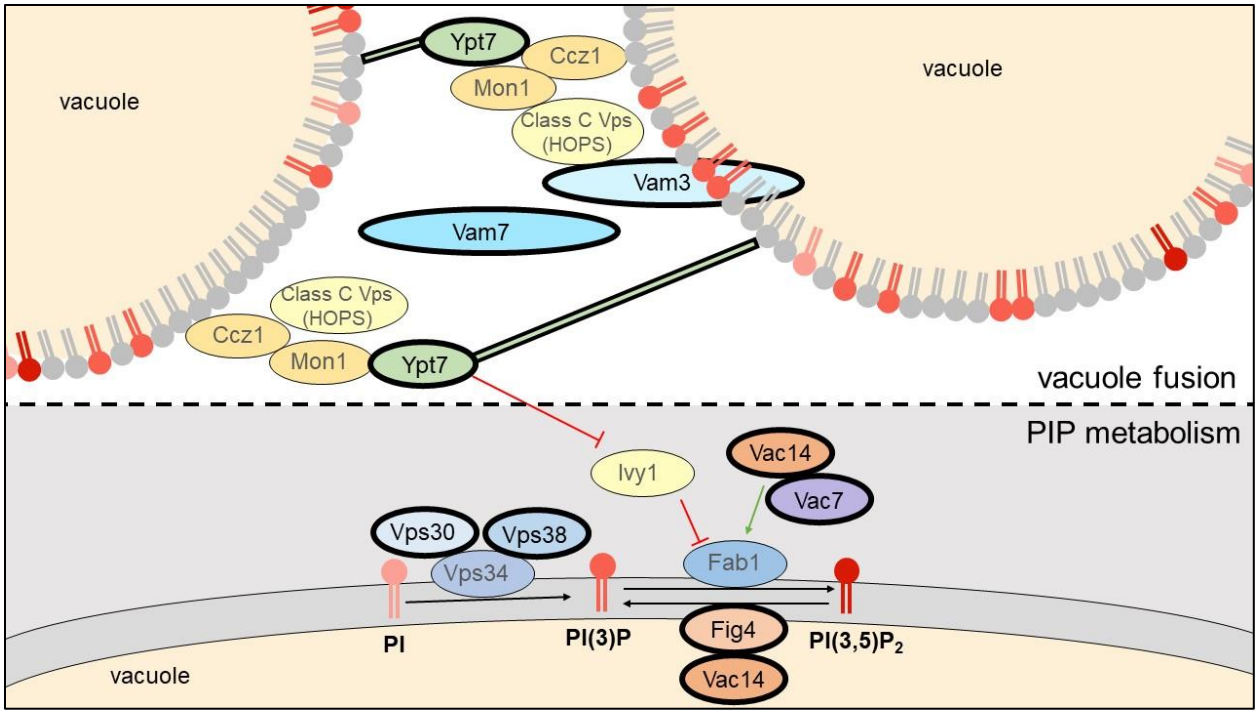

Figure 2.8: Functions of the second generation mini array sentinel proteins

A cartoon depicting the role of the eight sentinel proteins (Vac14, Fig4, Vac7, Vam3, Vam7, Vps30, Vps38, and Ypt7) in yeast vacuole fusion and PIP metabolism.

2.3.4.2 Technical design of second generation mini-array

The genetic interactions of the 97 PTEN variants with the sentinel strains were determined concurrently by the creation of a second generation mini array design. A cartoon depiction of the design can be found in Figure 2.9. For this mini array, one query plate at 96 colony density was created for each variant, wildtype *PTEN*, and empty pEGH vector. A set of master query plates were created at 1536 density which each contained seven variants and pEGH with rows of *PTEN* in alternating patterns such that each variant or pEGH spot is paired with a PTEN spot directly beneath. A novel sentinel deletion plate was created with the eight sentinels such that each sentinel is represented in 4x4 colonies 12 times on the plate.

In similarity with the original SDL screens and the first generation mini array, the master query and sentinel plates were mated then underwent sporulation under enriched sporulation conditions at 25℃ for 14 days. Haploid cells expressing the query plasmids and deletion strains were selected for as in the traditional SDL. For this experiment, there are two important in-plate controls. One is the wildtype *PTEN* expressing colony adjacent to each variant colony and the other is the empty pEGH vector colonies found on each plate. Similar to the first generation mini array screen, no separate control plates were generated. Colony sizes were determined using the analysis software Balony (Young and Loewen, 2013).

Similar to analysis in the first generation mini array screen, a functional score representing PTEN activity was calculated for each *PTEN* variant for each sentinel in this screen. However, in this screen we were able to compare the colony size of variants to *PTEN* and pEGH grown on the same plate. To calculate this, the variant colony size is subtracted from the *PTEN* colony size immediately next to it. This median of these values is divided by the median value of the size of the pEGH colonies subtracted from the *PTEN* colonies immediately proximal. This functional scores estimate activity level compared to *PTEN* expression such that a value close to 1 signifies, with a high likelihood, an activity level similar to *PTEN*, and a value close to 0 signifies a complete loss of activity and therefore a colony with a growth phenotype more similar to pEGH.

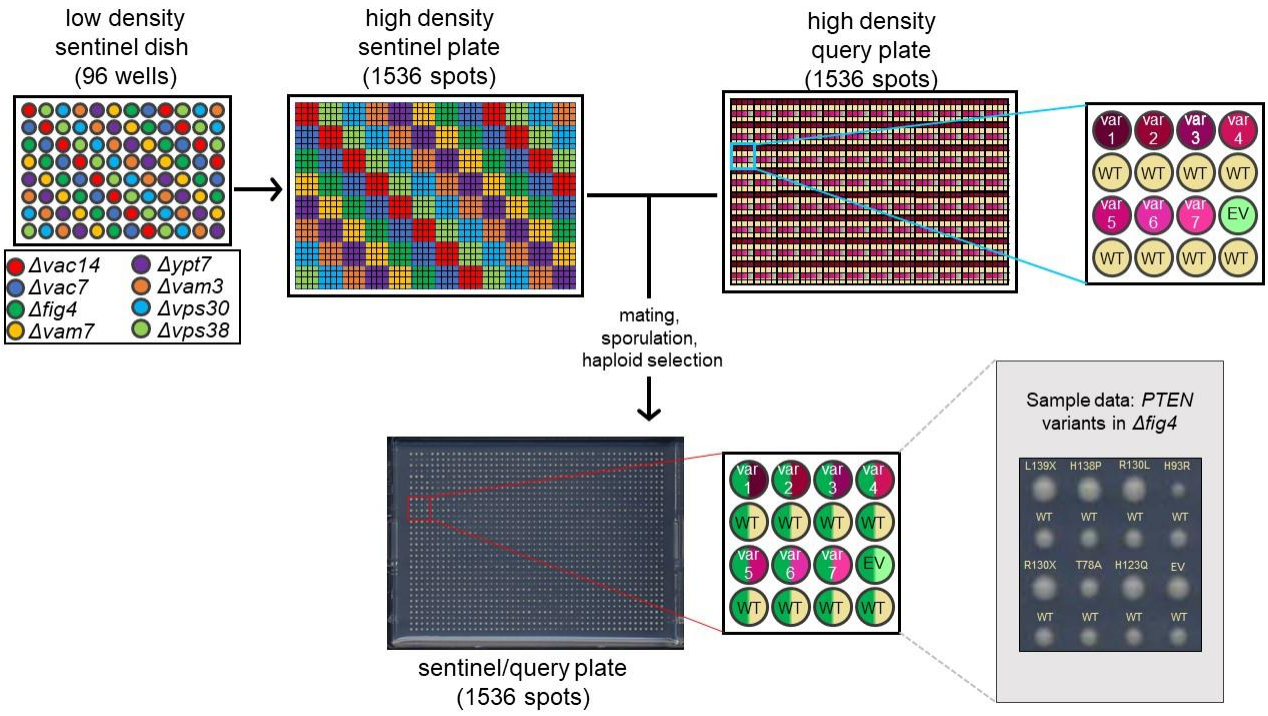

Figure 2.9: Technical design of the second generation mini array

The high density sentinel plate contained 4x4 blocks of each sentinel strain arrayed 12 times across the plate. Query plates contained alternating rows of *PTEN* (WT) such that each variant/*PTEN* or pEGH(EV)/*PTEN* pair is expressed in all of the 12 sentinel replicates. The high density sentinel plate was mated with all of the query plates followed by sporulation and haploid selection. Query gene expression was induced via incubation on SC-Histidine/Arginine/ Lysine/Uracil +Canavanine/Thialysine/G418 plates. A representative sample image depicts colony size of seven variants, pEGH and *PTEN* spots expressed in the sentinel strain *Δfig4*.

2.3.4.3 Understanding *PTEN* variant function through genetic interaction

With the goal of identifying variants of *PTEN* that reduced function, we were able to take advantage of the yeast SDL assay to first determine deletion strains indicative of genetic interactions specific to functional PTEN and then use those strains to screen almost 100 *PTEN* variants for maintenance of these genetic interactions. The eight sentinel strains which were finally selected for our screen are important due to both their genetic interaction with PTEN and

their known functions. Each of them had a range of colony growth between *PTEN* and pEGH expression which allowed us to get a quantitative range of functional scores for each variant instead of an all-or-nothing bimodal distribution (Figure 2.10). The sentinels correlate highly with each other which is depicted in their high correlation with *Δvac14* as indicated by the Spearman correlations (Figure 2.10i-o). The highest correlation between sentinels is between *Δvac14* and *Δfig4*.

As expected, the catalytically inactive C124S and C124S-4A show complete loss of function in all sentinels. G129E, a biochemical control with known loss of PTEN lipid phosphatase activity but retained protein phosphatase activity also showed complete loss of function (Myers et al., 1998). Y138L which has loss of protein phosphatase activity but retains some lipid phosphatase activity (Davidson et al., 2010; Shi et al., 2014) is complete loss of function in *Δvac7, Δvam7, Δvam3, Δvps30, Δvps38,* and *Δypt7* sentinel strains but partially functional in *Δvac14* and *Δfig4* sentinel strains. The constitutively active variant 4A is fully functional in all sentinels as expected. Interestingly, ASD-associated variants display a range of function including WT-like, partial LoF, and complete LoF with a total of 45 of the ASD-associated variants significantly less functional than wildtype PTEN in *Δvac14* . The somatic cancer variants are all significantly less functional than PTEN except I135V which shows loss of function in all sentinels except *Δvps30 and Δvps38* where it exhibits reduced function though not significantly. Interestingly, the majority of PHTS variants are not significantly less functional than wildtype PTEN in any of the sentinels. All of the predicted low impact bioinformatic variants and the majority of the population variants are fully functional in all sentinels while the predicted high impact variants show a range of functionality. This is not completely surprising since bioinformatic tools are known to over-predict pathogenicity (Andersen et al., 2019).

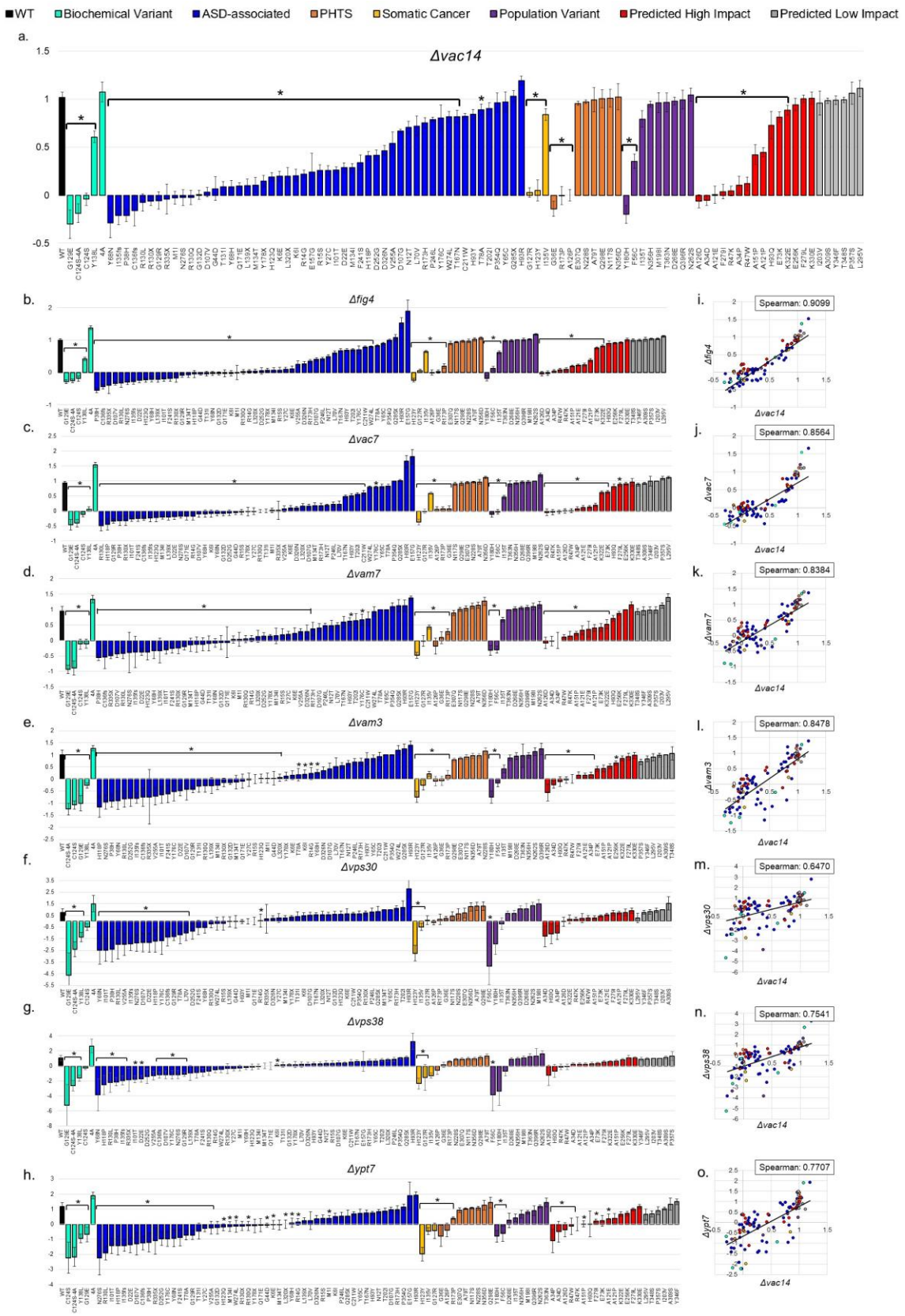

Figure 2.10: PTEN variant functional scores in all sentinels and correlation of all sentinels to *Δvac14*

The functional scores of all PTEN variants in all sentinels is represented (a-h). The correlation of all sentinels to Δvac14 are displayed in the scatterplots (i-o). Variants with significantly different normalized PTEN function compared to their paired wildtype PTEN ($p<0.05$) as determined by t-test are indicated with an *. The strength of correlation is indicated via the Spearman correlations and all Spearman correlations are statistically significant ($p<0.05$, two-tailed t-test).

2.3.5 Detection of the growth rate constant using yeast liquid growth assays

To corroborate the results of our mini array screen, we chose to conduct liquid growth assays to determine the growth rate constant for *PTEN* variants expressed in sentinel yeast. For this assay, the growth rates of yeast expressing variants in *Δvac14* cells were compared to the growth rates of yeast expressing either *PTEN* or pEGH in *Δvac14*. While the size of colonies from the mini array gives an insight into growth rate, it is an indirect proxy for determining yeast fitness. The liquid growth assay gives a quantitative measure of the growth rate constant. By actually measuring the growth rates we can validate our mini array data and get direct measures of growth rates. Our experimental design allowed for the measurement of ten variants, PTEN and pEGH in a single 96-well plate which allowed us to look at over 80 variants in total. By calculating a liquid growth functional score similar to the functional scores calculated from the second generation mini array we could directly compare the two scores for each of the variants tested. As expected, the functional scores determined from the growth rates in the liquid growth assays correlated highly with the functional scores determined from the colony size on agar plates in the second generation mini array (Figure 2.11). As we used *Δvac14* in the liquid growth assay we compared our results to the *Δvac14* mini array results. Our Spearman correlation between the two assays is 0.8815 which indicates a high correlation between the scores.

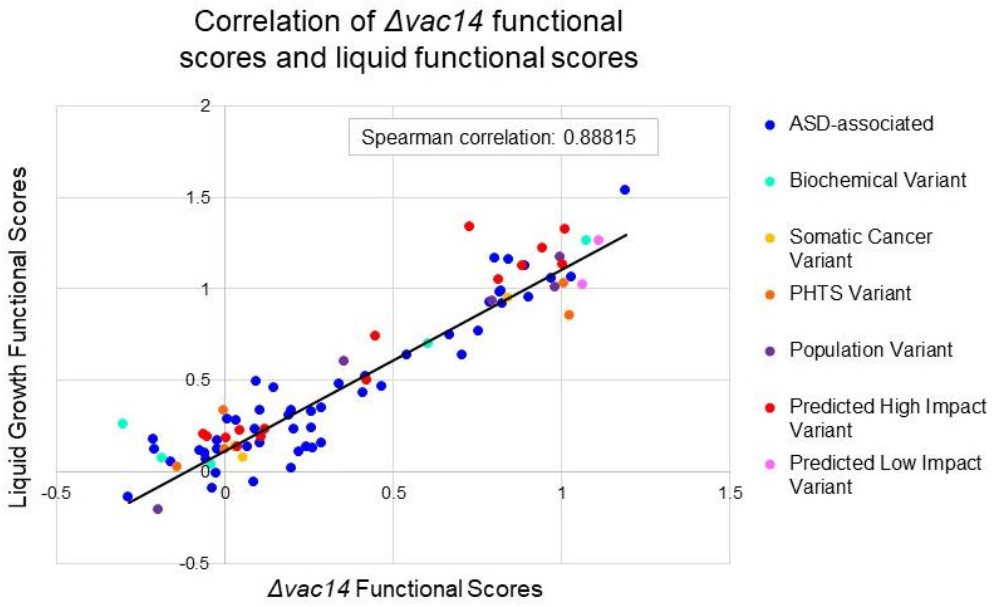

Figure 2.11: Correlation between SDL mini array functional scores and liquid growth functional scores

Functional scores from the *Δvac14* sentinel in the mini array are compared to liquid functional scores also in the *Δvac14* sentinel. The Spearman correlation indicates a high level of correlation between the two different assays. Spearman correlation is statistically significant ($p<0.05$, two-tailed t-test).

2.3.6 Predictive ability of bioinformatic scores

Bioinformatic scores from SNAP2 and CADD attempt to identify variants likely to be pathogenic. Therefore, we compared the mini array *Δvac14* functional scores of all variants to their SNAP2 and CADD phred scores (Figure 2.12a-b). We found a negative correlation between both bioinformatic scores and our functional data with SNAP2 having a slightly stronger correlation (Spearman correlation coefficient=-0.593) than CADD phred (Spearman correlation coefficient=-0.559). We expect a negative correlation because increased deleteriousness is

indicated by larger scores in both SNAP2 and CADD phred while non-functional variants in our assay will have lower functional scores.

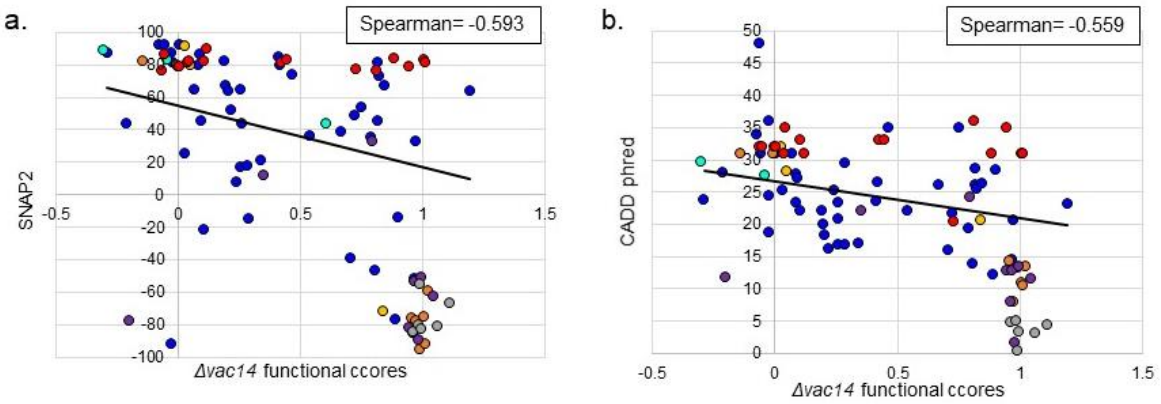

Figure 2.12: *Δvac14* functional scores by bioinformatic scores

Normalized *Δvac14* functional scores data for all variants compared to SNAP2 (a) and CADD phred (b) scores. Spearman correlation coefficients indicate the level of correlation. Spearman correlations are both statistically significant ($p<0.05$, two-tailed t-test).

2.3.7 Stratification of functional analysis by variant classifications

Results from the second generation mini array (Figure 2.10) indicate that variants associated with ASD, PHTS, and somatic cancer cause a reduction in the observed genetic interactions. If we stratify results from the second generation mini array by classification with a human condition (ASD, somatic cancer or PHTS) and compare that to functional scores from population variants we find that variants associated with ASD have significantly lower functional scores than population variants (Figure 2.13a). For this analysis, variants are classified by their main category regardless of whether or not they are also found in additional categories. The mean functional scores for somatic cancer and PHTS variants are not significantly different from each other nor from ASD-associated or population variants. While it appears that variants which

are associated with ASD are less functional than the other categories, it is important to remember the criteria we use to classify. Many of the variants identified in ASD have also been found in somatic cancer and PHTS. It is difficult to draw conclusions based solely on the primary categorization. Therefore, we also stratified ASD-associated variants by comorbidities including macrocephaly and DD but found no significant difference between any of these classifications (Figure 2.13b). We also stratified all variants by number of annotations in the COSMIC databank. In this analysis we found that variants never found in somatic cancer are significantly more functional than variants found greater than eight times in somatic cancer (Figure 2.13c). However, from the variants annotated eight or more times in somatic cancer, there are a few variants that have similar activity levels to wildtype PTEN in this assay. Interestingly, several of these variants are located within the WPD-loop which makes up a portion of the wall of PTEN's active site. Since PTEN is likely acting on a substrate other than PIP_3 in this assay, these variants might not be affecting the binding of this other substrate and therefore are not LoF in this assay.

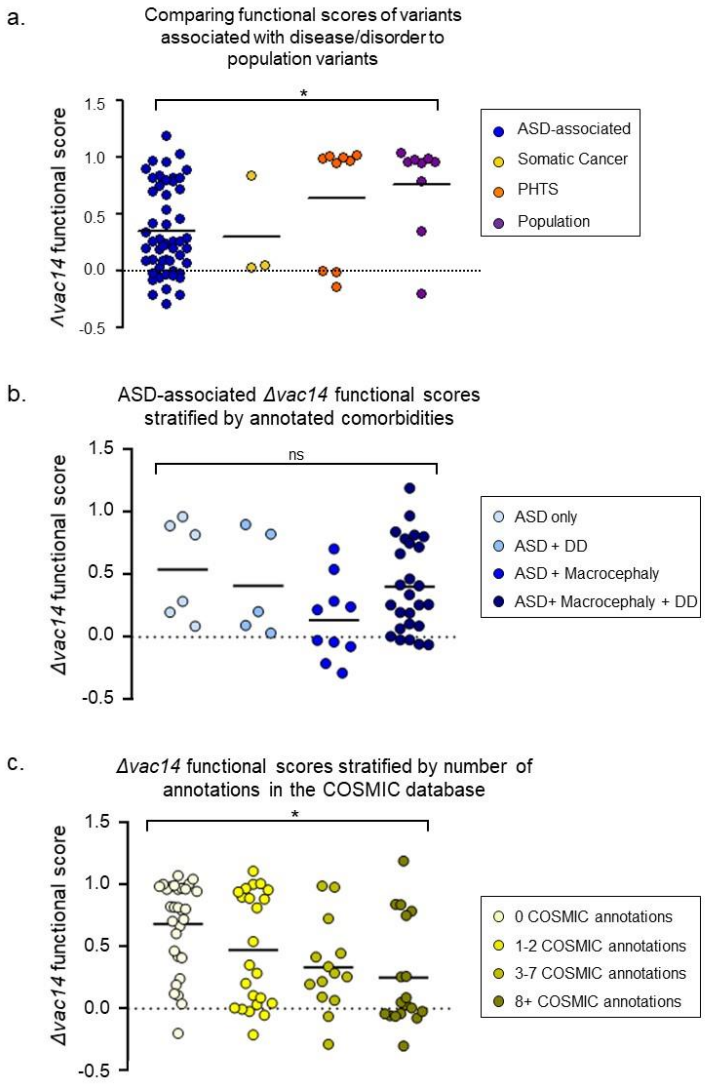

Figure 2.13: Comparison of functional scores to variants with different classifications

Stratification of variants are based on primary classification found functional scores of ASD-associated variants are significantly ($p<0.05$) less than functional scores of population variants as determined by a multiple comparison Kruskal-Wallis test (a). ASD-associated variants stratified by comorbidities found no differences between aby of the classifications as determined a multiple comparison Kruskal-Wallis test (b). All variants stratified by number of COSMIC annotations found functional scores of variants with 0 COSMIC annotations are significantly more functional than variants with greater than 8 COSMIC annotations ($p<0.05$) as determined by a multiple comparison Kruskal-Wallis test ASD-associated variants are significantly ($p<0.05$) less than functional scores of population variants (c).

2.4 Discussion

Determining how *de no*vo mutations impact protein function may give important insight into understanding how these mutations are contributing to brain and human development and identify deficits which could be potential targets for therapeutics. By estimating function through genetic interactions we were able to utilize the well-established, quantitative SDL assay to first identify genetic interactions specific to functional *PTEN,* then test if point mutations impact these interactions. Given that the main catalytic substrate of PTEN is not present in yeast, the identified genetic interactions also give insight into consequences of the point mutations which are not dependent on the presence of PTEN's main catalytic substrate. While this may cause us to miss additional interactions of PTEN which require its phosphatase activity on $PI(3,4,5)P_3$, the identification of sentinels which are either directly or indirectly part of phosphatidylinositide metabolism indicates that PTEN is having a non-trivial impact on this pathway. Furthermore, previous research supports PTEN impacting PIPs in addition to $PI(3,4,5)P_3$. PTEN is known to dephosphorylate $PI(3,4)P_2$ (Malek et al., 2017) and to bind PIPs, such as PI3P, via its C2-domain which is important for the recruitment of cytoplasmic PTEN to the endosomal membrane (Naguib, 2015).

The creation of our novel approach of first finding sentinels and then using those sentinels to parse loss of function due to point mutations allows for a time efficient and economical approach to looking at genetic interactions in variants. For our experiment, it was important to be able to directly compare the growth rates of yeast expressing the variants with the growth rates of yeast expressing wildtype *PTEN* and yeast expressing the pEGH empty vector. In designing our assay thus, we were able to reduce the impact of confounding variables

specific to plate composition, colony location on the plate, environmental conditions, and manufacturing variance of the pads used for all robot pinning steps. This allows us to test our hypothesis that point mutations in *PTEN* associated with disease cause a reduction in the function of the PTEN protein. It also allowed us to create a novel experimental design we could use to test human genes regardless of their function or whether they exist in yeast.

Our initial iteration of the mini array screen was a useful tool to determine if the sentinels would be able to discern subtle differences in function between variants. Many sentinel strains were included in this initial mini array design because it was unclear prior to doing the experiment exactly how many and which sentinel strains would be sensitive to point mutations. Interestingly, only a few of the sentinels seemed to be reporting on genetic interactions that were strong enough to determine the impact of point mutations. Because we were so inclusive in our selection of sentinel strains in our initial screen, it was not an ideal design to screen all variants. We hoped to learn from this first design to create a more optimized mini array. Because only ten sentinels were reporting genetic interactions that were sensitive to point mutations, we were able to redesign and optimize our technique in the second generation mini array. In the creation of our optimized final screen, we kept certain aspects from the first mini array while also prioritizing increased sample size and having pEGH empty vector controls on each plate. Doing so allowed us to compare differences in colony size of the variant, *PTEN*, and pEGH directly.

These assays were a first step towards understanding how point mutations disrupt PTEN protein function. Importantly, these assays allow us to see a full range of function including wildtype-like, reduced function, and loss of function. However, we do not know the direct mechanism that causes this loss of genetic interaction. Additionally, although yeast are an ideal model to study genetic interactions, it is important to be careful not to overinterpret results since

yeast diverged from humans over one billion years ago (Kachroo et al., 2015). Therefore yeast may not fully share relevant PTEN-associated molecular interactions found in humans such as the dephosphorylation of the actin cytoskeleton-organizing protein Drebrin (Kreis et al., 2013; Garcia-Junco-Clemente and Golshani, 2014) or the physical interaction of PTEN and CENP-C (Planchon et al., 2008; Gong et al., 2015). Since ASD is a disorder that is characterized by alterations in brain development as well as social and behavioral deficits, we must be careful not to overinterpret protein-level analysis. Further experiments are required if we are to make conclusions on whether or not these variants are the cause of the developmental and behavioral differences observed in ASD probands.

Chapter 3: Search for mechanisms of PTEN variant dysfunction

3.1 Introduction

In an effort to understand if VUS identified in people with ASD may be conferring risk, we have chosen to investigate the impact of *PTEN* point mutations on several aspects of the protein. In Chapter 2, we began by first identifying genetic interactions of wildtype PTEN then investigated how point mutations change these interactions. However, this only scratched the surface of the possible ways a point mutation may alter a protein. While the sentinels in the second generation mini array may indicate particular functions of PTEN, we wanted to further explore specific aspects of PTEN in the context of a human cell. Therefore, we next explored two hypotheses to determine the mechanisms by which point mutations reduce function.

The first mechanism we chose to investigate was protein stability. Amino acid substitutions are known to impact protein structure (Dill et al., 2008), however, predicting which changes will alter the thermal stability of the protein is challenging despite great progress made in computational models (Pucci et al., 2016). In our genetic interaction assay, if a point mutation renders the protein unstable, it will be look like a LoF variant. However, knowing that a variant's dysfunction is due to a reduction in available protein indicates haploinsufficiency as a potential mechanism for the disruption which could be therapeutically targeted (Montefiori and Nobrega, 2019). Since PTEN haploinsufficiency has been implicated in increasing tumor formation (Kwon et al., 2008) and evidence indicates that point mutations in PTEN disrupt protein stability (Matreyek et al., 2018), we wanted to know if the VUS associated with ASD disrupt protein stability and if incremental reduction in stability correlates with an incremental reduction in function.

In addition to stability, we chose to test how *PTEN* point mutations change the ability of PTEN to regulate the AKT pathway. This pathway has been identified as dysregulated in cancer (Myers et al., 1998; Vivanco and Sawyers, 2002; Martini et al., 2014) as well as brain disorders such as ASD and schizophrenia (Kwon et al., 2006; Chen et al., 2014; Rosina et al., 2019). PTEN is known to negatively regulate AKT signaling via its lipid phosphatase activity towards $PI(3,4,5)P_3$. A reduction in $PI(3,4,5)P_3$ levels in membranes is associated with an increase in activated, pAKT levels which effects cell growth, proliferation, cell metabolism and cell survival (Sansal and Sellers, 2004; Martini et al., 2014; Ersahin et al., 2015). To test both stability and regulation of pAKT in HEK293 cells we used a flow cytometry based approach. Using this method we are able to overexpress fluorescently tagged PTEN as well as utilize antibodies specific for AKT or pAKT in HEK293 cells to acquire single cell resolution measures of fluorescence indicating levels of PTEN, AKT, and pAKT. By determining which variants effect stability, pAKT levels, or both, we hope to understand how ASD-associated variants alter PTEN function as well as gain insight into differences between ASD-associated, PHTS-associated, and somatic cancer associated variants.

3.2 Methods

3.2.1 Cell culture

HEK293 cells (ATCC, CRL-1573) were grown in Dulbecco's Modified Eagle Medium (DMEM) (Gibco, D6046) + 10% fetal bovine serum (FBS) (Atlanta biologicals, cat#S12450) and transfected with X-tremeGENE 9 (Roche, cat #XTG9). For the pAKT/AKT assay cells, cells were washed with DMEM 24 hours after transfection then starved for 19 hours in DMEM without FBS. They were then stimulated for 10 minutes with DMEM + 10% FBS + 10nM insulin (Gibco, cat #12585014). They were treated with Trypsin-EDTA (Gibco, cat #25200072)

for 5 minutes to create a single-cell suspension and then fixed for 10 minutes in 3.2% paraformaldehyde (VWR, cat#CAJT2106-4). Cells were permeabilized with Flow Cytometry Permeabilization/Wash Buffer I (R&D Systems, cat #FC005) and stained with 1:100 of rabbit anti-pAKT (Cell Signaling, S473 cat #9271) and 1:100 of mouse anti-pan-AKT (Cell Signaling, cat #2920) for one hour on ice. Cells were washed and then stained with 1:100 of goat anti-Rabbit IgG-Alexa Fluor 647 (Thermo Fisher Scientific, cat #A21244) and 1:100 of Goat anti-Mouse IgG-Alexa Fluor 405 (Thermo Fisher Scientific , cat #A31553) for one hour on ice. Cells were washed and re-suspended in Flow Cytometry Staining Buffer (R&D Systems, cat #FC001) before loading into an Attune Nxt Flow Cytometer (Thermo Fisher Scientific).

3.2.2 Flowcytometry analysis

Data was recorded on a Attune Nxt Flow Cytometer using VL-1 (Alexa Fluor 405), BL-1 (sfGFP), YL-1 (mtagRFP-T) and RL-1 (Alexa Fluor 647) channels, which were single-stain compensated. FlowJo software was used to select cells via forward scatter-height (FSC-H)/side scatter-height (SSC-H) and single cells were determined using SSC-H/side scatter-area (SSC-A). For pAKT/AKT measures, cells with sfGFP fluorescence above untransfected cells to 100-fold above untransfected (300 arbitrary units (a.u.) to 20,000 a.u.) were selected. For stability measures, cells with higher mtagRFP-T fluorescence than untransfected and sfGFP fluorescence <20,000 a.u. were selected. The median pAKT/AKT levels as determined by (647-background)/(405-background) was calculated for cells with RFP and sfGFP staining and cells without RFP and sfGFP staining in each sample. The difference between cells without RFP and sfGFP staining from cells with RFP and sfGFP staining was measured as the effect measure, normalizing each sample by its in-well untransfected control. Samples were then normalized to WT=1 for each day of experiment. For stability measures, the median ratio of (sfGFP-

background)/(mtagRFP-T–background) was calculated as the relative stability of PTEN compared to its 1:1 transfection control. Samples were then normalized to WT=1 for each day of experiment. Each variant was tested 3-8 times using flow cytometry.

3.2.3 Yeast media and growth conditions

Yeast liquid and agar media was prepared using standard methods (Amberg et al., 2005). Transformations were performed with standard methods. The expression of wild-type human *PTEN* or *PTEN* variants was induced in the Y7093 yeast strain (created through modification of the Y7092 (genotype: can1delta::STE2pr-Sp_his5 lyp1delta his3delta1 leu2delta0 ura3delta0 met15delta0) strain, a gift from Dr. Charles Boone, Toronto, CA) using homologous recombination (Amberg et al., 2005). Successful homologous recombination was verified via plasmid recovery and transformation into *Escherichia coli* followed by sequence validation.

3.2.4 Yeast western blot

Yeast were grown overnight under inducing conditions (2% galactose, 2% raffinose). A_{600} was measured for all samples then each were diluted to an OD of 0.1 and allowed to grow approximately 6-8 hours or until they reached between 0.8-1.0OD. An equivalent of 1.0 OD was collected for each sample. Samples were pelleted and supernatant discarded before being frozen at -20C. After freezing, glass beads and sample buffer containing sodium dodecyl sulfate (SDS) and 2.0% β-mercaptoethanol were added to the pellet then bead bashed for 2 minutes. Samples were then boiled at 90C for 5 minutes and centrifuged before being loaded into a 4-12% Bis-Tris protein gel (Thermo Fisher Scientific, cat #NP0322) and run with MOPS running buffer (Thermo Fisher Scientific, cat #NP0001). Gels were transferred in buffer containing 20% methanol and 0.03% SDS for 2 hours at 4C at 28V onto 0.2μM nitrocellulose membranes (Bio-Rad, cat #1620097). Membranes were blocked in tris-buffered saline and Tween 20 (TBST) (Acros

Organics, cat #AC233360010) containing 5% milk (Bio Basic, cat #NB0669). Membranes were then stained with PTEN antibody (1:1000) (R&D Systems, cat# MAB847,) followed by goat-anti-mouse HRP (Thermo Fisher Scientific, cat#62-6520) and imaged with ChemiDoc MP (Bio-Rad). Membranes were then stripped with a pH 2.2 stripping buffer containing 1.5% glycine, 0.1%SDS and 1% Tween 20 and re-stained for loading using the anti-glucose-6-phosphate dehydrogenase antibody (Millipore, cat #A9521). Blots were again imaged on ChemiDoc MP (Bio-Rad). Band density was determined using ImageJ. All PTEN bands were normalized first to their loading control then to wildtype PTEN measured in each experiment. All variants were tested in triplicate.

3.2.5 HEK293 CRISPR/Cas9 mediated knockout of PTEN

HEK293 cells were co-transfected with sgRNA/Cas9 (SC-400103, Santa Cruz) and 600ng puromycin homology directed repair (HDR) (SC-400103-HDR, Santa Cruz) with XtremeGene9 (#6365779001, Roche) according to manufacturer's instructions. Successful transfection was confirmed by GFP fluorescence. 3 days after selection with 0.6μg/ml puromycin, single cell dilution was performed to obtain 0.8 cells/well of a poly-lysine coated 96 well plate. Single cell colonies from each well were expanded to a 24-well plate before PCR and Western analysis were performed to identify PTEN KO clone. The following primers were used to detect CRISPR event at exon 2 (f: GTTTGATTGCTGCATATTTCA, r: CACACTGTAATGGAATCCAG) and exon 4 (f: CACAGCATAATATGTGTCACA, r: GCACTTTAGTCTTCCTGACAA) of the PTEN gene.

3.2.6 Analysis methods

Values are expressed in graphs as means ± the standard error of the mean (SEM). Multiple t-tests and one-way ANOVA's were performed using Graphpad Prism 8 to determine

difference from 0 or wildtype PTEN respectively. Bonferoni post hoc tests were performed when it was necessary to compare all categories to each other while Dunnett's post hoc test was performed to compare all variants to wildtype PTEN. Heatmaps were generated and normalized using Graphpad Prism 8. Statistical significance was defined as $p<0.05$ and indicated in figures with a (*). Spearman and Pearson correlation coefficients were calculated using Microsoft Excel.

3.3 Results

3.3.1 Stability as a mechanism of dysfunction

To investigate if the *PTEN* point mutations in this study impact protein stability, we first tested PTEN protein abundance in yeast using Western blot analysis (Figure 3.1a). For this assay we determined the protein level of PTEN expressed in the query yeast strain, Y7093. Each variant was analyzed compared to wildtype PTEN and a normalized protein abundance score was calculated comparing the variant to wildtype. In this way, wildtype PTEN protein abundance is normalized to one and a reduction in this ratio indicates a variant with reduced protein expression compared to wildtype PTEN. We found a significant reduction in PTEN protein levels in 32 point mutations, 20 of which have less than 50% abundance compared to WT. Interestingly, almost half (23/48) of the ASD variants tested for stability had significantly reduced levels while only 1/3 somatic cancer variants and 3/15 PHTS variants had significantly reduced PTEN protein levels.

While it was important to understand how point mutations alter the protein abundance in the yeast assay, we also wanted to verify that the protein abundance results from yeast were indicative of PTEN stability in human cells. Therefore, we overexpressed 103 PTEN variants in HEK293 cells to measure stability. For this experiment a sfGFP-PTEN fusion protein was expressed at a 1:1 ratio with mTagRFP-T via the P2A plasmid system (Lo et al., 2015). The

relative level of green-to-red fluorophore for each transfected cell was quantified with flowcytometry. This ratio was normalized to wildtype PTEN expressing cells analyzed from the same experiment. A reduction in the ratio of green-to-red compared to wildtype PTEN green-to-red ratio indicates a reduction in protein stability (Figure 3.1b). Of the 103 variants tested, 76 were significantly less stable than PTEN with 50/103 demonstrating less than 50% stability compared to PTEN. Of the 52 ASD-associated variants tested, the majority (47/52) are significantly less stable than wildtype PTEN in this assay with 32 exhibiting less than 50% stability compared to wildtype PTEN. Five ASD variants were not significantly different from wildtype PTEN: P354Q, R130L, R15S, R14G, and T78A. Interestingly, of these five, P354Q is one of the top three most common variants identified in the normal population, R15S has been found in a case of DD but not ASD, and T78A was found in a high-functioning ASD proband (see Table 2.2). All of the biochemical and 4/5 of the somatic cancer variants show reduced stability while 9/19 PHTS variants have reduced stability with only 4/19 having less than 50% stability compared to wildtype PTEN. Although the yeast and HEK293 stability measures were acquired using different methods, the stability data is well correlated (Figure 3.1c) with a Spearman correlation coefficient of 0.6066 indicating that they are both sensitive to disrupted protein stability.

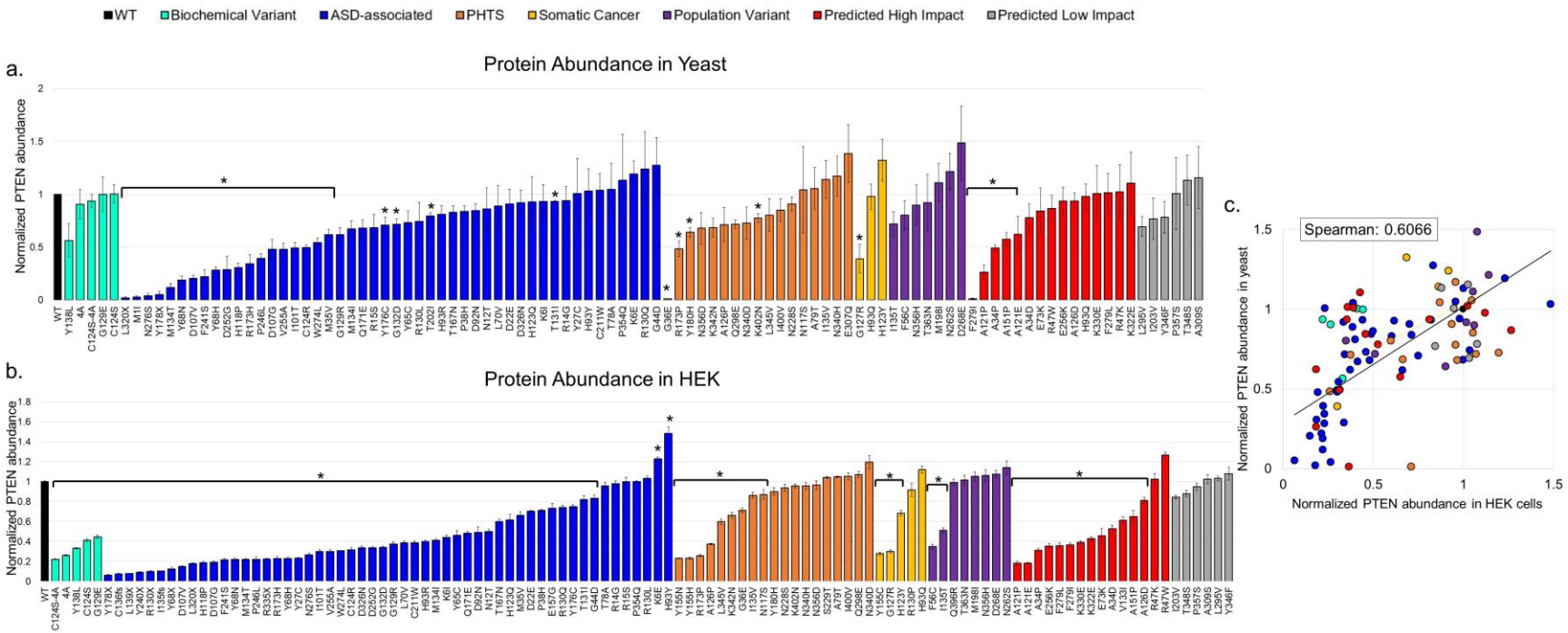

Figure 3.1: Normalized stability of PTEN in yeast and HEK293 cells

All variants tested for stability in yeast (n=3-6) and HEK293 cells (n=3-8) are displayed with SEM in a and b respectively. Variants with significantly different normalized PTEN levels compared to wildtype PTEN ($p<0.05$) as determined by one-way ANOVA with Dunnet's correction are indicated with an *. For each variant, the stability data for both methods is plotted in the scatterplot (c) and the level of correlation between the HEK293 flowcytometry data and the yeast western blot data is indicated by a Spearman correlation coefficient of 0.6066. Spearman correlation is statistically significant ($p<0.05$, two-tailed t-test).

3.3.2 Altered levels of pAKT as a mechanism of dysfunction

PTEN's function as a negative regulator of the AKT pathway has many downstream ramifications. Alterations in this pathway have been implicated in somatic cancers (Myers et al., 1998; Vivanco and Sawyers, 2002; Martini et al., 2014) as well as neurodevelopment (Wang et al., 2017) and severe disruption to this pathway has been linked with more severe cases of ASD (Rosina et al., 2019). We were interested in determining how point mutations of interest effect this pathway. Additionally, we hope to establish if there is a correlation between our yeast mini array results and changes in phosphorylation of AKT. Since yeast do not have the lipid $P(3,4,5)P_3$, and PTEN regulates the AKT signaling pathway through lipid phosphatase activity towards $P(3,4,5)P_3$, it was unclear how well our yeast results would predict dysfunction in this pathway. Therefore, to investigate this question we used HEK293 cells as a human cell model to determine if expression of PTEN variants alters the ratio of phosphorylated AKT to unphosphorylated AKT.

In order to determine how PTEN variants alter AKT and AKT levels, we first synchronized cells in the cell cycle via FBS starvation followed by stimulation with insulin to increase the activity and phosphorylation of AKT (Kim et al., 1999). Cells transfected with tagged PTEN and a transfection control in a P2A multi-gene expression system were fixed and treated with antibodies to determine the levels of AKT and pAKT. Using flow cytometry on these cells allowed us to get single-cell measures of PTEN, AKT, and pAKT levels. For this assay, transfected cells were defined as cells expressing sfGFP and mtag-RFPt. The pAKT/AKT levels from the pools of transfected cells were compared to the pAKT/AKT of in-well untransfected cells. As expected, overexpression of PTEN or the constitutively active 4A variant

causes a decrease in the pAKT/AKT ratio while expression of C124S actually has a slightly but significantly increased ratio of pAKT/AKT compared to untransfected cells (Figure 3.2a-b).

To directly compare all variants as well as compare results from the pAKT assay to the functional scores derived from the yeast mini array we then generated a functional score for wildtype PTEN and each variant normalized to wildtype PTEN (Figure 3.2c). In this way, a score of 1 indicates a variant which is negatively regulating the phosphorylation of AKT similar to wildtype PTEN while a score of 0 indicates a variant whose expression does not impact the phosphorylation of AKT differently compared to untransfected controls. Statistical difference from wildtype PTEN was determined via a one-way ANOVA with Dennett's post hoc test. As expected, the biochemical variants C124S, C124S-4A and G129E have reduced function as does Y138L, although it remains more functional than C124S, C124S-4A and G129E. Also expected, the constitutively active variant 4A has increased function compared to wildtype PTEN. Of the 51 ASD-associated variants tested, 46 have a significantly reduced normalized pAKT/AKT ratio while 4/19 PHTS variants, 5/5 somatic cancer, and 2/8 population variants also have reduced function. All predicted low impact variants show a similar normalized pAKT/AKT ratio to wildtype PTEN while 11/15 predicted high impact variants exhibit significantly reduced function.

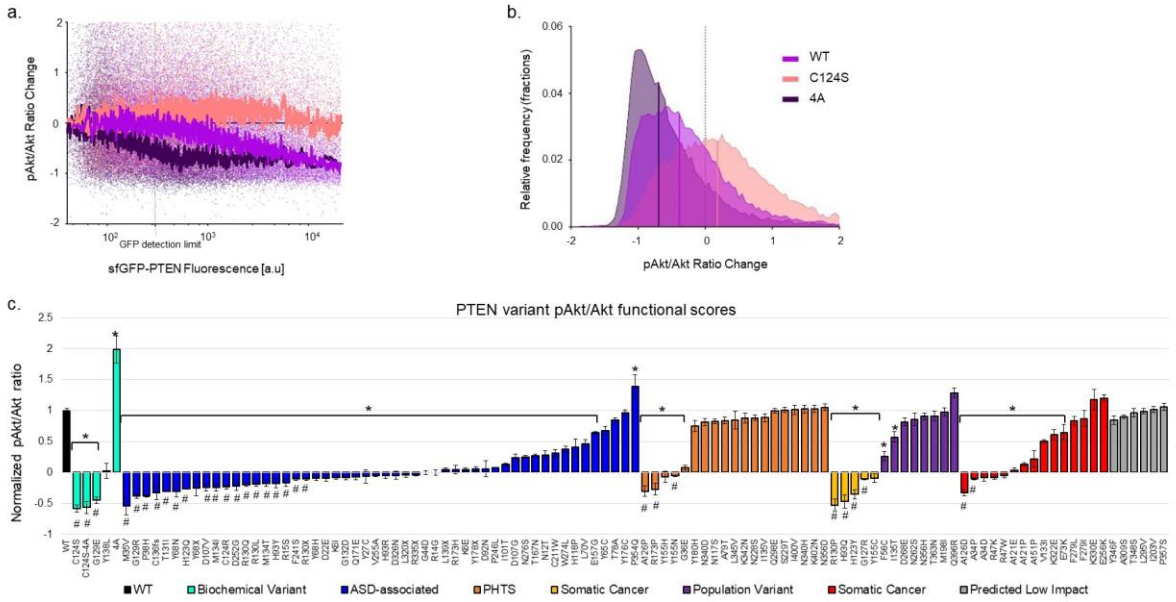

Figure 3.2: Normalized pAKT/AKT ratio of *PTEN* variants

The rolling median of the difference between the pAKT/AKT ratio of HEK293 cells expressing PTEN, C124S or 4A and the in-well untransfected control cells (a) and the corresponding frequency histogram (b). The pAKT/AKT of *PTEN* variants normalized to untransfected control and PTEN pAKT/AKT ratios (c). Variants with significantly different pAKT/AKT ratio to PTEN ($p<0.05$) as determined by one-way ANOVA with Bonferroni correction are indicated with an * while variants with a pAKT/AKT ratio significantly less than 0 as determined by t-test are indicated with a #.

Interestingly, in addition to being significantly less than the normalized wildtype PTEN ratio, many variants also have a normalized pAKT/AKT ratio significantly less than 0 as determined by multiple t-tests. This indicates that the expression of these variants increases pAKT/AKT ratio greater than no variant expression thereby indicating that these variants may have a dominant negative effect on PTEN's regulation of the phosphorylation of AKT. Previous reports have indicated a dominant negative action of several PTEN variants including C124S (Papa et al., 2014; Xu et al., 2018), G129E (Papa et al., 2014; Xu et al., 2018), D252G (Fricano-

Kugler et al., 2018), F241S (Fricano-Kugler et al., 2018), W274L (Fricano-Kugler et al., 2018), and P38S (Matreyek et al., 2018). Although we did not test all of these variants, we found not only C124S, G129E, D252G, and F241S to be significantly less than 0, but an additional 27 variants as well. A total of 18/51 ASD variants indicate significant dominant negativity as do 3/19 PHTS, 4/5 somatic cancer, and 2/15 predicted high impact variants.

To investigate if the indication of dominant negativity observed in the overexpression assay is due to homodimerization with endogenous PTEN, we used a CRISPR-Cas9 approach to create a bi-allelic knockout of PTEN (PTEN-KO) in HEK293 cells. With these new cells, we then repeated the pAKT/AKT assay on those variants which indicated dominant negativity in our initial study as well as wildtype PTEN and the 4A variant. We found that the majority of these variants were no longer significantly less than 0 indicating that they require endogenous PTEN for this effect (Figure 3.3a). Additionally, while the constitutively active 4A variant remained significantly more active than wildtype PTEN in the PTEN-KO cells, its normalized pAKT/AKT ratio was less than in the presence of PTEN indicating that it may enhance endogenous PTEN activity in addition to being constitutively active. We looked at the protein stability of these variants in the PTEN-KO cells (Figure 3.3b) and found many variants to be significantly less stable than wildtype PTEN. A comparison between stability of these variants in HEK293 cells versus PTEN-KO cells indicates a high correlation between the two (Pearson correlation coefficient= 0.939) (Figure 3.3c).

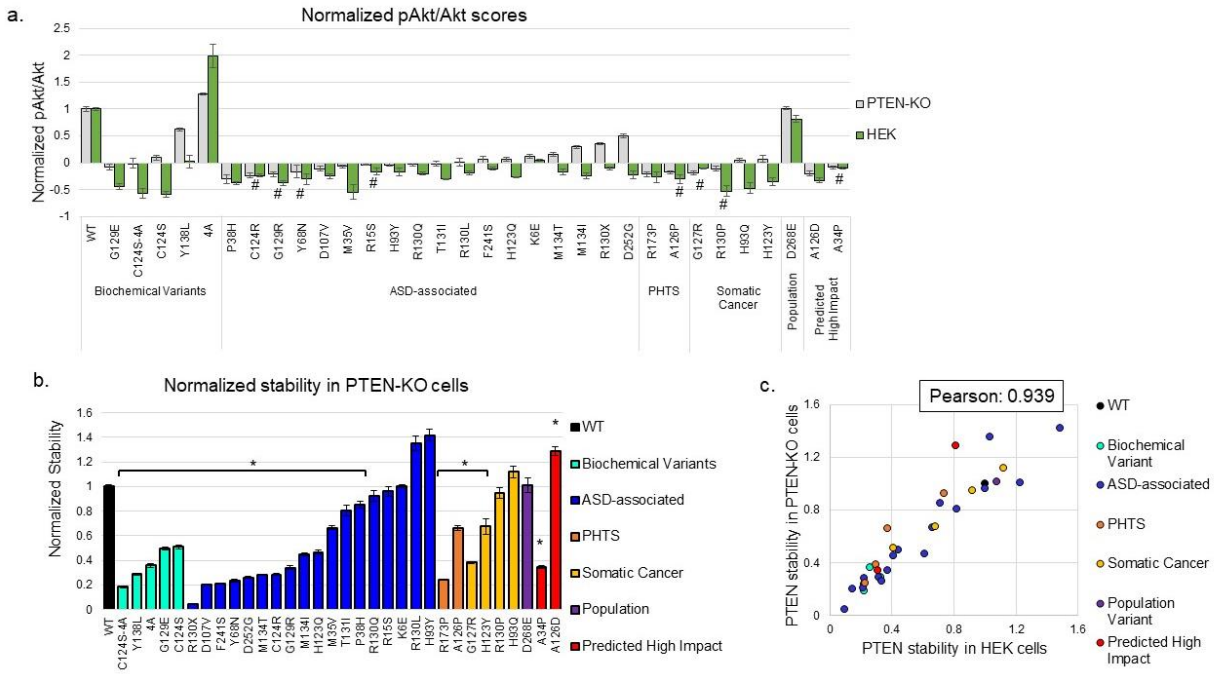

Figure 3.3: *PTEN* overexpression in HEK293 PTEN-KO cells

The pAKT/AKT of *PTEN* variants in HEK293 and HEK293 PTEN-KO cells normalized to untransfected control and wildtype *PTEN* pAKT/AKT ratios (a). Variants with a pAKT/AKT ratio significantly less than 0 in PTEN-KO cells as determined by t-test ($p<0.05$) are indicated with a #. The stability *PTEN* variants in HEK293 PTEN-KO cells normalized to wildtype *PTEN* (b). Variants with significantly different ($p<0.05$) stability compared to wildtype PTEN as determined by one-way ANOVA followed by Dunnett's post-hoc correction are indicated with an *. Comparison between protein stability in HEK293 cells versus PTEN-KO cells yields a high correlation with a Pearson correlation coefficient =0.939 (c). The Pearson correlation is statistically significant ($p<0.05$, two-tailed t-test).

3.3.3 Comparison between stability and pAKT/AKT and genetic interaction results

To determine if the loss of function observed in certain variants in the yeast genetic interaction assay and the HEK293 pAKT assay was due to a loss of stability of the variants, we compared the stability in HEK293 cells and yeast compared to pAKT/AKT function (Figure 3.4b) or *Δvac14* functional score as a representation of the yeast data (Figure 3.4a). If stability

was the main cause of the disfunction observed, we would see a high correlation between the stability data and the functional data. However, we find that the stability data does not explain all of the loss of function observed. Instead, we find that loss of function variants exhibited a range of stability. This indicates that for some variants the loss of function is caused by reduced protein stability while others are dysfunctional despite the protein being present. The correlation between function and stability is low as indicated by the Spearman correlation coefficients from the yeast (0.3698) and HEK293 (0.4059) assays.

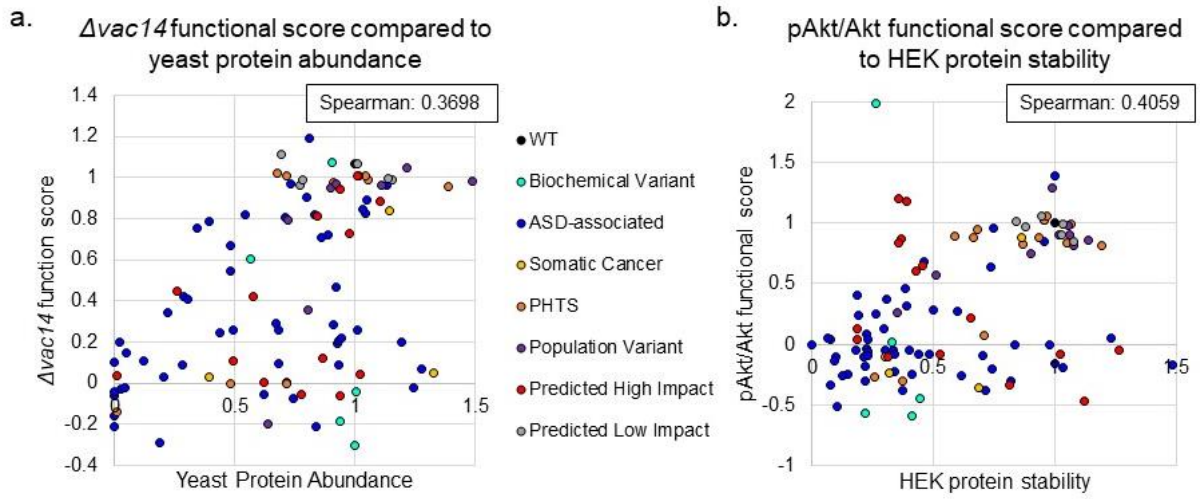

Figure 3.4: Comparison of function and stability

The relationship between protein stability and function in yeast (a) and HEK293 cells (b) indicates a low overall correlation as indicated by the Spearman correlation coefficients of 0.3698 for the yeast data and 0.4059 for the HEK293 data. The Spearman correlations are statistically significant ($p<0.05$, two-tailed t-test).

3.3.4 Comparison between yeast functional scores and pAKT results

From our initial experiments in the yeast mini array assay, we identified variants which caused a reduction in the genetic interactions of PTEN. While some of the sentinel strains

identified in this assay were involved in PIP metabolism, we were interested to determine if any of the sentinels could predict the change in normalized pAKT/AKT ratio ascertained in the HEK293 assay. Interestingly, we found strong correlations between measures of PTEN variant function in the HEK293 pAKT/AKT data set and those from genetic interactions in yeast (Figure 3.5 a-h). The sentinels with the highest correlations to the pAKT/AKT data were *Δvac14*, *Δfig4*, and *Δvac7* with Spearman correlation coefficients >0.72. Knockout of these three genes likely causes alterations to the pools of $PI(3,5)P_2$ and PI(3)P due to their functions either directly or indirectly regulating the generation of these PIPs (Dove et al., 2002; Gary et al., 2002; Rudge et al., 2004). The presence of a genetic interaction between these genes and PTEN indicates that PTEN is likely acting as a lipid phosphatase in this pathway. Although AKT phosphorylation is downstream of PTEN's phosphatase activity on $PI(3,4,5)P_3$, the results from these sentinels may be indicative of the overall lipid phosphatase activity of PTEN.

The functional scores from the sentinels *Δvam3* and *Δvam7* also have a high correlation (Spearman correlation coefficients >0.70) with the pAKT/AKT data. While these two proteins are not involved in lipid metabolism, VAM3 and VAM7 are t-SNARE proteins involved in vacuole fusion and known to bind PI(3)P (Sato et al., 1998; Boeddinghaus et al., 2002). VAM7 is known to be sensitive to levels of available PI(3)P (Boeddinghaus et al., 2002), so it is also likely that PTEN's lipid phosphatase is at least partly responsible for these interactions as well. Since VPS30 and VPS38 belong to a complex homologous to PI3K which synthesizes PI(3)P, it is plausible that these sentinels are reporting on functions of PTEN specific to PI(3)P (Kihara et al., 2001). Finally, YPT7, which has also has a high Spearman correlation coefficient (0.648) with HEK293 pAKT/AKT, is known to influence $PI(3,5)P_2$ synthesis as well as vacuole fusion (Schimmoller and Riezman, 1993; Yang and Klionsky 2009). Therefore, it seems that all eight

sentinels are reacting to the lipid phosphatase activity of PTEN in yeast, although perhaps at different sensitivities, and this is likely indicative of PTEN's overall lipid phosphatase activity including for $PI(3,4,5)P_3$.

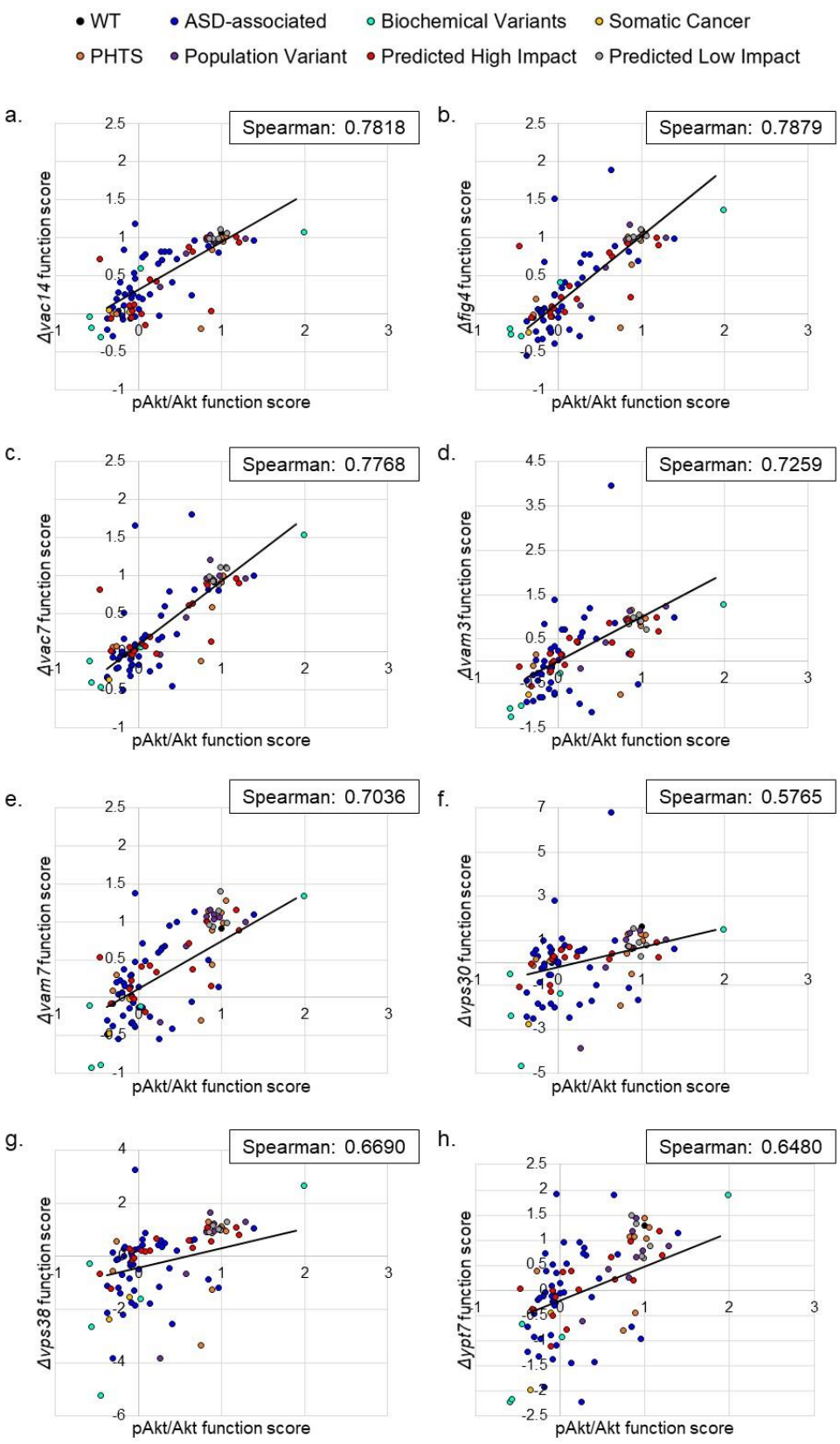
WT
ASD-associated
Biochemical Variants
Somatic Cancer
PHTS
Population Variant
Predicted High Impact
Predicted Low Impact
a.
Spearman: 0.7818
Δvac14 function score
pAkt/Akt function score
b.
Spearman: 0.7879
Δfig4 function score
pAkt/Akt function score
c.
Spearman: 0.7768
Δvac7 function score
pAkt/Akt function score
d.
Spearman: 0.7259
Δvam3 function score
pAkt/Akt function score
e.
Spearman: 0.7036
Δvam7 function score
pAkt/Akt function score
f.
Spearman: 0.5765
Δvps30 function score
pAkt/Akt function score
g.
Spearman: 0.6690
Δvps38 function score
pAkt/Akt function score
h.
Spearman: 0.6480
Δypt7 function score
pAkt/Akt function score

Figure 3.5: HEK293 pAKT function score compared to yeast sentinel function scores

The normalized pAKT/AKT ratio from the HEK293 assay is compared to results from the yeast mini array. Spearman correlations are depicted representing the correlation between each variant's functional score from each sentinel and pAKT/AKT functional score. The Spearman correlations are statistically significant ($p<0.05$, two-tailed t-test).

While the functional scores from sentinel genetic interactions and HEK293 pAKT/AKT ratios are well correlated across all variants, we were also interested to see if particular variants have similar functional scores in the different measures. To determine this, we used a heatmap to depict all functional scores for all variants (Figure 3.6). Because the functional scores have different ranges, they have been normalized for ease of viewing such that all of their values fall between 0 and 1. By graphing the data according to primary classifications, it is again clear that population and predicted low impact variants generally have functional scores similar to wildtype PTEN. The majority of ASD-associated, somatic cancer, and PHTS variants have similar functional scores across the pAKT/AKT and sentinel data, however, there are exceptions to this. For example, Y180H, a PHTS variant, shows loss of function across all sentinels but remains relatively functional in the HEK293 assay while ASD-associated variants H93Y and H93R are dysfunctional in the HEK293 assay but functional in yeast. This is an interesting finding since H93 resides in PTEN's WPD-loop. The WPD-loop is part of the active site of PTEN and proper mobility of loops near the catalytic site are crucial for phosphatase activity. Since the ratio of pAKT/AKT is downstream of PTEN's phosphatase activity towards $PI(3,4,5)P_3$ but yeast do not have $PI(3,4,5)P_3$, it is conceivable that a mutation at H93 alters PTEN's phosphatase activity towards $PI(3,4,5)P_3$ but not other PIPs. These and other variants indicate that specific variants may be more sensitive to different assays.

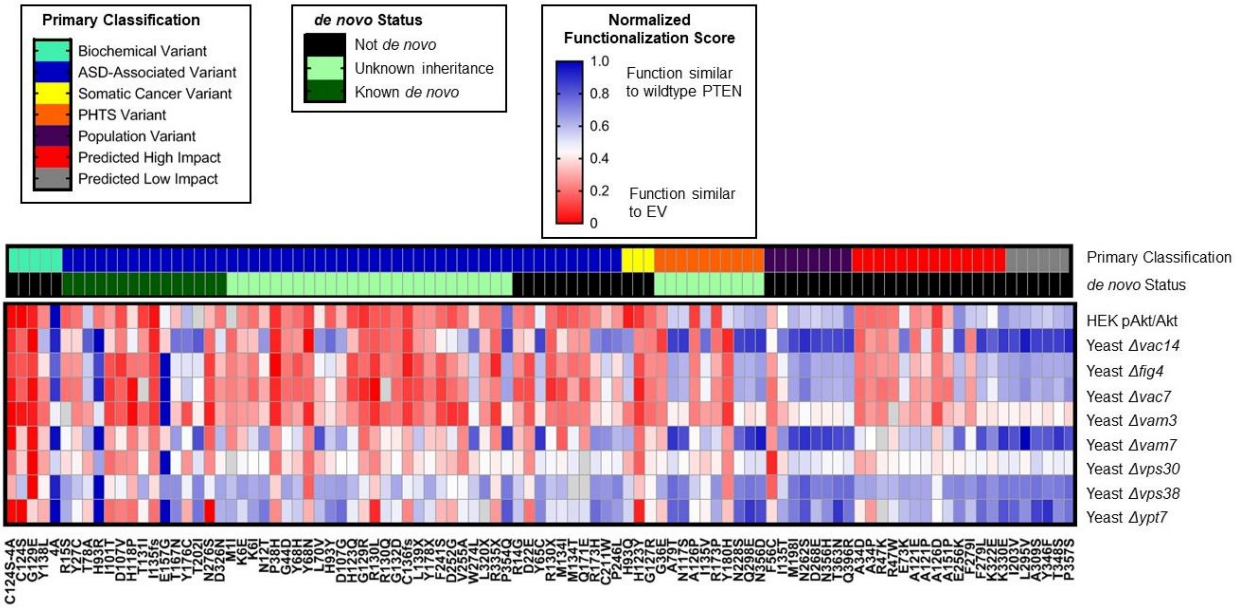

Figure 3.6: Yeast and HEK293 functional scores for all variants

A heatmap representation of the primary classifications, *de novo* status, and normalized functional scores for all variants.

3.3.5 Ability of stability to predict loss of function

We were also interested in determining if there are particular portions of the PTEN protein which are more sensitive to mutation in our assays and whose mechanism of dysfunction cannot be entirely explained by haploinsufficiency. To determine if some variants have a reduction in the regulation of the phosphorylation of AKT in HEK293 cells or a reduction in genetic interactions in yeast not related to reduced protein levels, we generated a stability impact index (SII) for each variant. The SII is calculated by subtracting the normalized stability of each variant (in either HEK293 cells or yeast) from the normalized functional scores (either pAKT/AKT or sentinel interaction functional score). If the reduction in function of a particular

variant is due to a loss of stability, the SII score will be around zero while a variant that is stable but dysfunctional will have a SII score below zero.

We represented SII scores for all variants in the HEK293 pAKT/AKT assay as well as all sentinels used in the yeast mini array in a heatmap arranged by amino acid order (Figure 3.7). We found that the majority of variants studied for which loss of stability does not predict the magnitude of observed loss of function reside in the phosphatase domain of PTEN. Specifically, many of these variants reside in the P-loop (aa 123-130) and the WPD-loop (aa 88-98) as well as the N-terminal domain including the phosphatase binding domain (aa 1-13) through to approximately the first 50 amino acids in the phosphatase domain. By separating these variants into two groups, one including the variants within the first 55 amino acids, P-loop and WPD-loop (depicted in red in Figure 3.7b) and the other containing all other variants (depicted in blue in Figure 3.7b), we are able to see a dramatic difference on the relationship between stability and function (Figure 3.7c). Variants in the P-loop, WPD-loop, and N-terminal of the phosphatase domain show little to no correlation between stability and function, with a Spearman correlation coefficient of 0.001 while the rest of the VUS show a relatively high correlation between stability and function with a Spearman correlation coefficient of 0.62.

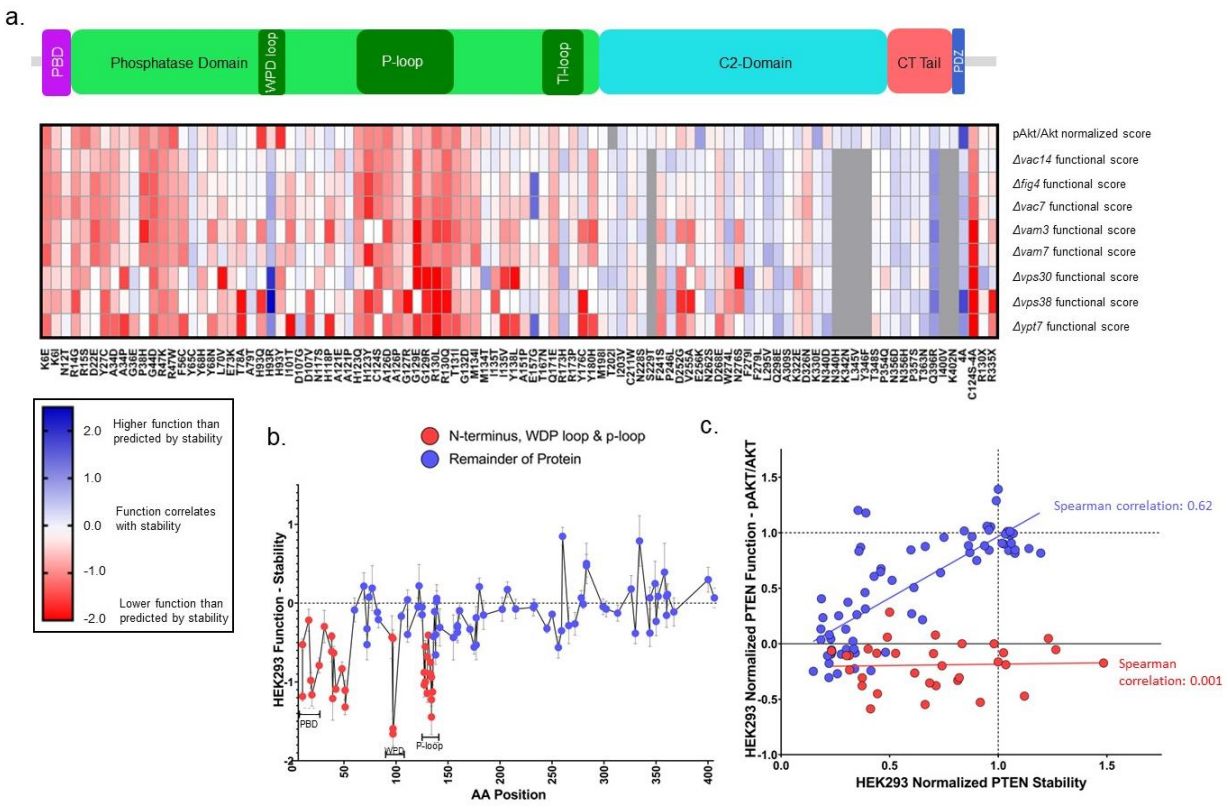

Figure 3.7: Stability impact index for PTEN variants

A normalized plot (a) displaying stability values subtracted from function values for each assay. Variants are displayed according to their amino acid position. A depiction of PTEN from N- to C- terminus is shown above with exaggerated scaling of the functional domains to indicate the variants in this study. For each row, normalized function – stability scores are depicted as a heat map. A score of 0 indicates loss of function explained completely by loss of stability. A positive score indicates higher stability than dysfunction while negative score indicates a variant more dysfunctional than unstable. A scatter plot (b) of HEK293 SII scores for all variants arranged according to amino acid position where variants in the first 55 amino acids, the P-loop, and the WPD-loop are colored red and the rest of the VUS are blue. By then depicting the HEK293 pAKT/AKT function and stability scores for these two groups of variants (c), we find that function and stability are well correlated (Spearman correlation coefficient or 0.062) for the variants not in the N-terminal, P-loop or WPD-loop

3.3.6 Variant results by classifications

In addition to determining if certain portions of the PTEN protein are differentially effected by mutation, we wanted to examine if there is an association between human phenotypes and the deleteriousness of a variant. Previously, in our analysis of the yeast mini array, we found that variants associated with ASD are significantly less functional than population variants (Figure 2.13). Similarly, in our pAKT/AKT assay, we find that variants associated with ASD and somatic cancer are significantly less functional than variants identified in the population (Figure 3.8). We also find that PHTS variants are significantly different from both ASD-associated and somatic cancer variants but not population variants. In the pAKT/AKT assay, the majority of somatic cancer variants are complete loss of function or have function significantly less than 0 while ASD-associated variants have a range of function from similar to wildtype-PTEN to less than 0. A major challenge in interpreting this data is the way we are categorizing variants. Some variants are present in multiple categories, however, for this graph they are only represented in their primary categorization. Therefore, it is necessary to be careful when interpreting these result as it is possible that if we chose different inclusion criteria the conclusions would be slightly different. Additionally, because phenotype data is rare for most ASD-associated and PHTS variants, it is difficult to make conclusions about the correlation of our biological results with patient phenotypes.

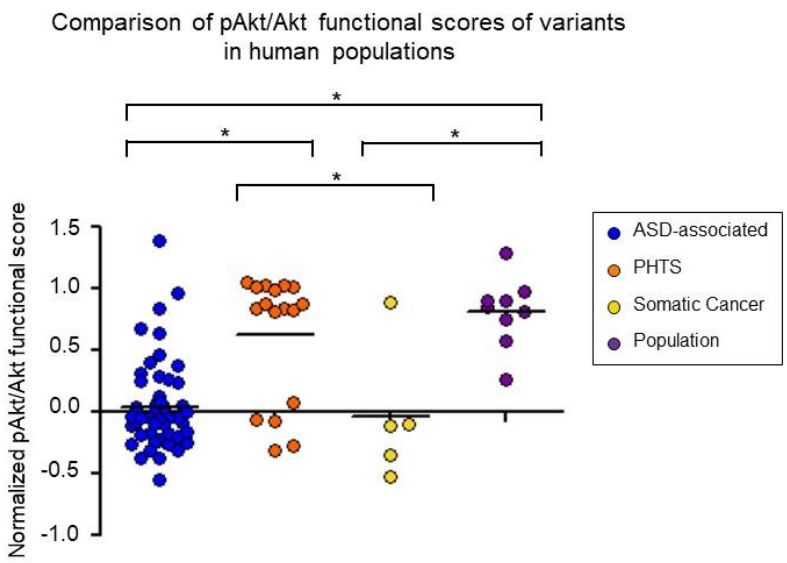

Figure 3.8: Stratification of variants by classification

All variants tested in the pAKT/AKT HEK293 cell assay are grouped based on their primary categorization. A one-way ANOVA with Bonferroni correction was performed and significance is defined as $p<0.05$.

However, some phenotype data is available for many of the ASD-associated variants. The most common comorbidities denoted are macrocephaly and DD. Therefore, we stratified the normalized pAKT/AKT data for all ASD-associated variants based on if they were found in conjunction with macrocephaly, DD, both, or neither (Figure 3.9a) as in Chapter 2. Variants are categorized as ASD alone if they are found in ASD probands for which we have no annotation data or if they were found in people who had neither macrocephaly nor DD. Unfortunately, because variants were not all annotated to a similar degree, it is likely that we do not have all the phenotype data for all variants. A one-way ANOVA with Bonferroni correction found no significant differences between any of the categories. However, with the phenotype data available to us, we find no significant difference between variants identified in people diagnosed

with ASD without macrocephaly and DD and people diagnosed with ASD with ether macrocephaly, DD, or both.

Unlike ASD phenotype data, there is more information for somatic cancer variants which helps to stratify variants found in cancers. For all variants investigated, we collected information on the number of times that particular variant was found in a somatic cancer as detailed by the COSMIC database (Tate et al., 2019). In this way, we are able to get a frequency value for identification in cancer for all variants. If we stratify variants by the number of times they have been identified in somatic cancer, we find that those which have been found in cancer, even only once, are significantly less functional in our HEK293 cell pAKT/AKT assay than those variants never identified in cancer (Figure 3.9b). Additionally, variants identified a total of eight or more times are significantly less functional than those found between one and seven times.

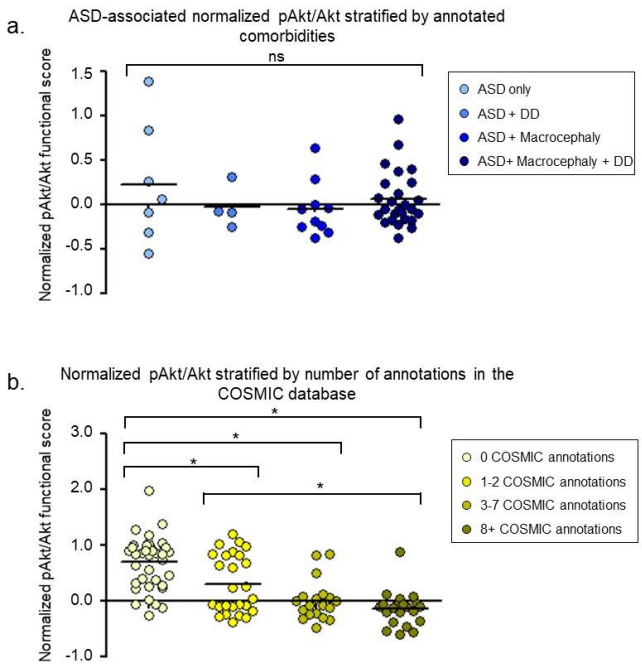

Figure 3.9: Variants stratified by ASD-associated comorbidities and by number of times identified in somatic cancer

All ASD-associated variants are stratified by their comorbidity with macrocephaly, developmental delay (DD), or both (a). A one-way ANOVA with Bonferroni correction was performed but no significance was found between any group (ns). All variants tested in the pAKT/AKT HEK293 cell assay are grouped based on the number of times they have been identified in cancer as determined by the COSMIC database (b). A one-way ANOVA with Bonferroni correction was performed and significance is designated with a (*) and defined as $p<0.05$.

3.3.7 Predictive ability of bioinformatic scores

In addition to phenotype data, bioinformatic scores give indications about the likely pathogenicity or deleteriousness of point mutations. Therefore we compared the normalized pAKT/AKT functional data for all variants to their SNAP2 and CADD phred scores (Figure 3.10 a-b). We found a negative correlation between both SNAP2 and CADD phred (Spearman correlation coefficients of -0.617 and -0.511 respectively) and our data. Our functional scores are

defined such that a value close to one indicates a variant with function similar to wildtype PTEN while a score close to zero is similar to no PTEN overexpression. Since high scores in both SNAP2 and CADD indicate deleterious mutations, we expected to see a negative correlation. Between the two, SNAP2 correlates better with our functional data than CADD phred does. We also compared our HEK293 stability data to SNAP2 and CADD phred scores (Figure 3.10 c-d). We found that CADD phred scores correlate slightly better than SNAP2 scores with the stability data. However, we also found that both bioinformatic scores are better at predicting function than stability. This is not unexpected because while both are trained on data sets that include some measures related to protein structure, their main goal is to identify mutations which alter protein function leading to either non-neutrality or pathogenicity (Bromberg and Rost, 2007; Hecht et al., 2014; Kircher et al., 2014; Reeb et al., 2016).

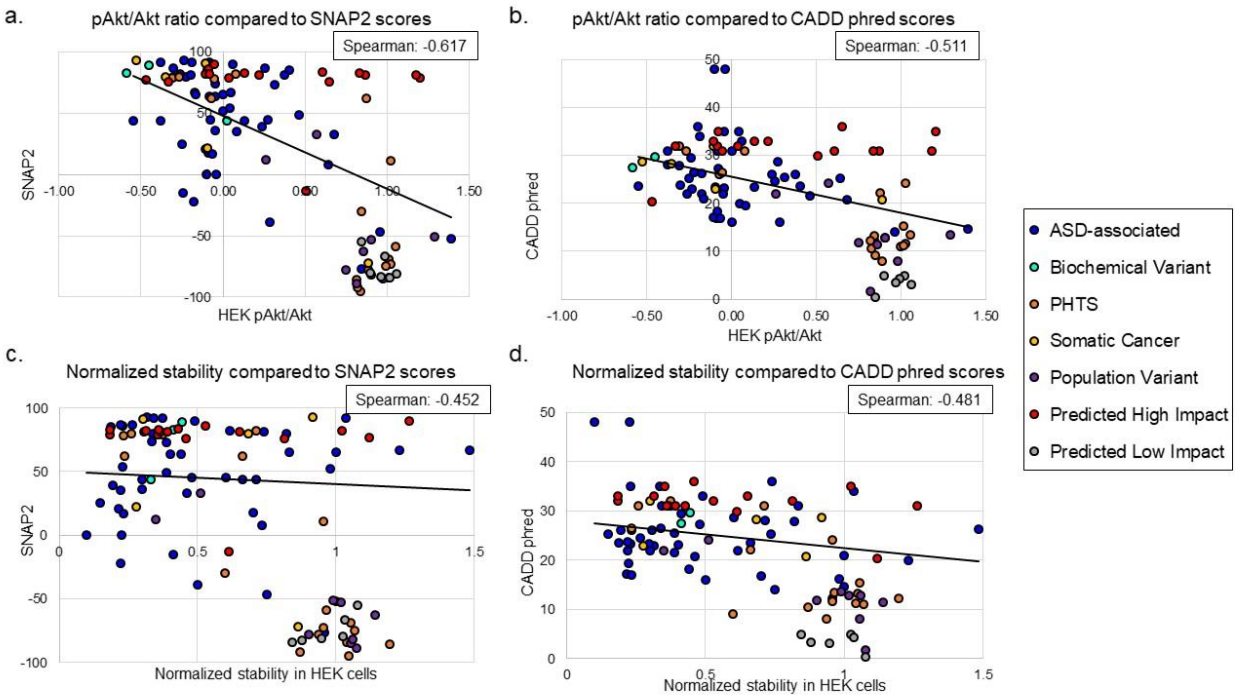

Figure 3.10: HEK293 pAKT/AKT and stability data compared to bioinformatic scores

Normalized HEK293 pAKT/AKT data for all variants compared to SNAP2 (a) and CADD phred (b) scores as well as normalized stability in HEK293 cells for all variants compared to SNAP2 (c) and CADD phred (d) scores. Spearman correlation coefficients indicate the level of correlation. The Spearman correlations are all statistically significant ($p<0.05$, two-tailed t-test).

3.3.8 Pathogenic classification of variants

To identify pathogenic variants, we classified the PTEN variants studied in HEK293 cells and yeast by their level of function in the different assays (Figure 3.11a). In this way it is easy to get an overview of how variants function in the different models. For this heatmap, variants are considered dominant negative (in HEK293 cells) if they are significantly less than zero, LoF if they have less than 25% wildtype function, partial LoF/hypomorphic if they have between 25% and 75% wildtype function, WT if they have greater than 75% function and gain of function (GoF) if they have significantly greater function than wildtype PTEN.

Additionally, recent effort has been made to develop guidelines for interpreting pathogenicity of *PTEN* variants (Mester et al., 2018). Therefore, it is important to relate variants included in our study to the guidelines generated in an effort to relate our data to the overall analysis of *PTEN* variants in the field (Figure 3.11b). For this analysis, we needed to make our data binary. Therefore, if a variant displays less than 50% function it is considered LoF while if it exhibits greater than 50% function it is considered wildtype (WT). According to the guidelines set out by Mester et al., for a variant to be considered pathogenic, it must reach a certain pathogenic criteria. For classifying our variants, we denote a variant as pathogenic if it falls under Mester et al.'s category of strong or very strong evidence for pathogenicity which includes a 50% reduction in function in in-vivo functional studies. Therefore, any null variants as well as

variants with less than 50% function in the HEK293 pAKT/AKT assay and all sentinels are classified as pathogenic. Variants which have less than 50% function in HEK293 pAKT/AKT and the majority of sentinels are classified as likely pathogenic. Variants which are stable and functional in pAKT/AKT and 7/8 sentinels are classified as likely benign. Variants which do not fall into the pathogenic, likely pathogenic, or likely benign criteria are designated as VUS. In classifying the variants this way we found 26 likely benign variants, 18 VUS, 12 likely pathogenic variants, and 35 pathogenic variants.

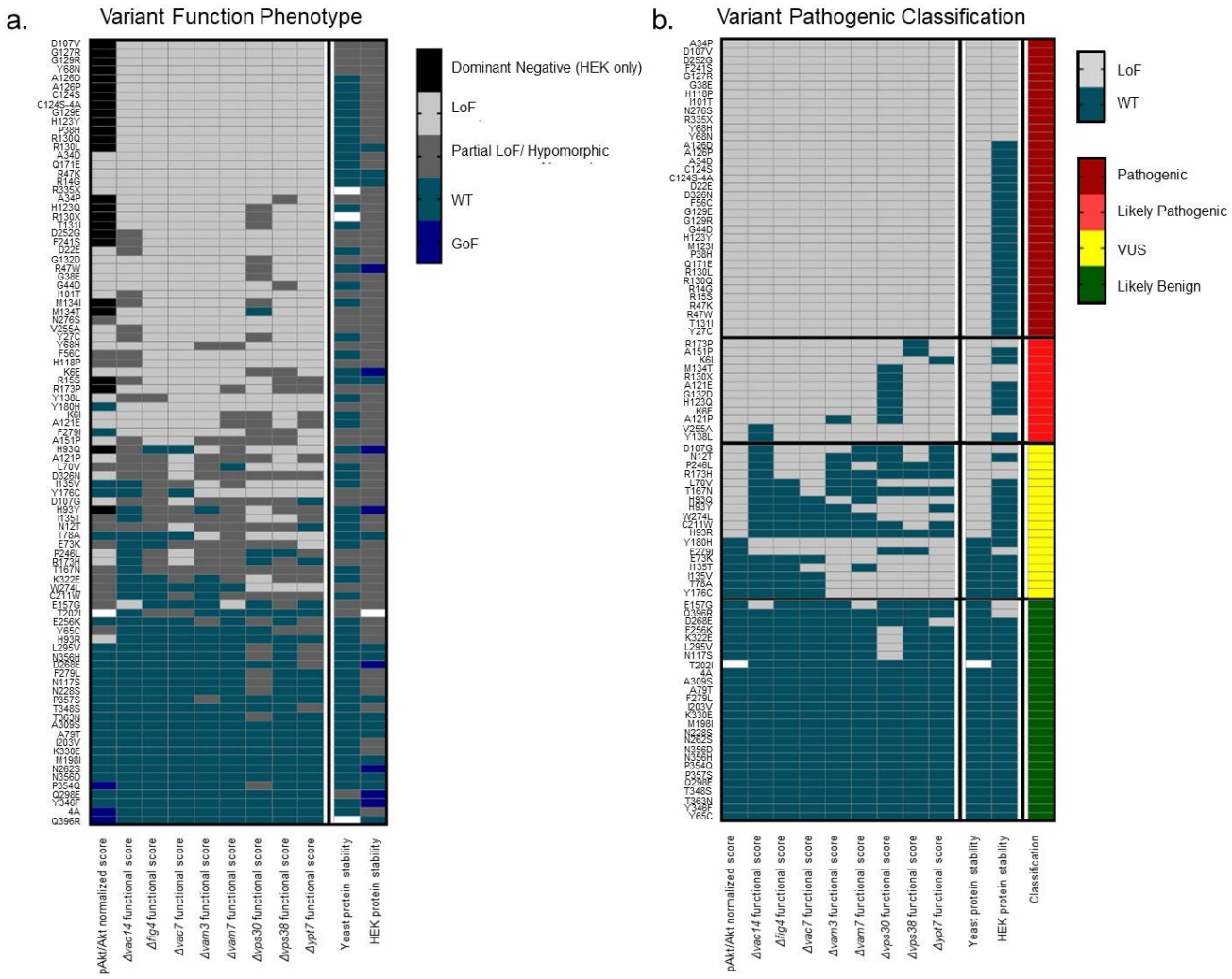

Figure 3.11: Pathogenic classification of variants

The results of HEK293 pAKT/AKT and the eight sentinel functional scores are summarized in (a) and are shaded according to the following criteria: WT-like variants exhibit no differences to wildtype PTEN, gain of function (GoF) variants are greater than WT, Partial LoF / Hypomorphs are greater than null but less than WT, LoF exhibit no differences to zero, Dominant Negative / Antimorphs are less than both null and WT. All differences (greater or less than) stated above imply statistically significant differences of at least $p<0.05$ compared to the wildtype PTEN overexpression effects, null or both in their respective assay. White cells indicate that the variant was not put through that assay. In b., PTEN variant impact is classified based on pathogenicity guidelines (Mester et al., 2018). Each variant is assigned a binary notation of either LoF (<50% of WT effect) or WT-like (≥50% of WT effect).

3.4 Discussion

To detect the ramifications of point mutations in *PTEN*, we tested how VUS impact PTEN protein stability as well as how they alter the ability of PTEN to regulate the AKT pathway. We were able to compare these results to the results obtained in Chapter 2 where we investigated genetic interactions of PTEN and established how point mutations alter these interactions. Using two well-established model systems allowed us to measure molecular function of variants in an effort to understand how point mutations impact various roles of the protein. Understanding how point mutations, including *de novo* mutations found in ASD probands, impact function may impart insight into potential mechanisms leading to altered development or cellular functioning observed in human populations.

Previous advances in studying PTEN variant stability (Matreyek et al., 2018) and function (Rodriguez-Escudero et al., 2005; Rodriguez-Escudero et al., 2011; Mighell et al., 2018) revealed that PTEN is sensitive to point mutations. A consistent outcome of our assessment was the full range of impact on protein function observed including GoF, WT-like, partial and complete LoF, and dominant negative phenotypes. Our findings from the SII calculations recapitulate what has been previously found (Mighell et al., 2018) which recognized that residues across the protein, but especially those in the phosphatase domain near the catalytic P-loop, are sensitive to mutation. However, our analysis adds that not only are these residues sensitive to mutation, but the mutations in these regions cause a disruption to function that cannot be explained by loss of stability alone while dysfunction in residues outside this domain correlate well with the level of instability.

By looking at the stability and function of variants in the same study we were able to identify variants whose dysfunction is caused by instability and variants whose dysfunction is

not predicted by instability. For a large subset of PTEN variants, there was strong correlation between protein instability and dysfunction, resulting in a corresponding range of partial to complete LoF. For these variants, we conclude that reduced protein abundance is the molecular mechanism underlying dysfunction in our assays, and in humans expressing these variants haploinsufficiency likely underlies disease states. These variants are distributed across the entire protein. For other variants, including stable variants, levels of instability did not correlate with protein function, indicating distinct molecular mechanisms underlying dysfunction. For these variants, many of which are within the N-terminal phosphatase domain and active site, dysfunction is not due to protein instability but may be due to reduced binding of substrate perhaps due to a change in protein structure. Alternatively, these mutations could be causing a mis-localization which keeps the protein from reaching it's catalytic target. Additionally, our HEK293 pAKT/AKT assay indicated that a large number of variants are exhibiting dominant negative phenotypes. This adds to the number of previously identified dominant negative PTEN variants (Papa et al., 2014; Fricano-Kugler et al., 2018; Matreyek et al., 2018). Identifying dominant negative variants is clinically important since PTEN mutations are found in a heterozygous genetic context and dominant negativity may produce outcomes distinct from haploinsufficiency. Overall, these results suggest multiple mechanisms of MS mutation-induced PTEN dysfunction: instability producing haploinsufficiency, and loss of substrate binding or catalytic function causing from partial to complete LoF or dominant negativity.

We were also able to compare ASD-associated variants with variants found in other human populations which identified that ASD-associated variants are more unstable than PHTS and somatic cancer variants, especially as determined by Western blot. However, conclusions on the difference between human populations are difficult to draw because so many variants are

found in more than one population. While variants in our study found in people with ASD and comorbidities did not have significantly different function from variants found in people with ASD without comorbidities, it is difficult to draw conclusions from this given the lack of consistent pathology recordings. For further analysis to be done on how severity of disease or disorder correlates with variant dysfunction we need to increase the available human phenotype data.

Chapter 4: Conclusions and future directions

4.1 Overview of findings

4.1.1 PTEN and ASD

Since the recognition of ASD in humans, researchers and clinicians have been redefining and improving our diagnosis and understanding of potential mechanisms of the disorder. The contribution of genetic factors has been repeatedly established (Geschwind, 2011; Woodbury-Smith and Scherer, 2018) and advances in sequencing technologies have begun to identify *de novo* mutations associated with ASD (O'Roak et al., 2011; O'Roak et al., 2012a; Iossifov et al., 2012; Iossifov et al., 2014; de Rubeis et al., 2014; Codina-Sola et al., 2015). A focus has been placed on identifying *de novo* mutations in non-syndromic ASD probands to determine the contribution of non-inherited genetic variations to ASD pathogenesis. In doing so, the high-confidence ASD-associated *PTEN* has been identified to have *de novo*, as well as inherited, missense and LGD mutations (Buxbaum et al, 2007; Herman et al, 2007; Orrico et al, 2009; Varga et al, 2009; McBride et al, 2010; Hobert et al, 2014; Klein et al, 2013; Vanderver et al, 2014). While the presence of a *de novo* mutation does not indicate causality, the prevalence of *de novo* mutations in *PTEN* gives credence to the hypothesis that dysfunction in PTEN may be impacting the pathogenesis of the disorder.

A multi-model collaboration was initiated to investigate if and how point mutations in *PTEN*, as well as other genes, alter protein function, neuron development and animal development and behavior. To examine these *PTEN* variants in multiple cellular environments, a total of five model systems were selected for variant profiling. The overall objective of this thesis was to determine how point mutations of PTEN associated with ASD alter protein function and stability. To do this, I used *Saccharomyces cerevisiae* and HEK293 cells as model systems for

elucidating the range of function observed in VUS found in different human populations. Previously validated variants with well characterized disruptions to PTEN function, referred to as biochemical variants, were used as controls to evaluate function of VUS. The catalytically inactive C124S and C124S-4A as well as the lipid phosphatase dead G129E variants showed loss-of-function in the yeast genetic interaction assay as well as dominant negativity in the HEK293 pAKT/AKT assay. Y138L, the protein phosphatase-inactive variant was a partial loss-of-function in the yeast genetic interaction assay and loss-of-function in the HEK293 pAKT/AKT assay while the 4A, constitutively active variant was functional in the genetic interaction assay and gain-of-function in HEK293 pAKT/AKT. All ASD-associated variants as well as variants identified in somatic cancer, PHTS, and control populations were selected for inclusion in our study in addition to variants classified by bioinformatic tools as predicted high or low impact. In total, these variants are distributed across the entire PTEN protein in all known functional domains.

Using the methods detailed in Chapters 2 and 3, we were able to characterize VUS associated with ASD for their ability to function similar to the wildtype protein. In Chapter 2 we first identified genetic interactions specific to catalytically functional PTEN. Interestingly, when we analyzed the sentinel strains, we found that many of the knockout genes had functions specific to PIP metabolism (Figure 2.8). Despite yeast not containing a homolog of PTEN nor PTEN's main catalytic substrate, $PI(3,4,5)P_3$, this indicated that in yeast, PTEN is likely acting as a lipid phosphatase towards other phospholipids. Of the sentinels not involved in PIP metabolism, the SNARE protein Vam3, is known to bind PI(3)P and be involved in vacuole fusion along with Vam7 (Sato et al., 1998; Boeddinghaus et al., 2002) and Ypt7 (Ungermann et al., 2000). All of these sentinels give indications of what human PTEN is doing in yeast. Our

results add to previous studies which investigated other central nervous system conditions and identified important molecular mechanisms of disorder using yeast as a model system (Pandolfo, 1999; Dhungel et al., 2015; Kryndushkin and Shewmaker, 2011; Abeliovich and Gitler, 2016; Jo et al., 2017).

When we performed our second generation mini array screen with eight sentinels, we found point mutations alter these interactions in a manner that suggests that many of the VUS are neither fully functional nor fully dysfunctional but that their loss of function lies on a continuum. These changes in functionality were highly correlated between sentinel strains (Figure 2.10, Figure 3.6). This indicates that they are likely sensitive to a similar function of PTEN. However, their differences may indicate specific differences of single variants between the sentinels or that different sentinels have different levels of sensitivity. We also found that the bioinformatic tools used to give predictive scores of pathogenicity to each variant are not sufficient for predicting the results of our biological experiment (Figure 2.12). Overall, both SNAP2 and CADD tend to overestimate pathogenicity. While of itself this is not a problem, it may cause some variants to be categorized as likely pathogenic due to bioinformatics, when they do not effect protein function in biological experiments. This is especially important if treatment options are based on causality established by bioinformatic predictions. Finally, when we compared results of ASD-associated variants to somatic cancer associated, PHTS associated, and population variants, we found ASD-associated variants are significantly less functional than population variants (Figure 2.13a).

In Chapter 3 we investigated two additional mechanisms of PTEN dysfunction. We hypothesized that VUS of PTEN may be disrupting protein stability and/or reducing PTEN's ability to negatively regulate the AKT pathway. We found that variants alter PTEN stability and its ability to regulate AKT across the functional range including functional, partial LoF, LoF, and

some even show signs of dominant negativity. When we compare the results from our mini array with results from the pAKT/AKT assay, we find a high level of correlation (Figure 3.5) which further suggests that the sentinels are sensitive to PTEN lipid phosphatase activity.

Through investigation of protein stability, we identified a number of variants whose level of dysfunction can be explained by changes in protein stability (Figure 3.7). Haploinsufficiency has been identified as a mechanism of *PTEN* dysfunction in mice (Clipperton-Allen and Page, 2014; Clipperton-Allen and Page, 2015), and these results further indicate that these variants are likely causing haploinsufficiency in humans expressing them. However, haploinsufficiency does not explain all of our results. Variants in the N-terminal and phosphatase domain of PTEN, especially the catalytic P-loop and WPD-loop, have a reduction in function that cannot be explained by their stability. While mutations in these regions may also lead to protein haploinsufficiency, the indication of dominant negativity in several of these variants signifies that haploinsufficiency is not the only mechanism of *PTEN* dysfunction. Similar to results in the mini array, ASD- and somatic cancer-associated variants are significantly less functional than PHTS associated and population variants (Figure 3.9) in their ability to regulate the phosphorylation of AKT. However, we again have to be careful in the interpretation of these results given the criteria we use to classify variants. Many of the ASD-associated variants are also somatic cancer and/or PHTS associated which we are unable to see when selecting primary categorizations. Because variants exist in multiple human diseases, the classification of variants can be arbitrary and vary between studies. Therefore, due to the lack of clear divisions, looking for differences between PTEN variants associated with ASD, PHTS or somatic cancer can be speculative and problematic. We also found that bioinformatic scores overestimate the pathogenicity of variants for both the regulation of AKT and PTEN stability (Figure 3.10).

Understanding how specific variants alter function has clinical and therapeutic implications and elucidating this will hopefully lead to a push for therapeutics specific to the exact mechanism of disruption.

4.1.2 Use of the multi-model approach

In addition to the experiments in yeast and HEK293 cells, experiments have been completed by other labs in our multi-lab collaboration (detailed in section 1.6) looking at the same variants in different model systems. The overarching goal of this collaboration was to ascertain functional data of *PTEN* VUS in five model systems thus learning the impact of variation on protein function, protein stability, organism development, neuron development, and animal behavior with the added desire to use these same assays to determine the effect of VUS in other ASD-associated genes. Therefore, in addition to the work presented in this thesis an additional effort was made to create a streamlined process for creating reagents for all of the assays. An overview of the molecular biology pipeline designed to address the creation of plasmids can be found in Figure 4.1.

All variants were placed in appropriate vectors for experimentation in yeast, *Drosophila* and HEK293 cells. After initial results from yeast and fly assays, a total of 22 variants were selected for experimentation in the low-throughput, high-resolution rat hippocampal culture and worm behavior assays. For this smaller subset, *PTEN* as well as the *C124S* biochemical variant were selected as controls necessary in all model systems. Additional variants were selected based on variant category, protein stability, and known inheritance status. The three most common population variants, *A79T*, *P354Q*, and *D268E* were also chosen. Prioritization of ASD variants was based on reduction of function as observed in high-throughput assays as well as high protein stability with a preference for *de novo* variants. Selected variants were moved into the worm

destination vector using Gateway cloning. Rat hippocampal experiments were performed with plasmids containing silent PTEN mutations to render the plasmid short hairpin RNA (shRNA)-resistant. A list of all plasmids created for this collaboration can be found in the Appendix.

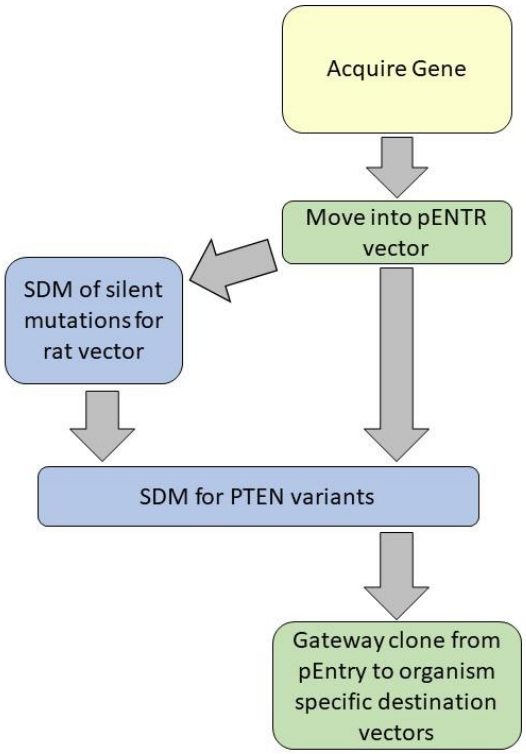

Figure 4.1: Molecular biology pipeline

Representation of the pipeline created to make the various plasmids necessary to test PTEN function in all of the five model systems tested. Following arrival of the gene of interest, it was moved into a Gateway entry vector and site directed mutagenesis (SDM) was initiated. Upon sequence validation, variants were moved into the destination vectors.

Comparison with these other model systems further advances our understanding of the impact of these point mutations on several aspects of PTEN function in varying cellular environments. Briefly, the assays used in fly, worm, and rat hippocampal cultures investigated the ability of *PTEN* variants to recapitulate alterations in development and behavior caused by

overexpression of wildtype *PTEN*. *Drosophila melanogaster* provide another high-throughput model in which we could assess rates of development in an invertebrate model through expression of *PTEN* variants and analysis of the rates of fly eclosion. This is a phenotype sensitive to both PI3K-dependent and –independent pathways (Gao et al 2000). A smaller subset of variants was selected for the lower-throughput rat hippocampal neuron culture and *Caenorhabditis elegans* experiments. In neuron cultures, the effects of *PTEN* variant expression on soma size, neuronal dendritic and axonal growth, and excitatory synaptogenesis were analyzed. These are processes found to be disrupted in ASD (Courchesne et al 2019; Rubenstein and Merzenich 2003) and mediated likely by both lipid and protein phosphatase PTEN activities (Zhang et al., 2012; Fricano-Kugler et al., 2017). Finally, expression of VUS in *Caenorhabditis elegans* allowed for the identification of disordered behavior and sensorimotor processing. Results from chapters 2 and 3 are compared to the fly, rat, and worm results in Figure 4.2.

When we compare the five model systems, we see that a consistent outcome is that we observe a full range of PTEN function in all models including a range of partial LoF. We also observe that the yeast and HEK293 results are predictive of results in some variants, such as C124S and P38H where we found dominant negativity in HEK293 cells, LoF in yeast sentinels, and LoF in the additional model systems. However, the ASD-associated variant H123Q, which is dominant negative in HEK293 cells and LoF in the majority of yeast sentinels, is LoF in the fly development and worm behavioral assay but has function levels similar to wildtype PTEN in all neuron attributes tested. This indicates that neuron development in the rat hippocampal culture assay is less sensitive to this mutation.

Unfortunately, it is difficult to draw broad conclusions with the small number of variants tested in the low-throughput systems. However, the complex pattern of effects observed in the different

measures of chemotaxis behavior and neuronal development suggest multiple functions of PTEN may be involved. Support for this also comes from previous experiments using hippocampal slices where PTEN lipid as well as protein phosphatase activity impacted synaptogenesis (Zhang et al., 2012). While the yeast and HEK293 assays are great for understanding specific protein functions, additional experiments are required to understand variant impact in the context of brain development and ASD progression.

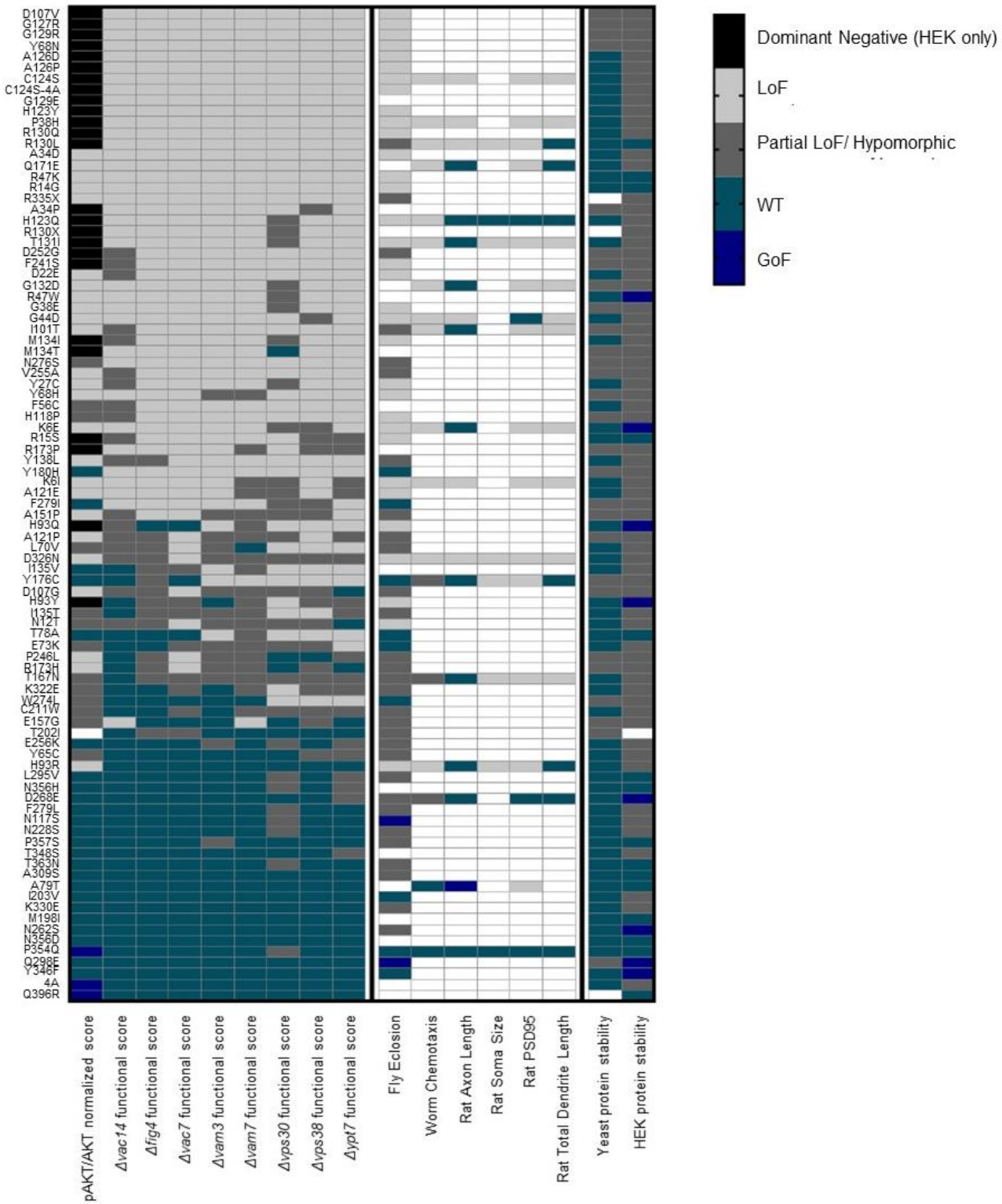
Dominant Negative (HEK only)
LoF
Partial LoF/ Hypomorphic
WT
GoF
D107V
G127R
G129R
Y68N
A126D
A126P
C124S
C124S-4A
G129E
H123Y
P38H
R130Q
R130L
A34D
Q171E
R47K
R14G
R335X
A34P
H123Q
R130X
T131I
D252G
F241S
D22E
G132D
R47W
G38E
G44D
I101T
M134I
M134T
N276S
V255A
Y27C
Y68H
F56C
H118P
K6E
R15S
R173P
Y138L
Y180H
K6I
A121E
F279I
A151P
H93Q
A121P
L70V
D326N
I135V
Y176C
D107G
H93Y
I135T
N12T
T78A
E73K
P246L
R173H
T167N
K322E
W274L
C211W
E157G
T202I
E256K
Y65C
H93R
L295V
N356H
D268E
F279L
N117S
N228S
P357S
T348S
T363N
A309S
A79T
I203V
K330E
M198I
N262S
N356D
P354Q
Q298E
Y346F
4A
Q396R
pAKT/AKT normalized score
Δvac14 functional score
Δfig4 functional score
Δvac7 functional score
Δvam3 functional score
Δvam7 functional score
Δvps30 functional score
Δvps38 functional score
Δypt7 functional score
Fly Eclosion
Worm Chemotaxis
Rat Axon Length
Rat Soma Size
Rat PSD95
Rat Total Dendrite Length
Yeast protein stability
HEK protein stability

Figure 4.2: Multi-model classification of variants

The results of HEK293 pAKT/AKT, the eight yeast sentinel functional scores, and results from fly worm and rat hippocampal neuron culture experiments are summarized. Shading is done according to: WT-like variants exhibit no differences to wildtype PTEN, gain of function (GoF) variants are greater than WT, partial LoF/hypomorphs are greater than null but less than WT, LoF exhibit no differences to zero, dominant negative are less than both null and WT. All differences (greater or less than) stated above imply statistically significant differences of at least $p<0.05$ compared to the WT-PTEN overexpression effects, null or both in their respective assay. White cells indicate that the variant was not put through the assay.

4.2 Application of the multi-model approach

Our work highlights the importance of using multiple assays to study the function of a gene and the implications of VUS. While we have learned much from the yeast and HEK293 cell assays, we still do not have a clear understanding of the full range of implications a VUS has on protein function and stability. While it is important to study many variants, doing so at the expense of depth might lead to incomplete conclusions. Our analysis is a first step towards functionalizing variants, or functional variomics. Currently, the standard clinical and therapeutic interventions for people with ASD do not take into account the genetic or biological underpinnings specific to each person. However, hopefully this will not always be the case. There is no single medical treatment that is beneficial to all symptoms of ASD (Sanchack and Thomas, 2016), and several of the treatments have adverse side effects that reduce their benefits (McPheeters et al., 2011). Ideally, ASD research could follow in the footsteps of cancer research where recently, huge efforts have yielded progress in connecting genotypes to phenotypes, prognosis, and treatment of the disease (Yi et al., 2017). A push towards personalized treatments

would be benefited by additional studies looking at the impact of VUS on multiple aspects of protein function.

4.3 Comparison with previously published data on PTEN stability and function

As previously mentioned, two studies were published recently which also attempted to understand the impact of point mutations on PTEN function. Mighell et al., took advantage of a previously validated assay in which yeast were humanized via overexpression of the catalytic subunit of PI3K which is lethal to yeast without the concurrent expression of functional PTEN (Rodriguez-Escudero et al., 2005; Rodriguez-Escudero et al., 2011; Mighell et al., 2018). Therefore, their assay is a measure of the ability of PTEN and PTEN variants to antagonize the function of PI3K. Matreyek et al. measured PTEN protein abundance using saturation mutagenesis followed by flow-cytometry and VAMP-seq. We were interested to see how the results from these other studies compared to our HEK293 and yeast function and stability scores. In order to compare measures from the two papers with the functional and stability scores from HEK293 and the yeast *Δvac14* sentinel, all scores were normalized such that a normalized score of 1 corresponds to high levels of function or stability similar to wildtype PTEN and a normalized score of 0 corresponds to low levels of function or stability similar to empty vector. We find that the Mighell et al. normalized PTEN functional scores are much higher than our functional scores in the HEK293 pAKT/AKT and *Δvac14* genetic interaction assays (Figure 4.3). This was not surprising since their assay found 67% of nonsynonymous mutations to have wildtype levels of function with about half of all PTEN amino acid positions tolerant to mutation (Mighell et al., 2018). Comparing the stability results from both yeast and HEK293 to the results from Mighell et al., we find a high level of correlation (Figure 4.3b-c) (Spearman correlations of 0.680 and 0.849 respectively). We see a higher level of correlation between our HEK293 data

and the data from Matreyek et al. which was also collected from HEK293 cells using flow cytometry. This indicates that the degradation of variants may be slightly different in yeast cells than in HEK293 cells.

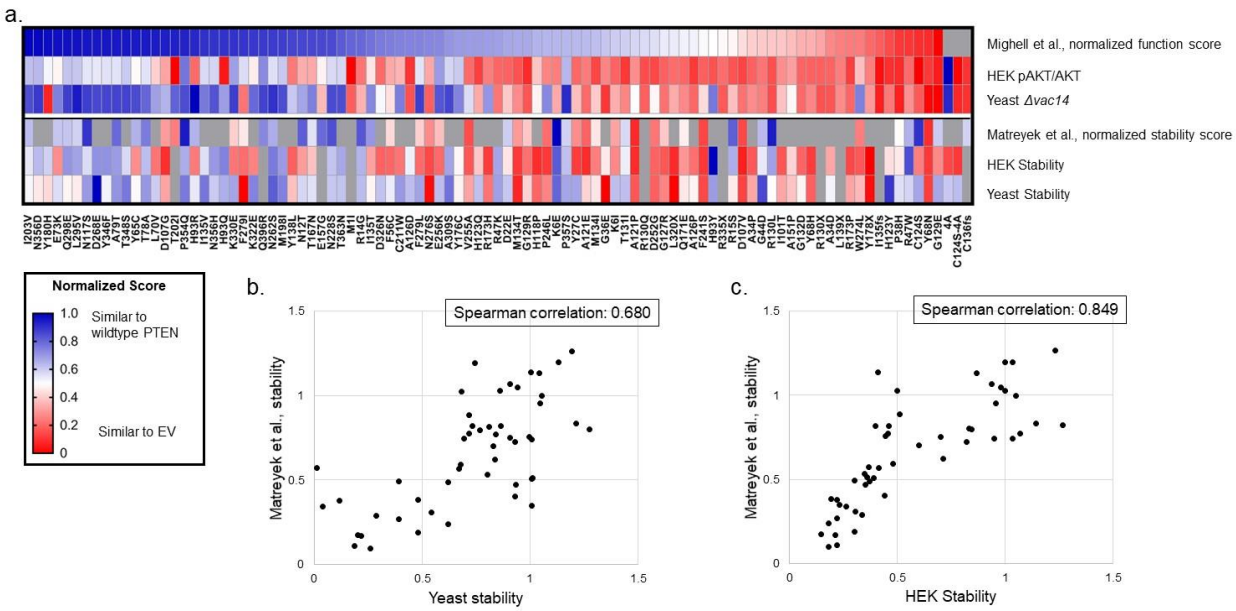

Figure 4.3: Comparison of functional and stability scores from yeast and HEK293 models to previous reports

A heatmap (a) displaying function scores from Mighell et al., yeast sentinel *Δvac14* and HEK293 pAKT/AKT as well as stability data from Matreyek et al., HEK293, and yeast. All scores were normalized in order to compare across the different measures. To better visualize correlation of stability scores, Matreyek et al. stability is compared to our yeast stability (b) and HEK293 stability (c). Both are highly correlated with Spearman correlation coefficients of 0.680 and 0.849 respectively.

4.4 Experimental considerations and limitations

4.4.1 Strengths of the yeast mini array and HEK293 cell pAKT/AKT assays

Both the yeast and HEK293 cell assays created opportunities to parse the impact of PTEN variant expression. In yeast, we were able to capitalize on decades of work fine tuning the SDL assay (Tong et al., 2001; Tong and Boone 2005; Botstein and Fink, 2011; Duina et al.,

2014; Giaver and Nislow, 2014; Ferreira et al., 2019) which made it ideal for identifying genetic interactions specific to *PTEN*. We were also able to create a novel mini-array assay which took advantage of technology perfected for expression of additional genetic modifications to the DMA (Tong et al., 2001; Giaever et al., 2002; Tong and Boone 2005). This design allowed us to efficiently and quantitatively test the consequences of VUS to pre-determined genetic interactions specific to *PTEN*. By performing the first generation mini array we learned which sentinels were ideal for screening variants and were able to take advantage of this knowledge in the design of the second generation mini array. HEK293 cells, which are derived from human epithelial cells, are likely to appropriately express, fold, and post-translationally modify exogenously expressed human proteins (Thomas and Smart, 2005) which makes them a good model to study the effect of expression of a human protein on human cells. The high-throughput nature of cell culture and the flow cytometry assay allowed us to concurrently test many variants. Flow cytometry allowed us to study the impact of exogenous *PTEN* expression on the phosphorylation of AKT on a single-cell level. Combined, the results from experiments in these two model systems begin to elucidate how *PTEN* variants alter various aspects of PTEN relative to wildtype function.

4.4.2 Limitations of the current model

It is worth noting that all experiments in the yeast and HEK293 assays, as well as the other experiments in the multi-lab collaboration, take advantage of overexpressing exogenous PTEN. By comparing all results to overexpression of wildtype PTEN, we attempt to control for alterations to cell physiology due to exogenous protein overexpression not specific to point mutations. In HEK293 cells we use the constitutive cytomegalovirus early enhancer/chicken β-actin (CAG) promoter to induce gene expression. While this is a constitutively active promoter,

it has been found to induce less overexpression in HEK293 cells than the constitutive cytomegalovirus (CMV) promoter (Qin et al., 2010). Additionally, we only included cells with detection levels of GFP less than a ten-fold increase over autofluorescence. However, plasmid expression can alter a cells physiology due to an overexpression of the protein of interest which can cause resource overload, stoichiometric imbalance, promiscuous interactions, and modulation of pathways regardless of function (Moriya, 2015). However, since most probands are heterozygous for PTEN mutations, it is important to understand how these VUS behave in the presence of endogenous PTEN protein.

A major limitation in our analysis is the lack of patient phenotype data. While information has been collated where data exists, we were unable to attain detailed information on the phenotype of most of the ASD and PHTS probands. Additional data on intellectual capabilities or comorbidities associated with ASD cases such as gastrointestinal disorders, ADHD, sleep disorders, epilepsy, etc. would greatly enhance our understanding of the implications of the variants on disorder outcome. It would be a great opportunity to relate the levels of PTEN function to severity of the patient's condition. Additionally, because our data indicates that PTEN function is disrupted along a continuum with certain variants displaying hypomorphic function, the guidelines previously set out by Mester et al. determining pathogenic variants on an all or nothing scale are reductive. Although great strides have been made within the community of scientists who study *PTEN*, restricting our data to the current guidelines (Mester et al., 2018) causes us to lose the distinction between LoF and partial LoF which might be relevant. Also, since we see LoF in some contexts and not in others, calling a variant either pathogenic or benign may be glossing over context-specific details. Therefore, a refined system

of variant curation would give more insight into the actual changes to function caused by point mutations.

Another limitation to the current work is that we are not studying all possible mutations of *PTEN*. This project was focused on determining the impact of variants associated with human populations, but in doing so we may be missing other variants not yet identified in humans. A saturation mutagenesis approach, which has been utilized by other labs (Matreyek et al., 2018; Mighell et al., 2018), would be a good approach to overcome this deficit. While this would greatly reduce the ability of our assays to be high-throughput, it would put the function of each variant into perspective of all other possible variants. Saturation mutagenesis could also be used as an efficient tool for making many point mutations in a library then selecting additional variants to study as needed. However, this approach would still not cover all mutations found in *PTEN* or even all mutations found in *PTEN* associated with disease or disorder. Various splice site mutations and mutations identified in untranslated regions may also impact *PTEN* transcription or protein function (Buxbaum et al., 2007; Rodriguez-Escudero et al., 2011; Chen et al., 2017). Now that whole genome sequencing is being taken advantage of with increasing frequency (Brandler et al., 2018; Guo et al., 2018; McKenna et al., 2018), it would be interesting to determine how frequently these types of mutations arise for PTEN.

4.5 Future directions

4.5.1 Next experiments for PTEN

While we were able to test PTEN's genetic interactions and its regulation of the phosphorylation of AKT, there are many other aspects of PTEN function which may be contributing to ASD. While some of these are being addressed in the fly, worm, and hippocampal neuron cultures in our collaboration, additional experiments could be added which

would contribute to a well-rounded understanding of PTEN function. Of specific interest is the impact of PTEN protein localization. PTEN is known to translocate to the nucleus (Planchon et al., 2008) where it is thought to regulate gene transcription and genomic stability (Planchon et al., 2008; Gong et al., 2015) as well as slow growth rate by causing cells to arrest in G0-G1 phase of growth (Planchon et al., 2008). Of note in the scope of our study is that certain ASD-associated PTEN variants have been found to be nuclear excluded (Fricano-Kugler et al., 2018). This previous study found three variants to be nuclear excluded, two of which we find to be complete LoF in the HEK293 pAKT/AKT assay and partial LoF in the yeast *Δvac14* sentinel (F241S and D252G) and one which is partial LoF in both the HEK293 pAKT/AKT assay and the yeast *Δvac14* sentinel (W274L). It would be very valuable to test these and the rest of the PTEN variants we studied for nuclear localization.

While the lipid phosphatase activity of PTEN is well known, it would also be interesting to determine how point mutations alter the ability of PTEN to act as a protein phosphatase. Several potential protein phosphatase targets of PTEN have already been identified including PTEN itself (Tibarewal et al., 2012), PTK6 (Wozniak et al., 2017), Drebrin (Kreis et al., 2013; Garcia-Junco-Clemente and Golshani, 2014), and CREB (Gu et al., 2011). Additionally, determining how variants impact the protein binding partners of PTEN would further elucidate how point mutations could disrupt PTEN function. PTEN has several known protein interactions such as NHERF, Neprilysin, MAGI-1b, MAGI-2, and p85 (Sumitomo et al., 2004; Kotelevets et al., 2005; Takahashi et al., 2006; Hu et al., 2007; Rabinovsky et al., 2009) and interaction could be determined via immunoprecipitation techniques. However, this would be hard to do in a high-throughput manner. Therefore, an assay utilizing BioID (Roux et al., 2013) would be a good choice for high-throughput screening for protein interactions.

Further neuron culture or transgenic animal experiments would increase our understanding of how PTEN variants impact neuron development and behavior. Expressing variants in a conditional knock-out mouse (Takeuchi et al., 2013; Garcia-Junco-Clemente and Golshani, 2014) or in neurons treated with shRNA to knockdown endogenous PTEN would enable us to learn how PTEN variants impact brain and neuron development without the presence of endogenous PTEN. Finally, there are alternate isoforms of PTEN which would be interesting to study. Specifically, the recently characterized PTEN-Long isoform has been found to be secreted and internalized by cells in addition to antagonizing the PI3K pathway (Hopkins et al., 2013; Hopkins et al., 2014). Determining if VUS alter PTEN-Long secretion and internalization would elucidate another way VUS could be impacting function.

4.5.2 Next genes for study

In addition to studying the impact of *PTEN* VUS on genetic interactions, we endeavored to create a novel mini array approach which could be used to study ASD-associated genes regardless of protein function. The benefit of this assay is that despite yeast not having a PTEN homolog or PTEN's main catalytic substrate we were able to identify genetic interactions specific to PTEN's catalytic function. Additionally, the results from these sentinels were highly correlated with the pAKT/AKT values generated in the HEK293 cell assay. This indicates that this assay identified differences relevant to known functions of PTEN and gives credence to the belief that it could also be used to study other genes. While this assay is ideal for genes with a large number of variants we wish to study, it is conceivable that we would also be interested in studying genes which have few ASD-associated VUS. According to the SFARI database, there are currently 1,089 genes associated with ASD. In addition to PTEN, there are 24 other genes designated as high-confidence genes and 66 strong-candidate genes (Abrahams et al., 2013). By

selecting the next gene of study from this list there will likely be more VUS associated with ASD to test. However, it is also important to consider the ability of whatever gene is selected to be testable in all model systems.

4.5.3 Better pipeline generation

As we have progressed in generating the multi-lab, multi-model pipeline, we have found ways to streamline where possible. We found that using Gateway cloning techniques makes plasmid construction easier, faster, and less error prone. Therefore, shifting all model systems to Gateway vectors in future will increase the speed at which we generate clones for use in experiments. For the yeast mini array creation, we were initially unsure about how many deletion strains to include. Then we were challenged to design the novel mini array such that we are able to test all relevant sentinel hits with all variants efficiently. Now that we are experienced in selecting sentinels and designing refined mini-arrays, we are well prepared to create mini arrays for future genes. Additionally, since completing our PTEN mini-array, we realized the importance of being able to compare VUS to empty vector as well as wildtype PTEN. For future designs, it would be prudent to include an empty vector colony next to each variant in addition to having an adjacent wildtype colony.

The HEK293 cell pAKT/AKT assay, while great for studying PTEN is less ideal for a platform designed to test all genes regardless of function because it is unlikely that all ASD-associated genes alter the pAKT/AKT ratio. Therefore, when selecting next genes for study it is important to determine potential targets to assay via flow cytometry in HEK293 cells on a gene-by-gene basis. While pAKT may not be relevant to a large number of genes, there are many other signaling pathways that may be influenced by multiple genes such as the extracellular signal-related kinase (ERK)/mitogen activated protein kinase (MAPK) (Subramanian et al.,

2015; Vithayathil et al., 2018) and Wnt (Kumar et al., 2019). Additionally, we learned that an important mechanism of dysfunction for variants is reduced protein stability. Therefore, in the future it might be prudent to test for protein stability prior to running both high and low-throughput experiments. If a variant is unstable, it will not need to be tested in functional assays.

4.6 Future implications

A variety of model systems have recently been used to functionally analyze the consequence of VUS found in several ASD-associated genes on biological processes (den Hoed et al., 2018; Fricano-Kugler et al., 2018; Mighell et al., 2018; Widowati et al., 2018; Wong et al., 2019). As our results indicate, variants have different dysfunction profiles depending on the assay performed. This kind of heterogeneity of effect is best characterized using a diversity of experimental functionalization and may underly the clinical heterogeneity of ASD phenotypes. Therefore, to understand the impact of VUS on gene function, we should continue to expand our repertoire of functionalization experiments. While functionalization of variants does not take into account the genetic background a variant was identified in, it likely broadens our understanding of the effects of point mutations as well as potentially elucidating unknown gene functions and the underlying etiology of ASD. Understanding the biological underpinnings specific to each person will go a long way towards achieving our overarching goal of determining the pathophysiology of this disorder in an effort to personalize therapeutic and clinical interventions.

Bibliography

Abeliovich A, and Gitler AD. (2016) Defects in Trafficking Bridge Parkinson's Disease Pathology and Genetics. *Nature* 539(7628):207–16.

Abrahams BS, Arking DE, Campbell DB, Mefford HC, Morrow EM, Weiss LA, Menashe I, Wadkins T, Banerjee-Basu S, and Packer A. (2013) SFARI Gene 2.0: A Community-Driven Knowledgebase for the Autism Spectrum Disorders (ASDs). *Molecular Autism* 4(1):36.

Aceti M, Creson TK, Vaissiere T, Rojas C, Huang W, Wang Y, Petralia RS, Page DT,. Miller CA, and Rumbaugh G. (2015) Syngap1 Haploinsufficiency Damages a Postnatal Critical Period of Pyramidal Cell Structural Maturation Linked to Cortical Circuit Assembly." *Biological Psychiatry* 77(9):805–15.

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, and Sunyaev SR. (2010) A Method and Server for Predicting Damaging Missense Mutations. *Nature Methods* 7(4):248–49.

Ahmad SF, Ansari MA, Nadeem A, Bakheet SA, AL-Ayadhi LY, Alotaibi MR, Alhoshani AR, Al-Hosaini KA, and Attia SM. (2018) Dysregulation of the Expression of HLA-DR, Costimulatory Molecule, and Chemokine Receptors on Immune Cells in Children with Autism. *International Immunopharmacology* 65:360–65.

Alonso-Gonzalez A, Rodriguez-Fontenla C, and Carracedo A. (2018) De Novo Mutations (DNMs) in Autism Spectrum Disorder (ASD): Pathway and Network Analysis. *Frontiers in Genetics* 9: 406.

Amaral DG, Schumann CM, and Nordahl CW. (2008) Neuroanatomy of Autism. *Trends in Neurosciences* 31(3):137–45.

Amberg DC, Burke DJ, and Strathern JN. (2005) *Methods in Yeast Genetics*, 2005 edition, Cold Spring Harbor Press, New York.

An JY, Lin K, Zhu L, Werling DM, Dong S, Brand H, Wang HZ, Zhao X, Schwartz GB, Collins RL, Currall BB, Dastmalchi C, Dea J, Duhn C, Gilson MC, Klei L, Liang L, Markenscoff-Papadimitriou E, Pochareddy S, Ahituv N, Buxbaum JD, Coon H, Daly MJ, Kim YS, Marth GT, Neale BM, Quinlan AR, Rubenstein JL, Sestan N, State MW, Willsey AJ, Talkowski ME, Devlin B, Roeder K, and Sanders SJ. (2018) Genome-Wide de Novo Risk Score Implicates Promoter Variation in Autism Spectrum Disorder. *Science* 362(6420).

Andersen LL, Terczyńska-Dyla E, Mørk N, Scavenius C, Enghild JJ, Höning K, Hornung V, Christiansen M, Mogensen TH, and Hartmann R. (2019) Frequently used bioinformatis tools overestimate the damaging effect of allelic variants. *Genes Immun* 20(1):10-22.

Auyeung B, Lombardo MV, and Baron-Cohen S. (2013) Prenatal and Postnatal Hormone Effects on the Human Brain and Cognition. *Pflügers Archiv - European Journal of Physiology* 465(5):557–71.

Backman SA, Stambolic V, Suzuki A, Haight J, Elia A, Pretorius J, Tsao MS, Shannon P, Bolon B, Ivy GO, and Mak TW. (2001) Deletion of Pten in Mouse Brain Causes Seizures, Ataxia and Defects in Soma Size Resembling Lhermitte-Duclos Disease. *Nature Genetics* 29(4):396–403.

Barnard-Brak L, Sulak T, and Ivey Hatz JK. (2011) Macrocephaly in Children With Autism Spectrum Disorders. *Pediatric Neurology* 44(2):97–100.

Baron-Cohen, S, Auyeung B, Nørgaard-Pedersen B, Hougaard DM, Abdallah MW, Melgaard L, Cohen AS, Chakrabarti B, Ruta L, and Lombardo MV. (2015) Elevated Fetal Steroidogenic Activity in Autism. *Molecular Psychiatry* 20(3):369–76.

Bermúdez Brito M, Goulielmaki E, and Papakonstanti EA.. (2015) Focus on PTEN Regulation. *Frontiers in Oncology* 5:166.

Bauman MD, and Schumann CM. (2018) Advances in Nonhuman Primate Models of Autism: Integrating Neuroscience and Behavior. *Experimental Neurology* 299:252–65.

Bier E. (2005) Drosophila, the Golden Bug, Emerges as a Tool for Human Genetics. *Nature Reviews Genetics* 6(1):9–23.

Bischof J, Bjorklund M, Furger E, Schertel C, Taipale J, and Basler K. (2013) A Versatile Platform for Creating a Comprehensive UAS-ORFeome Library in Drosophila. *Development* 140(11):2434–42.

Blumenthal GM, and Dennis PA. (2008) PTEN Hamartoma Tumor Syndromes. *European Journal of Human Genetics* 16(11):1289–1300.

Boeddinghaus C, Merz AJ, Laage R, and Ungermann C. (2002) A Cycle of Vam7p Release from and PtdIns 3-P–Dependent Rebinding to the Yeast Vacuole Is Required for Homotypic Vacuole Fusion. *The Journal of Cell Biology* 157(1):79–90.

Bölte S, Girdler S, and Marschik PB. (2019) The Contribution of Environmental Exposure to the Etiology of Autism Spectrum Disorder. *Cellular and Molecular Life Sciences* 76(7):1275-1297.

Bonaguidi MA, Wheeler MA, Shapiro JS, Stadel RP, Sun GJ, Ming G, and Song H. (2011) In Vivo Clonal Analysis Reveals Self-Renewing and Multipotent Adult Neural Stem Cell Characteristics. *Cell* 145(7):1142–55.

Boone C, Bussey H, Andews BJ. (2007) Exploring genetic interactions and networks with yeast. *Nat Reviews* 8:437-449.

Botstein D, and Fink GR. (2011) Yeast: An Experimental Organism for 21st Century Biology. *Genetics* 189(3):695–704.

Bourgeron, Thomas. (2015) From the Genetic Architecture to Synaptic Plasticity in Autism Spectrum Disorder. *Nature Reviews Neuroscience* 16(9):551–63.

Bozdagi O, Sakurai T, Papapetrou D, Wang X, Dickstein DL, Takahashi N, Kajiwara Y, Yang M, Katz AM, Scattoni ML, Harris MJ, Saxena R, Silverman JL, Crawley JN, Zhou Q, Hof PR, and Buxbaum JD. (2010) Haploinsufficiency of the Autism-Associated Shank3 Gene Leads to Deficits in Synaptic Function, Social Interaction, and Social Communication. *Molecular Autism* 1(1):15.

Brandler WM, Antaki D, Gujral M, Kleiber ML, Whitney J, Maile MS, Hong O, Chapman TR, Tan S, Tandon P, Pang T, Tang SC, Vaux KK, Yang Y, Harrington E, Juul S, Turner DJ, Thiruvahindrapuram B, Kaur G, Wang Z, Kingsmore SF, Gleeson JG, Bisson D, Kakaradov B, Telenti A, Venter JC, Corominas R, Toma C, Cormand B, Rueda I, Guijarro S, Messer KS, Nievergelt CM, Arranz MJ2, Courchesne E, Pierce K, Muotri AR, Iakoucheva LM, Hervas A, Scherer SW, Corsello C, and Sebat J. (2018) Paternally Inherited Cis-Regulatory Structural Variants Are Associated with Autism." *Science* 360(6386):327–31.

Bromberg Y and Rost B. (2007) SNAP: predict effect of non-synonymous polymorphisms on function. *Nucleic Acids Res* 35(11): 3823-35.

Brown AS, Surcel H, Hinkka-Yli-Salomäki S, Cheslack-Postava K, Bao Y, and Sourander A. (2015) Maternal Thyroid Autoantibody and Elevated Risk of Autism in a National Birth Cohort. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 57: 86–92.

Brueggeman L, Koomar T, and Michaelson J. (2018) Forecasting Autism Gene Discovery with Machine Learning and Genome-Scale Data. Preprint. *bioRxiv* Preprint.

Buler MG. (2005) Subset of Individuals with Autism Spectrum Disorders and Extreme Macrocephaly Associated with Germline PTEN Tumour Suppressor Gene Mutations. *Journal of Medical Genetics* 42(4):318–21.

Busa T, Chabrol B, Perret O, Longy M, and Philip N. (2013) Novel PTEN Germline Mutation in a Family with Mild Phenotype: Difficulties in Genetic Counseling. *Gene* 512(2):194–97.

Busa T, Milh M, Degardin N, Girard N, Sigaudy S, Longy M, Olshchwang S, Sobol H, Chabrol B, and Philip N. (2015) Clinical Presentation of PTEN Mutations in Childhood in the Absence of Family History of Cowden Syndrome. *European Journal of Paediatric Neurology* 19(2):188–92.

Butler MG. (2005) Subset of Individuals with Autism Spectrum Disorders and Extreme Macrocephaly Associated with Germline PTEN Tumour Suppressor Gene Mutations. *Journal of Medical Genetics* 42(4):318–21.

Buxbaum JD, Cai G, Chaste P, Nygren G, Goldsmith J, Reichert J, Anckarsäter H, Rastam M, Smith CJ, Silverman JM, Hollander E, Leboyer M, Gillberg C, Verloes A, and Betancur C. (2007) Mutation Screening of The PTEN Gene in Patients with Autism Spectrum Disorders and Macrocephaly. *American Journal of Medical Genetics: Neuropsychiatric Genetics* 144B(4):484–91.

Caux F, Plauchu H, Chibon F, Faivre L, Fain O, Vabres P, Bonnet F, Selma ZB, Laroche L, Gérard M, and Longy M. (2007) Segmental overgrowth, lipomatosis, arteriovenous malformation and epidermal nevus (SOLAMEN) syndrome is related to mosaic PTEN nullizygosity. *Eur J Hum Genet* 15(7):767-73.

Chahrour M, O'Roak BJ, Santini E, Samaco RC, Kleiman RJ, and Manzini MC. (2016) Current Perspectives in Autism Spectrum Disorder: From Genes to Therapy. *The Journal of Neuroscience* 36(45):11402–10.

Chaste P, Klei L, Sanders SJ, Hus V, Murtha MT, Lowe JK, Willsey AJ, Moreno-De-Luca D, Yu TW, Fombonne E, Geschwind D, Grice DE, Ledbetter DH, Mane SM, Martin DM, Morrow EM, Walsh CA, Sutcliffe JS, Martin CL, Beaudet AL, Lord C, State MW, Cook Jr EH, Devlin B. (2015) A Genome-wide association study of autism using the Simons Simplex Collection: Does reducing phenotypic heterogeneity in autism increase genetic homogeneity? *Biological Psychiatry* 77(9):775-784.

Chen HJ, Romigh T, Sesock K, and Eng C. (2017) Characterization of Cryptic Splicing in Germline *PTEN* Intronic Variants in Cowden Syndrome: CHEN et Al. *Human Mutation* 38(10):1372–77.

Cherry JM, Hong EL, Amundsen C, Balakrishnan R, Binkley G, Chan ET, Christie KR, Costanzo MC, Dwight SS, Engel SR, Fisk DG, Hirschman JE, Hitz BC, Karra K, Krieger CJ, Miyasato SR, Nash RS, Park J, Skrzypek MS, Simison M, Weng S, and Wong ED. (2012) Saccharomyces Genome Database: The Genomics Resource of Budding Yeast. *Nucleic Acids Research* 40(D1):D700–705.

Ciaccio C, Saletti V, D'Arrigo S, Esposito S, Alfei E, Moroni I, Tonduti D, Chiapparini L, Pantaleoni C, and Milani D. (2018) Clinical Spectrum of PTEN Mutation in Pediatric Patients. A Bicenter Experience. *European Journal of Medical Genetics* Epub ahead of print.

Clipperton-Allen AE, and Page DT. (2014) Pten Haploinsufficient Mice Show Broad Brain Overgrowth but Selective Impairments in Autism-Relevant Behavioral Tests. *Human Molecular Genetics* 23(13):3490–3505.

Clipperton-Allen AE, and Page DT. (2015) Decreased Aggression and Increased Repetitive Behavior in *Pten* Haploinsufficient Mice: Aggression and Repetitive Behavior in *Pten* $^{+/-}$ Mice. *Genes, Brain and Behavior* 14(2):145–57.

Courchesne E, Pramparo T, Gazestani VH, Lombardo MV, Pierce K, and Lewis NE. (2019) The ASD Living Biology: From Cell Proliferation to Clinical Phenotype. *Molecular Psychiatry* 24(1):88-107.

Codina-Solà M, Rodríguez-Santiago B, Homs A, Santoyo J, Rigau M, Aznar-Laín G, Del Campo M, Gener B, Gabau E, Botella MP, Gutiérrez-Arumí A, Antiñolo G, Pérez-Jurado LA, and Cuscó I. (2015) Integrated Analysis of Whole-Exome Sequencing and Transcriptome Profiling in Males with Autism Spectrum Disorders." *Molecular Autism* 6(1).

Conti S, Condò M, Posar A, Mari F, Resta N, Renieri A, Neri I, Patrizi A, Parmeggiani A. (2012) Phosphatase and Tensin Homolog (PTEN) Gene Mutations and Autism: Literature Review and a Case Report of a Patient With Cowden Syndrome, Autistic Disorder, and Epilepsy. *Journal of Child Neurology* 27(3):392–97.

Costanzo M, VanderSluis B, Koch EN, Baryshnikova A, Pons C, Tan G, Wang W, Usaj M, Hanchard J, Lee SD, Pelechano V, Styles EB, Billmann M, van Leeuwen J, van Dyk N, Lin ZY, Kuzmin E, Nelson J, Piotrowski JS, Srikumar T, Bahr S, Chen Y, Deshpande R, Kurat CF, Li SC, Li Z, Usaj MM, Okada H, Pascoe N, San Luis BJ, Sharifpoor S, Shuteriqi E1, Simpkins SW, Snider J, Suresh HG, Tan Y, Zhu H, Malod-Dognin N, Janjic V, Przulj N, Troyanskaya OG, Stagljar I, Xia T, Ohya Y, Gingras AC, Raught B, Boutros M, Steinmetz LM, Moore CL, Rosebrock AP, Caudy AA, Myers CL, Andrews B, and Boone C. (2016) A Global Genetic Interaction Network Maps a Wiring Diagram of Cellular Function. *Science* 353(6306):aaf1420–aaf1420.

Crawley, Jacqueline N. (2012) Translational Animal Models of Autism and Neurodevelopmental Disorders." *Dialogues in Clinical Neuroscience* 14(3):293–305.

Cupolillo D, Hoxha E, Faralli A, De Luca A, Rossi F, Tempia F, and Carulli D. (2016) Autistic-Like Traits and Cerebellar Dysfunction in Purkinje Cell PTEN Knock-Out Mice. *Neuropsychopharmacology* 41(6):1457–66.

Dabell MP, Rosenfeld JA, Bader P, Escobar LF, El-Khechen D, Vallee SE, Dinulos MB, Curry C, Fisher J, Tervo R, Hannibal MC, Siefkas K, Wyatt PR, Hughes L, Smith R, Ellingwood S, Lacassie Y, Stroud T, Farrell SA, Sanchez-Lara PA, Randolph LM, Niyazov D, Stevens CA, Schoonveld C, Skidmore D, MacKay S, Miles JH, Moodley M, Huillet A, Neill NJ, Ellison JW, Ballif BC, and Shaffer LG. (2013) Investigation of *NRXN1* Deletions: Clinical and Molecular Characterization. *American Journal of Medical Genetics* 161(4):717–31.

Daum G, Lees ND, Bard M, and Dickson R. (1998) Biochemistry, Cell Biology and Molecular Biology of Lipids Of Saccharomyces Cerevisiae. *Yeast* 14(16):1471–1510.

Davidson L, Maccario H, Perera NM, Yang X, Spinelli L, Tibarewal P, Glancy B, Gray A, Weijer CJ, Downes CP, and Leslie NR. (2010) Suppression of cellular proliferation and invasion by the concerted lipid and protein phosphatase activities of PTEN. *Oncogene* 29(5):687-697.

Davies R, Vogelsang P, Jonsson R, and Appel S. (2016) An Optimized Multiplex Flow Cytometry Protocol for the Analysis of Intracellular Signaling in Peripheral Blood Mononuclear Cells. *Journal of Immunological Methods* 436:58–63.

de la Torre-Ubieta L, Won H, Stein JL, Geschwind GH (2016) Advancing the understanding of autism disease mechanisms through genetics. *Nature Medicine* 22:345-361.

den Hoed J, Sollis E, Venselaar H, Estruch SB, Deriziotis P, Fisher SE. (2018) Functional characterization of TBR1 variants in neurodevelopmental disorder. *Sci Rep* 8(1):14279.

DeStefano F and Shimabukuro TT. (2019) The MMR Vaccine and Autism. *Annual Review of Virology* 6(1).

Devasahayam G, Burke DJ, and Sturgill TW. (2007) Golgi Manganese Transport Is Required for Rapamycin Signaling in *Saccharomyces Cerevisiae*. *Genetics* 177(1):231–38.

Devlin B, and Scherer SW. (2012) Genetic Architecture in Autism Spectrum Disorder. *Current Opinion in Genetics & Development* 22(3):229–37.

Dhungel N, Eleuteri S, Li L, Kramer NJ, Chartron JW, Spencer B, Kosberg K, Fields JA, Stafa K, Adame A, Lashuel H, Frydman J, Shen K, Masilah E, Gitler AD. (2014) Parkinson's disease genes VPS35 and EIF4G1 interact genetically and converge on α-synuclein. *Neuron* 85:76-87.

Dill KA, Ozkan SB, Shell MS, and Weikl TR. (2008) The Protein Folding Problem. *Annual Review of Biophysics* 37(1):289–316.

Duina, Andrea A., Mary E. Miller, and Jill B. Keeney. (2014) Budding Yeast for Budding Geneticists: A Primer on the *Saccharomyces Cerevisiae* Model System. *Genetics* 197(1):33–48.

Elia M, Amato C, Bottitta M, Grillo L, Calabrese G, Esposito M, and Carotenuto M. (2012) An atypical patient with Cowden syndrome and PTEN gene mutation presenting with cortical malformation and focal epilepsy. *Brain Dev* (10):873-6.

Ersahin T, Tuncbag N, and Cetin-Atalay R. 2015 The PI3K/AKT/MTOR Interactive Pathway. *Molecular BioSystems* 11(7):1946–54.

Ferreira R, Limeta A, and Nielsen J. (2019) Tackling Cancer with Yeast-Based Technologies. *Trends in Biotechnology* 37(6):592–603.

Fombonne E. (2003) Modern Views of Autism. *The Canadian Journal of Psychiatry* 48(8):503–5.

Fombonne E. (2005) Epidemiology of Autistic Disorder and Other Pervasive Developmental Disorders. *J Clin Psychiatry* 66(10):3-8.

Fraser MM, Bayazitov IT, Zakharenko SS, and Baker SJ. (2008) Pten Deficiency in Brain Causes Defects in Synaptic Structure, Transmission and Plasticity, and Myelination Abnormalities. *Neuroscience* 151(2):476-488.

Frazier TW, Embacher R, Tilot AK, Koenig K, Mester J, and Eng C. (2015) Molecular and Phenotypic Abnormalities in Individuals with Germline Heterozygous PTEN Mutations and Autism. *Molecular Psychiatry* 20(9):1132–38.

Fricano-Kugler CJ, Getz SA, Williams MR, Zurawel AA, DeSpenza T Jr, Frazel PW, Li M, O'Malley AJ, Moen EL, and Luikart BW. (2018) Nuclear Excluded Autism-Associated Phosphatase and Tensin Homolog Mutations Dysregulate Neuronal Growth. *Biological Psychiatry* 84(4):265–77.

Garcia-Junco-Clemente P, and Golshani P. (2014) PTEN: A Master Regulator of Neuronal Structure, Function, and Plasticity. *Communicative & Integrative Biology* 7(2): e28358.

Gaugler T, Klei L, Sanders SJ, Bodea CA, Goldberg AP, Lee AB, Mahajan M, Manaa D, Pawitan Y, Reichert J, Ripke S, Sandin S, Sklar P, Svantesson O, Reichenberg A, Hultman CM, Devlin B, Roeder K, & Buxbaum JD. (2014) Most Genetic Risk for Autism Resides with Common Variation. *Nature Genetics* 46(8):881–85.

Gary JD, Sato TK, Stefan CJ, Bonangelino CJ, Weisman LS, and Emr SD. (2002) Regulation of Fab1 Phosphatidylinositol 3-Phosphate 5-Kinase Pathway by Vac7 Protein and Fig4, a Polyphosphoinositide Phosphatase Family Member. *Molecular Biology of the Cell* 13(4):1238–51.

Geschwind DH. (2011) Genetics of Autism Spectrum Disorders. *Trends in Cognitive Sciences* 15(9):409–16.

Giaever G., Chu AM, Ni L, Connelly C, Riles L, Veronneau S, Dow S, Lucau-Danila A, Anderson K, Andre B, Arkin AP, Astromoff A, El Bakkoury M, Bangham R, Benito R, Brachat S, Campanaro

S, Curtiss M, Davis K, Deutschbauer A, Entian K, Flaherty P, Foury F, Garfinkel DJ, Gerstein M, Gotte D, Guldener U, Hegemann JH, Hempel S, Herman Z, Jaramillo DF, Kelly DE, Kelly SL, Kotter P, LaBonte D, Lamb DC, Lan N, Liang H, Liao H, Liu L, Luo C, Lussier M, Mao R, Menard P, Ooi SL, Revuelta JL, Roberts CJ, Rose M, Ross-Macdonald P, Scherens B, Schimmack G, Shafer B, Shoemaker DD, Sookhai-Mahadeo S, Storms RK, Strathern JN, Valle G, Voet M, Volckaert G, Wang C, Ward TR, Wilhelmy J,. Winzeler EA, Yang Y, Yen G, Youngman E, Yu K, Bussey H, Boeke JD, Snyder M, Philippsen P,. Davis RW, & Mark Johnston M. (2002) Functional profiling of the Saccharomyces cerevisiae genome. *Nature* 418:387-391.

Giaever G, and Nislow C. (2014) The Yeast Deletion Collection: A Decade of Functional Genomics. *Genetics* 197(2):451–65.

Goffin A, Hoefsloot LH, Bosgoed E, Swillen A, Fryns JP. (2001) PTEN mutation in a family with Cowden syndrome and autism. *Am J Med Genet* 105(6):521-4.

Gong L, M Govan JM,Evans EB, Dai H, Wang E, Lee S, Lin H, Lazar AJ, Mills GB, and Lin S. (2015) Nuclear PTEN Tumor-Suppressor Functions through Maintaining Heterochromatin Structure. *Cell Cycle* 14(14):2323–32.

Gotts SJ, Ramot M, Jasmin K, and Martin A. (2019) Altered Resting-State Dynamics in Autism Spectrum Disorder: Causal to the Social Impairment? *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 90:28–36.

Gregorian C, Nakashima J, Le Belle J, Ohab J, Kim R, Liu A, Smith KB, Groszer M, Garcia AD, Sofroniew MV, Carmichael ST, Kornblum HI, Liu X, and Wu H. (2009) Pten Deletion in Adult Neural Stem/Progenitor Cells Enhances Constitutive Neurogenesis. *Journal of Neuroscience* 29(6):1874–86.

Grove J, Ripke S, Als TD, Mattheisen M, Walters RK, Won H, Pallesen J, Agerbo E, Andreassen OA, Anney R, Awashti S, Belliveau R, Bettella F, Buxbaum JD, Bybjerg-Grauholm J, Bækvad-Hansen M, Cerrato F, Chambert K, Christensen JH, Churchhouse C, Dellenvall K, Demontis D, De Rubeis S, Devlin B, Djurovic S, Dumont AL, Goldstein JI, Hansen CS, Hauberg ME, Hollegaard MV, Hope S, Howrigan DP, Huang H, Hultman CM, Klei L, Maller J, Martin J, Martin AR, Moran JL, Nyegaard M, Nærland T, Palmer DS, Palotie A, Pedersen CB, Pedersen MG, dPoterba T, Poulsen JB, Pourcain BS, Qvist P, Rehnström K, Reichenberg A, Reichert J, Robinson EB, Roeder K, Roussos P, Saemundsen E, Sandin S, Satterstrom FK, Davey Smith G, Stefansson H, Steinberg S, Stevens CR, Sullivan PF, Turley P, Walters GB, Xu X; Autism Spectrum Disorder Working Group of the Psychiatric Genomics Consortium; BUPGEN; Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium; 23andMe Research Team, Stefansson K, Geschwind DH, Nordentoft M, Hougaard DM, Werge T, Mors O, Mortensen PB, Neale BM, Daly MJ, and Børglum AD. (2019) Identification of Common Genetic Risk Variants for Autism Spectrum Disorder. *Nature Genetics* 51(3):431–44.

Gu T, Zhang Z, Wang J, Guo J, Shen WH, and Yin Y. (2011) CREB Is a Novel Nuclear Target of PTEN Phosphatase. *Cancer Research* 71(8):2821–25.

Guo H, Duyzend MH, Coe BP, Baker C, Hoekzema K, Gerdts J, Turner TN, Zody MC, Beighley JS, Murali SC, Nelson BJ; University of Washington Center for Mendelian Genomics, Bamshad MJ, Nickerson DA, Bernier RA, Eichler EE. (2018) Genome Sequencing Identifies Multiple Deleterious Variants in Autism Patients with More Severe Phenotypes. *Genetics in Medicine* 21(7):1611–20.

Gupta S, Ellis SE, Ashar FN, Moes A, Bader JS, Zhan J, West AB, and Arking DE. (2014) Transcriptome Analysis Reveals Dysregulation of Innate Immune Response Genes and Neuronal Activity-Dependent Genes in Autism. *Nature Communications* 5(1):5748.

Hansen-Kiss E, Beinkampen S, Adler B, Frazier T, Prior T, Erdman S, Eng C, and Herman G. (2017) A Retrospective Chart Review of the Features of PTEN Hamartoma Tumour Syndrome in Children. *Journal of Medical Genetics* 54(7):471–78.

Haws ME, Jaramillo TC, Espinosa F, Widman AJ, Stuber GD, Sparta DR, Tye KM, Russo SJ, Parada LF, Stavarache M, Kaplitt M, Bonci A, and Powell CM. (2014) PTEN Knockdown Alters Dendritic Spine/Protrusion Morphology, Not Density: Pten and Dendritic Spine Subtypes. *Journal of Comparative Neurology* 522(5):1171–90.

He J, Zhang Z, Ouyang M, Yang F, Hao H, Lamb KL, Yang J, Yin Y, and Shen WH. (2016) PTEN Regulates EG5 to Control Spindle Architecture and Chromosome Congression during Mitosis. *Nature Communications* 7(1):12355.

Heald B, Mester J, Rybicki L, Orloff MS, Burke CA, and Eng C. (2010) Frequent gastrointestinal polyps and colorectal adenocarcinomas in a prospective series of PTEN mutation carriers. *Gastroenterology*. 139(6):1927-33.

Hecht M, Bromberg Y, Rost B. (2014) Better prediction of functional effects for sequence variants. *BMC Genomics* 16.

Hensch TK. (2005) Critical Period Plasticity in Local Cortical Circuits. *Nature Reviews Neuroscience* 6(11):877–88.

Henderson CJ, Ngeow J, Collins MH, Martin LJ, Putnam PE, Abonia JP, Marsolo K, Eng C, and Rothenberg ME. (2014) Increased Prevalence of Eosinophilic Gastrointestinal Disorders in

Pediatric PTEN Hamartoma Tumor Syndromes. Journal of Pediatric Gastroenterology and Nutrition 58(5):553–60.

Henle SJ, Carlstrom LP, Cheever TR, and Henley JR. (2013) Differential Role of PTEN Phosphatase in Chemotactic Growth Cone Guidance. *Journal of Biological Chemistry* 288(29):20837–42.

Herman GE, Butter E, Enrile B, Pastore M, Prior TW, Sommer A. (2007) Increasing knowledge of *PTEN* germline mutations: Two additional patients with autism and macrocephaly. Am J Med Genet 143A: 589-593.

Ho MN, Hill KJ, Lindorfers MA, and Stevens TH. (1993) Isolation of Vacuolar Membrane H+-ATPase-Deficient YeasMt Utants; the VMAS and VMA4 Genes Are Essential for Assembly and Activity of the Vacuolar H+-ATPase. *The Journal of Biological Chemistry* 7: 221-227.

Hobert JA, Embacher R, Mester JL, Frazier TW2, Eng C. (2014) Biochemical screening and PTEN mutation analysis in individuals with autism spectrum disorders and macrocephaly. Eur J Hum Genet 22(2): 273-276.

Hopkins BD, Fine B, Steinbach N, Dendy M, Rapp Z, Shaw J, Pappas K, Yu JS, Hodakoski C, Mense S, Klein J, Pegno S, Sulis ML, Goldstein H, Amendolara B, Lei L, Maurer M, Bruce J, Canoll P, Hibshoosh H, and Parson R (2013) A secreted PTEN phosphatase that enters cells to alter signaling and survival. *Science* 341:399-402.

Hopkins BD, Hodakoski C, Barrows D, Mense SM, and Parsons RE. (2014) PTEN function: the long and the short of it. *Trends Biochem. Sci* 39:183-190.

Hu Y, Li Z, Guo L, Wang L, Zhang L, Cai X, Zhao H, and Zha X. (2007) MAGI-2 Inhibits Cell Migration and Proliferation via PTEN in Human Hepatocarcinoma Cells. *Archives of Biochemistry and Biophysics* 467(1):1–9.

Huang H, Miao L, Yang L, Liang F, Wang Q, Zhuang P, Sun Y, and Hu Y. (2019) AKT-Dependent and -Independent Pathways Mediate PTEN Deletion-Induced CNS Axon Regeneration. *Cell Death & Disease* 10(3):203.

Iossifov I, Ronemus M, Levy D, Wang Z, Hakker I, Rosenbaum J, Yamrom B, Lee YH, Narzisi G, Leotta A, Kendall J, Grabowska E, Ma B, Marks S, Rodgers L, Stepansky A, Troge J, Andrews P, Bekritsky M, Pradhan K, Ghiban E, Kramer M, Parla J, Demeter R, Fulton LL, Fulton RS, Magrini VJ, Ye K, Darnell JC, Darnell RB, Mardis ER, Wilson RK, Schatz MC, McCombie WR, and Wigler M. (2012) De Novo Gene Disruptions in Children on the Autistic Spectrum. *Neuron* 74(2):285–99.

Iossifov I, O'Roak BJ, Sanders SJ, Ronemus M, Krumm N, Levy D, Stessman HA, Witherspoon KT, Vives L, Patterson KE, Smith JD, Paeper B, Nickerson DA, Dea J, Dong S, Gonzalez LE, Mandell JD, Mane SM, Murtha MT, Sullivan CA, Walker MF, Waqar Z, Wei L, Willsey AJ, Yamrom B, Lee YH, Grabowska E, Dalkic E, Wang Z, Marks S, Andrews P, Leotta A, Kendall J, Hakker I, Rosenbaum J, Ma B, Rodgers L, Troge J, Narzisi G, Yoon S, Schatz MC, Ye K, McCombie WR, Shendure J, Eichler EE, State MW, and Wigler M. (2014) The Contribution of de Novo Coding Mutations to Autism Spectrum Disorder. *Nature* 515(7526):216–21.

Ishii N, Maier D, Merlo A, Tada M, Sawamura Y, Diserens AC, and Van Meir EG. (1999) Frequent co-alterations of TP5, p16/CDKN2A, p14ARF, PTEN tumor suppressor genes in human glioma cell lines. *Brain Pathol* 9:469-79.

Iossifov I, O'Roak BJ, Sanders SJ, Ronemus M, Krumm N, Levy D, Stessman HA, Witherspoon KT, Vives L, Patterson KE, Smith JD, Paeper B, Nickerson DA, Dea J, Dong S, Gonzalez LE, Mandell JD, Mane SM, Murtha MT, Sullivan CA, Walker MF, Waqar Z, Wei L, Willsey AJ, Yamrom B, Lee YH, Grabowska E, Dalkic E, Wang Z, Marks S, Andrews P, Leotta A, Kendall

J, Hakker I, Rosenbaum J, Ma B, Rodgers L, Troge J, Narzisi G, Yoon S, Schatz MC, Ye K, McCombie WR, Shendure J, Eichler EE, State MW, and Wigler M. (2014) The Contribution of de Novo Coding Mutations to Autism Spectrum Disorder. *Nature* 515(7526):216–21.

Janecka M, Mill J, Basson MA, Goriely A, Spiers H, Reichenberg A, Schalkwyk L, and Fernandes C. (2017) Advanced Paternal Age Effects in Neurodevelopmental Disorders—Review of Potential Underlying Mechanisms.*Translational Psychiatry* 7(1):e1019–e1019.

Jo M, Chung AY, Yachie N, Seo M, Jeon H, Nam Y, Seo Y, Kim E, Zhong Q, Vidal M, Park HC2 Roth FP, Suk K. (2017) Yeast Genetic Interaction Screen of Human Genes Associated with Amyotrophic Lateral Sclerosis: Identification of MAP2K5 Kinase as a Potential Drug Target. *Genome Research* 27(9):1487–1500.

Jónsson H, Sulem P, Kehr B, Kristmundsdottir S, Zink F, Hjartarson E, Hardarson MT, Hjorleifsson KE, Eggertsson HP, Gudjonsson SA, Ward LD, Arnadottir GA, Helgason EA, Helgason H, Gylfason A, Jonasdottir A, Jonasdottir A, Rafnar T, Frigge M, Stacey SN, Magnusson OT, Thorsteinsdottir U, Masson G, Kong A, Halldorsson BV, Helgason A, Gudbjartsson DF, & Stefansson K. (2017) Parental Influence on Human Germline de Novo Mutations in 1,548 Trios from Iceland. *Nature* 549(7673):519–22.

Jung PP., Christian N, Kay DP, Skupin A, and Linster CL. (2015) Protocols and Programs for High-Throughput Growth and Aging Phenotyping in Yeast. *PLOS ONE* 10(3):e0119807.

Just MA, Cherkassky VL, Keller TA, and Minshew NJ. (2004) Cortical Activation and Synchronization during Sentence Comprehension in High-Functioning Autism: Evidence of Underconnectivity. *Brain* 127(8):1811–21.

Kachroo AH, Laurent JM, Yellman CM, Meyer AG, Wilke CO, Marcotte EM. (2015) Systemic humanization of yeast genes reveals conserved function and genetic modularity. *Science* 348(6237): 921-925.

Kaletta T and Hengartner MO. (2006) Finding function in novel targets: C. elegans as a model organism. *Nat Rev Drug Discov* 5(5):387-98.

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child 2:217*-250.

Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, Collins RL, Laricchia KM, Ganna A, Birnbaum DP, Gauthier LD, Brand H, Solomonson M, Watts NA, Rhodes D, Singer-Berk M, 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, The Genome Aggregation Database Consortium, Neale BM, Daly MJ, and MacArthur DG. (2019) Variation across 141,456 Human Exomes and Genomes Reveals the Spectrum of Loss-of-Function Intolerance across Human Protein-Coding Genes. *bioRxiv* 531210.

Katayama Y, Nishiyama M, Shoji H, Ohkawa Y, Kawamura A, Sato T, Suyama M, Takumi T, Miyakawa T, and Nakayama KI. (2016) CHD8 Haploinsufficiency Results in Autistic-like Phenotypes in Mice. *Nature* 537(7622):675–79.

Kaushik G, Xia Y, Pfau JC, and Thomas MA. (2017) Dysregulation of Autism-Associated Synaptic Proteins by Psychoactive Pharmaceuticals at Environmental Concentrations. *Neuroscience Letters* 661:143–48.

Kazdoba, Tatiana M, Prescott T Leach, Jill L Silverman, and Jacqueline N Crawley. (2014) Modeling Fragile X Syndrome in the Fmr1 Knockout Mouse. *Intractable & Rare Diseases Research* 3(4):118-133.

Keil KP, and Lein PJ. (2016) DNA Methylation: A Mechanism Linking Environmental Chemical Exposures to Risk of Autism Spectrum Disorders? *Environmental Epigenetics* 2(1):dvv012.

Kim Y, Nikoulina SE, Ciaraldi TP, Henry RR, and Kahn BB. (1999) Normal Insulin-Dependent Activation of Akt/Protein Kinase B, with Diminished Activation of Phosphoinositide 3-Kinase, in Muscle in Type 2 Diabetes. *Journal of Clinical Investigation* 104(6):733–41.

Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, and Shendure J. (2014) A general framework for estimating the relative pathogenicity of human genetic variants. *Nature Genetics* 46(3):310-318.

Klauck SM. (2006) Genetics of autism spectrum disorder. *European Journal of Human Genetics* 14: 714-720.

Klei L, Sanders SJ, Murtha MT, Hus V, Lowe JK, Willsey AJ, Moreno-De-Luca D, Yu TW, Geschwind D, Grice DE, Ledbetter DH, Lord C, Mane SM, Martin CL, Martin DM, Morrow EM, Walsh CA, Melhem NM, Chaste P, Sutcliffe JS, State MW, Cook Jr EH, Roeder K, Devlin B. (2012) Common genetic variants, acting additively, are a major source of risk for autism. *Molecular Autism* 3:9.

Klein S, Sharifi-Hannauer P, and Martinez-Agosto JA. (2013) Macrocephaly as a clinical indicator of genetic subtypes in autism. Autism Research 6(1): 51-56.

Klionsky DJ. (2005) The Molecular Machinery of Autophagy: Unanswered Questions. *Journal of Cell Science* 118(1):7–18.

Kogan JH, Gross AK, Featherstone RE, Shin R, Chen Q, Heusner CL, Adachi M, Lin A, Walton NM, Miyoshi S, Miyake S, Tajinda K, Ito H, Siegel SJ, and Matsumoto M. (2015) Mouse Model of Chromosome 15q13.3 Microdeletion Syndrome Demonstrates Features Related to Autism Spectrum Disorder. *Journal of Neuroscience* 35(49):16282–94.

Kotelevets L, van Hengel J, Bruyneel E, Mareel M, van Roy F, and Chastre E. (2005) Implication of the MAGI-1b/PTEN Signalosome in Stabilization of Adherens Junctions and Suppression of Invasiveness. *The FASEB Journal* 19(1):115–17.

Kreis P, Hendricusdottir R, Kay L, Papageorgiou IE, van Diepen M, Mack T, Ryves J, Harwood A, Leslie NR, Kann O, Parsons M, Eickholt BJ. (2013) Phosphorylation of the Actin Binding Protein Drebrin at S647 Is Regulated by Neuronal Activity and PTEN. *PLoS ONE* 8(8):e71957.

Kryndushkin D, and Shewmaker F. (2011) Modeling ALS and FTLD Proteinopathies in Yeast: An Efficient Approach for Studying Protein Aggregation and Toxicity. *Prion* 5(4):250–57.

Kucharczyk R, Dupre S, Avaro S, Haguenauer-Tsapis R, Słonimski PP, and Rytka J. (2000) The novel protein Ccz1p required for vacuolar assembly in Saccharomyces cerevisiae functions in the same transport pathway as Ypt7p. *The Journal of cell science* 113(23):4301-4311.

Kumar S, Reynolds K, Ji Y, Gu R, Rai S, and Zhou CJ. (2019) Impaired Neurodevelopmental Pathways in Autism Spectrum Disorder: A Review of Signaling Mechanisms and Crosstalk. *Journal of Neurodevelopmental Disorders* 11(1):10.

Kwon CH, Zhu X, Zhang J, Knoop LL, Tharp R, Smeyne RJ, Eberhart CG, Burger PC, and Baker SJ (2001) Pten regulates neuronal soma size: a mouse model of Lhermitte-Duclos disease. *Nat Genetics* 29:404-411.

Kwon CH, Zhu X, Zhang J, Baker SJ (2003) mTor is required for hypertrophy of Pten-deficient neuronal soma in vivo. *Proc Natl Acad Sci* 100:12923-12928.

Kwon CH, Luikart BW, Powell CM, Zhou J, Matheny SA, Zhang W, Li Y, Baker SJ, and Parada LF. (2006) Pten regulates neuronal arborization and social interaction in mice. *Neuron* 50, 377-388.

Kwon CH, Zhao D, Chen J, Alcantara S, Li Y, Burns DK, Mason RP, Lee EYP, Wu H, and Parada LF. (2008) Pten Haploinsufficiency Accelerates Formation of High-Grade Astrocytomas." *Cancer Research* 68(9):3286–94.

Lai CH, Chou CY, Chang LY, Liu CS, and Lin W. (2000) Identification of novel human genes evolutionarily conserved in Caenorhabditis elegans by comparative proteomics. *Genome Res*. 10(5):703-13.

Lee BK, Magnusson C, Gardner RM, Blomström A, Newschaffer CJ, Burstyn I, Karlsson H, and Dalman C. (2015) Maternal Hospitalization with Infection during Pregnancy and Risk of Autism Spectrum Disorders. *Brain, Behavior, and Immunity* 44:100–105.

Lee, JO, Yang H, Georgescu MM, Di Cristofano A, Maehama T, Shi Y, Dixon JE, Pandolfi P, and Pavletich NP. (1999) Crystal Structure of the PTEN Tumor Suppressor: Implications for Its Phosphoinositide Phosphatase Activity and Membrane Association. *Cell* 99(3)323-34.

Lee JT, Shan J, Zhong J, Li M, Zhou B, Zhou A, Parsons R, and Gu W. (2013) RFP-Mediated Ubiquitination of PTEN Modulates Its Effect on AKT Activation. *Cell Research* 23(4):552–64.

Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware J, Hill AJ2, Cummings BB, Tukiainen T, Birnbaum DP, Kosmicki JA, Duncan LE, Estrada K, Zhao F, Zou J, Pierce-Hoffman E, Berghout J, Cooper DN, Deflaux N, DePristo M, Do R, Flannick J, Fromer M1, Gauthier L, Goldstein J, Gupta N, Howrigan D, Kiezun A, Kurki MI, Moonshine AL, Natarajan P, Orozco L, Peloso GM, Poplin R, Rivas MA, Ruano-Rubio V, Rose SA, Ruderfer DM, Shakir K, Stenson PD, Stevens C, Thomas BP, Tiao G, Tusie-Luna MT, Weisburd B, Won HH, Yu D, Altshuler DM, Ardissino D, Boehnke M, Danesh J, Donnelly S,

Elosua R, Florez JC, Gabriel SB, Getz G, Glatt SJ, Hultman CM, Kathiresan S, Laakso M, McCarroll S, McCarthy MI, McGovern D, McPherson R, Neale BM, Palotie A, Purcell SM, Saleheen D, Scharf JM, Sklar P, Sullivan PF, Tuomilehto J, Tsuang MT, Watkins HC, Wilson JG, Daly MJ, and MacArthur DG; Exome Aggregation Consortium. (2016) Analysis of Protein-Coding Genetic Variation in 60,706 Humans. *Nature* 536(7616:285-291.

Leslie NR, Kriplani N, Hermida MA, Alvarez-Garcia V, and Wise HM. (2016) The PTEN Protein: Cellular Localization and Post-Translational Regulation. *Biochemical Society Transactions* 44(1):273–78.

Leppa VM, Kravitz SN, Martin CL, Andrieux J, Le Caignec C, Martin-Coignard D, DyBuncio C, Sanders SJ, Lowe JK, Cantor RM, and Geschwind DH. (2016) Rare Inherited and De Novo CNVs Reveal Complex Contributions to ASD Risk in Multiplex Families. *The American Journal of Human Genetics* 99(3):540–54.

Li Z, Vizeacoumar FJ, Bahr S, Li J, Warringer J, Vizeacoumar FS, Min R, Vandersluis B, Bellay J, Devit M, Fleming JA, Stephens A, Haase J, Lin ZY, Baryshnikova A, Lu H, Yan Z, Jin K, Barker S, Datti A, Giaever G, Nislow C, Bulawa C, Myers CL, Costanzo M, Gingras AC, Zhang Z, Blomberg A, Bloom K, Andrews B, and Boone C. (2011) Systematic Exploration of Essential Yeast Gene Function with Temperature-Sensitive Mutants. *Nature Biotechnology* 29(4):361–67.

Liang M, Schwickart M, Schneider AK, Vainshtein I, Del Nagro C, Standifer N, and Roskos LK. (2016) Receptor Occupancy Assessment by Flow Cytometry as a Pharmacodynamic Biomarker in Biopharmaceutical Development: Receptor Occupancy Assessment by Flow Cytometry. *Cytometry Part B: Clinical Cytometry* 90(2):117–27.

Liaw D, Marsh DJ, Li J, Dahia PL, Wang SI, Zheng Z, Bose S, Call KM, Tsou HC, Peacocke M, Eng C, and Parsons R. (1997) Germline Mutations of the PTEN Gene in Cowden Disease, an Inherited Breast and Thyroid Cancer Syndrome. *Nature Genetics* 16(1):64–67.

Lo C, Kays I, Emran F, Lin T, Cvetkovska V, and Chen BE. (2015) Quantification of Protein Levels in Single Living Cells. *Cell Reports* 13(11):2634–44.

Ludwig DL and Bruschi CV. (1991) The 2-micron plasmid as a nonselectable, stable, high copy number yeast plasmid. *Plasmid* 2:81-95.

Luikart BW, Schnell E, Washburn EK, Bensen AL, Tovar KR, and Westbrook GL. (2011) Pten Knockdown In Vivo Increases Excitatory Drive onto Dentate Granule Cells. *Journal of Neuroscience* 31(11):4345–54.

Maddika S, Kavela S, Rani N, Palicharla VR, Pokorny JL, Sarkaria JN, and Chen J. (2011) WWP2 Is an E3 Ubiquitin Ligase for PTEN. *Nature Cell Biology* 13(6):728–33.

Maehama T and Dixon JE. (1998) The Tumor Suppressor, PTEN/MMAC1, Dephosphorylates the Lipid Second Messenger, Phosphatidylinositol 3,4,5-Trisphosphate. *Journal of Biological Chemistry* 273(22):13375–78.

Maehama T, Taylor GS, and Dixon JE. (2001) PTEN and Myotubularin: Novel Phosphoinositide Phosphatases. *Annual Review of Biochemistry* 70(1):247–79.

Maeyama K, Tomioka K, Nagase H, Yoshioka M, Takagi Y, Kato T, Mizobuchi M, Kitayama S, Takada S, Nagai M, Sakakibara N, Nishiyama M, Taniguchi-Ikeda M, Morioka I, Iijima K, and Nishimura N. (2018) Congenital Cytomegalovirus Infection in Children with Autism Spectrum Disorder: Systematic Review and Meta-Analysis. *Journal of Autism and Developmental Disorders* 48(5):1483–91.

Malek M, Kielkowska A, Chessa T, Anderson KE, Barneda D, Pir P, Nakanishi H, Eguchi S, Koizumi A, Sasaki J, Juvin V, Kiselev VY, Niewczas I, Gray A, Valayer A, Spensberger D, Imbert M, Felisbino S, Habuchi T, Beinke S7, Cosulich S, Le Novère N, Sasaki T, Clark J, Hawkins PT, and Stephens LR. (2017) PTEN Regulates PI(3,4)P2 Signaling Downstream of Class I PI3K. *Molecular Cell* 68(3):566-580.e10.

Malia P, Numrich J, Nishimura T, González Montoro A, Stefan CJ, and Ungermann C. (2018) Control of Vacuole Membrane Homeostasis by a Resident PI-3,5-Kinase Inhibitor. *Proceedings of the National Academy of Sciences* 115(18):4684–89.

Mani R, St. Onge RP, Hartman JL, Giaever G, and Roth FP. (2008) Defining Genetic Interaction. *Proceedings of the National Academy of Sciences* 105(9):3461–66.

Marsh DJ, Coulon V, Lunetta KL, Rocca-Serra P, Dahia PL, Zheng Z, Liaw D, Caron S, Duboué B, Lin AY, Richardson AL, Bonnetblanc JM, Bressieux JM, Cabarrot-Moreau A, Chompret A, Demange L, Eeles RA, Yahanda AM, Fearon ER, Fricker JP, Gorlin RJ, Hodgson SV, Huson S, Lacombe D, Eng C, et al. (1998) Mutation spectrum and genotype-phenotype analyses in Cowden disease and Bannayan-Zonana syndrome, two hamartoma syndromes with germline PTEN mutation. *Hum Mol Genet* 7(3):507-15.

Martini M, De Santis MC, Braccini L, Gulluni F, and Hirsch E. (2014) PI3K/AKT Signaling Pathway and Cancer: An Updated Review. *Annals of Medicine* 46(6):372–83.

Matreyek KA, Starita LM, Stephany JJ, Martin B, Chiasson MA, Gray VE, Kircher M, Khechaduri A, Dines JN, Hause RJ, Bhatia S, Evans WE, Relling MV, Yang W, Shendure J, and Fowler DM. (2018) Multiplex Assessment of Protein Variant Abundance by Massively Parallel Sequencing. *Nature Genetics* 50(6):874–82.

McBride KL, Varga EA, Pastore MT, Prior TW, Manickam K, Atkin JF, Herman GE. (2010) Confirmation study of PTEN mutations among individuals with autism or developmental delays/mental retardation and macrocephaly. *Autism Res* 3:137-141.

McKenna B, Koomar T, Vervier K, Kremsreiter J, and Michaelson JJ. (2018) Whole-Genome Sequencing in a Family with Twin Boys with Autism and Intellectual Disability Suggests Multimodal Polygenic Risk. *Molecular Case Studies* 4(6):a003285.

McPheeters ML, Warren Z, Sathe N, Bruzek JL, Krishnaswami S, Jerome RN, and Veenstra-VanderWeele J. (2011) A Systematic Review of Medical Treatments for Children With Autism Spectrum Disorders. *PEDIATRICS* 127(5):e1312–21.

Measday V, Baetz K, Guzzo J, Yuen K, Kwok T, Sheikh B, Ding H, Ueta R, Hoac T, Cheng B, Pot I, Tong A, Yamaguchi-Iwai Y, Boone C, Hieter P, Andrews B. (2005) Systematic yeast synthetic lethal and synthetic dosage lethal screens identify genes required for chromosome segregation. *PNAS* 102(39): 13956-13961.

Meshalkina DA, Kizlyk MN, Kysil EV, Collier AD, Echevarria DJ, Abreu MS, Barcellos LJG, Song C, Warnick JE, Kyzar EJ, and Kalueff AV. (2018) Zebrafish Models of Autism Spectrum Disorder. *Experimental Neurology* 299:207–16.

Mester J and Eng C (2015) PTEN Hamartoma Tumor Syndrome. *Handbook of Clinical Neurology* 132:129–37

Mester JL, Tilot AK, Rybicki LA, Frazier TW, and Eng C. (2011) Analysis of prevalence and degree of macrocephaly in patients with germline PTEN mutations and of brain weight in Pten knock-in murine model. *Eur J Hum Genet* 19(7):763-8.

Mester JL, Ghosh R, Pesaran T, Huether R, Karam R, Hruska KS, Costa HA, Lachlan K, Ngeow J, Barnholtz-Sloan J, Sesock K, Hernandez F, Zhang L, Milko L, Plon SE, Hegde M, and Eng C.

(2018) Gene-Specific Criteria for *PTEN* Variant Curation: Recommendations from the ClinGen PTEN Expert Panel. *Human Mutation* 39(11):1581–92.

Michaillat L and Mayer A. (2013) Identification of Genes Affecting Vacuole Membrane Fragmentation in Saccharomyces Cerevisiae. *PLoS ONE* 8(2):e54160.

Mighell TL, Evans-Dutson S, and O'Roak BJ. (2018) A Saturation Mutagenesis Approach to Understanding PTEN Lipid Phosphatase Activity and Genotype-Phenotype Relationships. *The American Journal of Human Genetics* 102(5):943–55.

Mila M, Alvarez-Mora MI, Madrigal I, and Rodriguez-Revenga L. (2018) Fragile X Syndrome: An Overview and Update of the FMR1 Gene. *Clinical Genetics* 93(2):197–205.

Mitchell DA, Marshall TK, and Deschenes RJ. (1993) Vectors for the inducible overexpression of glutathione S-transferase fusion protein in yeast. *Yeast* 9(7):715-722.

Modabbernia, Amirhossein, Eva Velthorst, and Abraham Reichenberg. (2017) Environmental Risk Factors for Autism: An Evidence-Based Review of Systematic Reviews and Meta-Analyses. *Molecular Autism* 8(1):13.

Montefiori LE, and Nobrega MA. (2019) Gene Therapy for Pathologic Gene Expression." *Science* 363(6424):231–32.

Moriya, Hisao. (2015) Quantitative Nature of Overexpression Experiments. *Molecular Biology of the Cell* 26(22):3932–39.

Myers MP, Pass I, Batty IH, Van der Kaay J, Stolarov JP, Hemmings BA, and Wigler MH, Downes CP, Tonks NK. (1998)The lipid phosphatase activity of PTEN is critical for its tumor suppressor function. *PNAS*. 95:13413-18.

Naguib A, Bencze G, Cho H, Zheng W, Tocilj A, Elkayam E, Faehnle CR, Jaber N, Pratt CP, Chen M, Zong WX, Marks MS, Joshua-Tor L, Pappin DJ, and Trotman LC. (2015) PTEN Functions by Recruitment to Cytoplasmic Vesicles. *Molecular Cell* 58(2):255–68.

Nakatani J, Tamada K, Hatanaka F, Ise S, Ohta H, Inoue K, Tomonaga S, Watanabe Y, Chung YJ, Banerjee R, Iwamoto K, Kato T, Okazawa M, Yamauchi K, Tanda K, Takao K, Miyakawa T, Bradley A, and Takumi T. (2009) Abnormal Behavior in a Chromosome- Engineered Mouse Model for Human 15q11-13 Duplication Seen in Autism. *Cell* 137(7):1235–46.

Nelen MR, van Staveren WC, Peeters EA, Hassel MB, Gorlin RJ, Hamm H, Lindboe CF, Fryns JP, Sijmons RH, Woods DG, Mariman EC, Padberg GW, and Kremer H. (1997) Germline mutations in the PTEN/MMAC1 gene in patients with Cowden disease. *Hum Mol Genet* 6(8):1383-7.

Nelson SB, and Valakh V. (2015) Excitatory/Inhibitory Balance and Circuit Homeostasis in Autism Spectrum Disorders. *Neuron* 87(4):684–98.

Nelson CA 3rd, Zeanah CH, and Fox NA. (2019) How Early Experience Shapes Human Development: The Case of Psychosocial Deprivation. *Neural Plasti.* 1676285.

Numrich J, Péli-Gulli MP, Arlt H, Sardu A, Griffith J, Levine T, Engelbrecht-Vandré S, Reggiori F, De Virgilio C, and Ungermann C. (2015) The I-BAR protein Ivy1 is an effector of the Rab7 GTPase Ypt7 involved in vacuole membrane homeostasis. *J Cell Sci* 128(13):2278-92.

O'Roak BJ, Deriziotis P, Lee C, Vives L, Schwartz JJ, Girirajan S, Karakoc E, Mackenzie AP, Ng SB, Baker C, Rieder MJ, Nickerson DA, Bernier R, Fisher SE, Shendure J, and Eichler EE. (2011) Exome Sequencing in Sporadic Autism Spectrum Disorders Identifies Severe de Novo Mutations. *Nature Genetics* 43(6):585–89.

O'Roak BJ, Vives L, Girirajan S, Karakoc E, Krumm N, Coe BP, Levy R, Ko A, Lee C, Smith JD, Turner EH, Stanaway IB, Vernot B, Malig M, Baker C, Reilly B, Akey JM, Borenstein E, Rieder

MJ, Nickerson DA, Bernier R, Shendure J, and Eichler EE. (2012a) Sporadic Autism Exomes Reveal a Highly Interconnected Protein Network of de Novo Mutations. *Nature* 485(7397):246–50.

O'Roak BJ, Vives L, Fu W, Egertson JD, Stanaway IB, Phelps IG, Carvill G, Kumar A, Lee C, Ankenman K, Munson J, Hiatt JB, Turner EH, Levy R, O'Day DR, Krumm N, Coe BP, Martin BK, Borenstein E, Nickerson DA, Mefford HC, Doherty D, Akey JM, Bernier R, Eichler EE, and Shendure J. (2012b) Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science* 338(6114):1619-22.

Orrico A, Galli L, Buoni S, Orsi A, Vonella G, Sorrentino V. (2009) Novel Pten mutations in neurodevelopmental disorders and macrocephaly. Clin Genet 75: 195-198.

Pandolfo M. (1999) Molecular pathogenesis of Friedrich ataxia. *Arch Neurol* 56:1201-1208.

Pandolfo M. (2008) Friedreich Ataxia. *Arch Neurol* 65(10):8.

Papa A, Wan L, Bonora M, Salmena L, Song MS, Hobbs RM, Lunardi A, Webster K, Ng C, Newton RH, Knoblauch N, Guarnerio J, Ito K, Turka LA, Beck AH, Pinton P, Bronson RT, Wei W, and Pandolfi PP. (2014) Cancer-Associated PTEN Mutants Act in a Dominant-Negative Manner to Suppress PTEN Protein Function. *Cell* 157(3):595–610.

Parisi MA, Dinulos MB, Leppig KA, ybert VP, Eng C, and Hudgins L. (2001) The Spectrum and Evolution of Phenotypic Findings in PTEN Mutation Positive Cases of Bannayan-Riley-Ruvalcaba Syndrome. *Journal of Medical Genetics* 38(1):52–58.

Persico AM, and Napolioni V. (2013) Autism Genetics. *Behavioural Brain Research* 251:95–112.

Pinto D, Pagnamenta AT, Klei L, Anney R, Merico D, Regan R, Conroy J, Magalhaes TR, Correia C, Abrahams BS, Almeida J, Bacchelli E, Bader GD, Bailey AJ, Baird G, Battaglia A, Berney T, Bolshakova N, Bölte S, Bolton PF, Bourgeron T, Brennan S, Brian J, Bryson SE, Carson AR,

Casallo G, Casey J, Chung BH, Cochrane L, Corsello C, Crawford EL, Crossett A, Cytrynbaum C, Dawson G, de Jonge M, Delorme R, Drmic I, Duketis E, Duque F, Estes A, Farrar P, Fernandez BA, Folstein SE, Fombonne E, Freitag CM, Gilbert J, Gillberg C, Glessner JT, Goldberg J, Green A, Green J, Guter SJ, Hakonarson H, Heron EA, Hill M, Holt R, Howe JL, Hughes G, Hus V, Igliozzi R, Kim C, Klauck SM, Kolevzon A, Korvatska O, Kustanovich V, Lajonchere CM, Lamb JA, Laskawiec M, Leboyer M, Le Couteur A, Leventhal BL, Lionel AC, Liu XQ, Lord C, Lotspeich L, Lund SC, Maestrini E, Mahoney W, Mantoulan C, Marshall CR, McConachie H, McDougle CJ, McGrath J, McMahon WM, Merikangas A, Migita O, Minshew NJ, Mirza GK, Munson J, Nelson SF, Noakes C, Noor A, Nygren G, Oliveira G, Papanikolaou K, Parr JR, Parrini B, Paton T, Pickles A, Pilorge M, Piven J, Ponting CP, Posey DJ, Poustka A, Poustka F, Prasad A, Ragoussis J, Renshaw K, Rickaby J, Roberts W, Roeder K, Roge B, Rutter ML, Bierut LJ, Rice JP, Salt J, Sansom K, Sato D, Segurado R, Sequeira AF, Senman L, Shah N, Sheffield VC, Soorya L, Sousa I, Stein O, Sykes N, Stoppioni V, Strawbridge C, Tancredi R, Tansey K, Thiruvahindrapduram B, Thompson AP, Thomson S, Tryfon A, Tsiantis J, Van Engeland H, Vincent JB, Volkmar F, Wallace S, Wang K, Wang Z, Wassink TH, Webber C, Weksberg R, Wing K, Wittemeyer K, Wood S, Wu J, Yaspan BL, Zurawiecki D, Zwaigenbaum L, Buxbaum JD, Cantor RM, Cook EH, Coon H, Cuccaro ML, Devlin B, Ennis S, Gallagher L, Geschwind DH, Gill M, Haines JL, Hallmayer J, Miller J, Monaco AP, Nurnberger JI Jr, Paterson AD, Pericak-Vance MA, Schellenberg GD, Szatmari P, Vicente AM, Vieland VJ, Wijsman EM, Scherer SW, Sutcliffe JS, and Betancur C. (2010) Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* 446:368-372.

Pucci F, Bourgeas R, and Rooman M. (2016) Predicting Protein Thermal Stability Changes upon Point Mutations Using Statistical Potentials: Introducing HoTMuSiC. *Scientific Reports* 6(1):23257.

Pulido R. (2015) PTEN: A Yin-Yang Master Regulator Protein in Health and Disease. *Methods* 77–78:3–10.

Qin JY, Zhang L, Clift KL, Hulur I, Xiang AP, Ren B, and Lahn BT. (2010) Systematic Comparison of Constitutive Promoters and the Doxycycline-Inducible Promoter. *PLoS ONE* 5(5):e10611.

Qiu Z. (2018) Deciphering MECP2 -Associated Disorders: Disrupted Circuits and the Hope for Repair. *Current Opinion in Neurobiology* 48:30–36.

Rabinovsky R, Pochanard P, McNear C, Brachmann SM, Duke-Cohan JS, Garraway LA, and Sellers WR. (2009) P85 Associates with Unphosphorylated PTEN and the PTEN-Associated Complex. *Molecular and Cellular Biology* 29(19):5377–88.

Rahdar M, Inoue T, Meyer T, Zhang J, Vazquez F, and Devreotes P. (2008) A phosphorylation-dependent intramolecular interaction regulates the membrane association and activity of the tumor suppressor PTEN. *PNAS* 106(2):480-485.

Reeb J, Hecht M, Mahlich Y, Bromberg Y, Rost B. (2016) Predicted molecular effects of sequence variants link to system level of disease. *PLOS Com Bio* 12(8).

Rodríguez-Escudero I, Oliver MD, Andrés-Pons A, Molina M, Cid VJ, and Pulido R. (2011) A Comprehensive Functional Analysis of PTEN Mutations: Implications in Tumor- and Autism-Related Syndromes. *Human Molecular Genetics* 20(21):4132–42.

Rodríguez-Escudero I, Fernández-Acero T, Bravo I, Leslie NR, Pulido P, Molina M, and Cid VJ. (2015) Yeast-Based Methods to Assess PTEN Phosphoinositide Phosphatase Activity in Vivo. *Methods* 77–78:172–79.

Rosina E, Battan B, Siracusano M, Di Criscio L, Hollis F, Pacini L, Curatolo P, and Bagni C. (2019) Disruption of MTOR and MAPK Pathways Correlates with Severity in Idiopathic Autism. *Translational Psychiatry* 9(1):50.

Ross AH, and Gericke A. (2009) Phosphorylation Keeps PTEN Phosphatase Closed for Business. *Proceedings of the National Academy of Sciences* 106(5):1297–98.

Roux KJ, Kim DI, and Burke B. (2013) BioID: A Screen for Protein-Protein Interactions. *Current Protocols in Protein Science*. 74:19.23.

Rubenstein JLR, and Merzenich MM. (2003) Model of Autism: Increased Ratio of Excitation/Inhibition in Key Neural Systems: Model of Autism. *Genes, Brain and Behavior* 2(5):255–67.

Rudge SA, Anderson DM, and Emr SD. (2004) Vacuole Size Control: Regulation of PtdIns(3,5)P_2 Levels by the Vacuole-Associated Vac14-Fig4 Complex, a PtdIns(3,5)P_2 -Specific Phosphatase. *Molecular Biology of the Cell* 15(1):24–36.

Sanchack KE and Thomas CA. (2016) Autism Spectrum Disorder: Primary Care Principles." *Autism Spectrum Disorder* 94(12):9.

Sansal I, and Sellers WR. (2004) The Biology and Clinical Relevance of the *PTEN* Tumor Suppressor Pathway. *Journal of Clinical Oncology* 22(14):2954–63.

Sato TK, Darsow T, and Emr SD. (1998) Vam7p, a SNAP-25-Like Molecule, and Vam3p, a Syntaxin Homolog, Function Together in Yeast Vacuolar Protein Trafficking. *Molecular and Cellular Biology* 18(9):5308–19.

Schaaf CP, Sabo A, Sakai Y, Crosby J, Muzny D, Hawes A, Lewis L, Akbar H, Varghese R, Boerwinkle E, Gibbs RA, and Zoghbi HY. (2011) Oligogenic heterozygosity in individuals with high-functioning autism spectrum disorders. *Hum Mol Genet* 20(17):3366-75.

Schimmöller F and Riezman H. (1993) Involvement of Ypt7p, a Small GTPase, in Traffic from Late Endosome to the Vacuole in Yeast. *Journal of Cell Science* 106:823-830.

Schmeisser K, and Parker JA. (2018) Worms on the Spectrum - C. Elegans Models in Autism Research. *Experimental Neurology* 299:199–206.

Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, Walsh T, Yamrom B, Yoon S, Krasnitz A, Kendall J, Leotta A, Pai D, Zhang R, Lee Y, Hicks J, Spence SJ, Lee AT, Puura K, Lehtimaki T, Ledbetter D, Gregersen PK, Bregman J, Sutcliffe JS, Jobanputra V, Chung W, Warburton D, King M, Skuse D, Geschwind DH, Gilliam TC, Ye K, Wigler M. (2007). Strong association of copy number mutations with autism. *Science* 316(5823):445-449.

Serezani CH, Kane S, Medeiros AI, Cornett AM, Kim SH, Marques MM, Lee SP, Lewis C, Bourdonnay E, Ballinger MN, White ES, and Peters-Golden M. (2012) PTEN Directly Activates the Actin Depolymerization Factor Cofilin-1 During PGE2-Mediated Inhibition of Phagocytosis of Fungi. *Science Signaling* 5(210):ra12.

Shaye DD and Greenwald I. (2011) OrthoList: a compendium of C. elegans genes with human orthologs. *PLoS One* 6(5):e20085.

Shen WH, Balajee AS, Wang J, Wu H, Eng C, Pandolfi PP, and Yin Y. (2007) Essential Role for Nuclear PTEN in Maintaining Chromosomal Integrity. *Cell* 128(1):157–70.

Shi Y, Paluch BE, Wang X, and Jiang X. (2012) PTEN at a Glance. *Journal of Cell Science* 125(20):4687–92.

Shi Y, Wang J, Chandarlapaty S, Cross J, Thompson C, Rosen N, and Jiang X. (2014) PTEN Is a Protein Tyrosine Phosphatase for IRS1. *Nature Structural & Molecular Biology* 21(6):522–27.

Shin JJ, Aftab Q, Austin Q, McQueen JA, Poon T, Li SC, Young BP, Roskelley CD, Loewen CJ. (2016) Systematic identification of genes involved in metabolic acid stress resistance in yeast and their potential as cancer targets. *Dis Model Mech* 9:1039-1049.

Sim N, Kumar P, Hu J, Henikoff S, Schneider G, and Ng PC. (2012) SIFT Web Server: Predicting Effects of Amino Acid Substitutions on Proteins." *Nucleic Acids Research* 40(W1):W452–57.

Singh I, Pass R, Togay SO, Rodgers JW, and Hartman JL. (2009) Stringent Mating-Type-Regulated Auxotrophy Increases the Accuracy of Systematic Genetic Interaction Screens with *Saccharomyces Cerevisiae* Mutant Arrays. *Genetics* 181(1):289–300.

Smith IN, and Briggs JM. (2016) Structural Mutation Analysis of PTEN and Its Genotype-Phenotype Correlations in Endometriosis and Cancer: Structural Analysis of PTEN Mutations. *Proteins: Structure, Function, and Bioinformatics* 84(11):1625–43.

Smith IN, Thacker S, Jaini R, and Eng C. (2019) Dynamics and Structural Stability Effects of Germline *PTEN* Mutations Associated with Cancer versus Autism Phenotypes. *Journal of Biomolecular Structure and Dynamics* 37(7):1766–82.

Song MS, Salmena L, and Pandolfi PP. (2012) The Functions and Regulation of the PTEN Tumour Suppressor. *Nature Reviews Molecular Cell Biology* 13(5):283–96.

Spinelli L, Black FM, Berg JN, Eickholt BJ, and Leslie NR. (2015) Functionally distinct groups of inherited PTEN mutations in autism and tumour syndromes. *J Med Genet* 52(2):128-34.

Subramanian M, Timmerman CK, Schwartz JL, Pham DL, and Meffert MK. (2015) Characterizing Autism Spectrum Disorders by Key Biochemical Pathways. *Frontiers in Neuroscience* 9.

Sukoff Rizzo SJ, and Silverman JL. (2016) Methodological Considerations for Optimizing and Validating Behavioral Assays: Optimizing and Validating Behavioral Assays. *Current Protocols in Mouse Biology* 2016 edition, John Wiley & Sons, Inc., New Jersey.

Sumitomo M, Iwase A, Zheng R, Navarro D, Kaminetzky D, Shen R, Georgescu MM, and Nanus DM. (2004) Synergy in Tumor Suppression by Direct Interaction of Neutral Endopeptidase with PTEN. *Cancer Cell* 5(1):67–78.

Sun S, Yang F, Tan G, Costanzo M, Oughtred R, Hirschman J, Theesfeld CL, Bansal P, Sahni N, Yi S, Yu A, Tyagi T, Tie C, Hill DE, Vidal M, Andrews BJ, Boone C, Dolinski K, and Roth FP. (2016) An Extended Set of Yeast-Based Functional Assays Accurately Identifies Human Disease Mutations. *Genome Research* 26(5):670–80.

Sztainberg Y, and Zoghbi HY. (2016) Lessons Learned from Studying Syndromic Autism Spectrum Disorders. *Nature Neuroscience* 19(11):1408–17.

Stepanenko AA, and Dmitrenko VV. (2015) HEK293 in Cell Biology and Cancer Research: Phenotype, Karyotype, Tumorigenicity, and Stress-Induced Genome-Phenotype Evolution. *Gene* 569(2):182–90.

Swierczek NA, Giles AC, Rankin CH, and Kerr RA. (2011) High-Throughput Behavioral Analysis in C. Elegans. *Nature Methods* 8(7):592–98.

Takahashi Y, Morales FC, Kreimann EL, and Georgescu MM. (2006) PTEN Tumor Suppressor Associates with NHERF Proteins to Attenuate PDGF Receptor Signaling. *The EMBO Journal* 25(4):910–20.

Takeuchi K, Gertner MJ, Zhou J, Parada LF, Bennett MVL, Zukin RS. (2013) Dysregulation of synaptic plasticity precedes appearance of morphological defects in a Pten conditional knockout mouse model of autism. *PNAS* 110(12):4738-4743.

Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, Boutselakis H, Cole CG, Creatore C, Dawson E, Fish P, Harsha B, Hathaway C, Jupe SC, Kok CY, Noble K, Ponting L, Ramshaw CC, Rye CE, Speedy HE, Stefancsik R, Thompson SL, Wang S, Ward S, Campbell PJ, and

Forbes SA. (2019) COSMIC: The Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Research* 47(D1):D941–47.

Tatsukawa T, Raveau M, Ogiwara I, Hattori S, Miyamoto H, Mazaki E, Itohara S, Miyakawa T, Montal M, and Yamakawa K. (2019) Scn2a Haploinsufficient Mice Display a Spectrum of Phenotypes Affecting Anxiety, Sociability, Memory Flexibility and Ampakine CX516 Rescues Their Hyperactivity. *Molecular Autism* 10(1):15.

Thomas P and Smart TG. (2005) HEK293 Cell Line: A Vehicle for the Expression of Recombinant Proteins." *Journal of Pharmacological and Toxicological Methods* 51(3):187–200.

Tibarewal P, Zilidis G, Spinelli L, Schurch N, Maccario H, Gray A, Perera NM, Davidson L, Barton GJ, and Leslie NR. (2012) PTEN Protein Phosphatase Activity Correlates with Control of Gene Expression and Invasion, a Tumor-Suppressing Phenotype, But Not with AKT Activity. *Science Signaling* 5(213):ra18–ra18.

Tong AHY, Evangelista M, Parsons AB, Xu H, Bader GD, Page N, Roninson M, Raghibizadeh S, Hogue CWV, Bussey H, Andrews B, Tyers M, Boone C. (2001) Systemic genetic analysis with ordered arrays of yeast deletion mutants. *Science* 294:2364-2368.

Ungermann C, Price A, and Wickner W. (2000) A New Role for a SNARE Protein as a Regulator of the Ypt7/Rab-Dependent Stage of Docking. *Proceedings of the National Academy of Sciences* 97(16):8889–91.

van Diepen MT and Eickholt BJ. (2008) Function of PTEN during the formation and maintenance of neuronal circuits in the brain. *Developmental Neuroscience* 30: 59-64.

Van Themsche C, Leblanc V, Parent S, and Asselin E. (2009) X-Linked Inhibitor of Apoptosis Protein (XIAP) Regulates PTEN Ubiquitination, Content, and Compartmentalization. *Journal of Biological Chemistry* 284(31):20462–66.

Vanderver A, Tonduti D, Kahn I, Schmidt J, Medne L, Vento J, Chapman KA, Lanpher B, Pearl P, Gropman A, Lourenco C, Bamforth JS, Sharpe C, Pineda M, Schallner J, Bodamer O, Orcesi S, Oberstein SA, Sistermans EA, Yntema HG, Bonnemann , Waldman AT, van der Knaap MS. (2014) Characteristic brain magnetic resonance imaging pattern in patients with macrocephaly and PTEN mutations. *Am J Med Genet* 164A(3):627-633.

Varga EA, Pastore M, Prior T, Herman GE, McBride KL. (2009) The prevalence of PTEN mutations in a clinical pediatric cohort with autism spectrum disorders, developmental delay, and macrocephaly. *Genet Med* 11(2):111-117.

Varghese M, Keshav N, Jacot-Descombes S, Warda T, Wicinski B, Dickstein DL, Harony-Nicolas H, De Rubeis S, Drapeau E, Buxbaum JD, and Hof PR. (2017) Autism Spectrum Disorder: Neuropathology and Animal Models. *Acta Neuropathologica* 134(4):537–66.

Vazquez F, Grossman SR, Takahashi Y, Rokas MV, Nakamura N, and Sellers WR. (2001) Phosphorylation of the PTEN tail acts as an inhibitory switch by preventing its recruitment into a protein complex. *JBC* (276)52:48627-48630.

Vithayathil J, Pucilowska J, and Landreth GE. (2018) ERK/MAPK Signaling and Autism Spectrum Disorders. *Progress in Brain Research* 241:63–112.

Vivanco I, and Sawyers CL. (2002) The Phosphatidylinositol 3-Kinase–AKT Pathway in Human Cancer. *Nature Reviews Cancer* 2(7):489–501.

Volpe J. (2000) Overview: Normal and Abnormal Human Brain development. *Mental Retardation and Developmental Disabilities Research Reviews*. 6:1-5.

Walberg MW. (2000) Applicability of yeast genetics to neurologic disease. *JAMA* 57:1129- 1134.

Wang X, Trotman LC, Koppie T, Alimonti A, Chen Z, Gao Z, Wang J, Erdjument-Bromage H, Tempst P, Cordon-Cardo C, Pandolfi PP, and Jiang X. (2007) NEDD4-1 Is a Proto-Oncogenic Ubiquitin Ligase for PTEN. *Cell* 128(1):129–39.

Wang L, Zhou K, Fu Z, Yu D, Huang H, Zang X, and Mo X. (2017) Brain Development and Akt Signaling: The Crossroads of Signaling Pathway and Neurodevelopmental Diseases. *Journal of Molecular Neuroscience* 61(3):379–84.

Wegiel J, Flory M, Kuchna I, Nowicki K, Ma SY, Imaki H, Wegiel J, Cohen IL, London E, Wisniewski T, and Brown WT. (2014) Stereological study of the neuronal number and volume of 38 brain subdivisions of subjects diagnosed with autism reveals significant alterations restricted to the striatum, amygdala and cerebellum. *Acta Neuropathol Commun* 2:141.

Widowati EW, Ernst S, Hausmann R, Müller-Newen G, Becker W. (2018) Functional characterization of DYRK1A missense variants associated with a syndromic form of intellectual deficiency and autism. *Biol Open* 7(4)pii: bio032862.

Wong WR, Brugman KI, Maher S, Oh JY, Howe K, Kato M, and Sternberg PW. (2019) Autism-associated missense genetic variants impact locomotion and neurodevelopment in Caenorhabditis elegans. *Hum Mol Genet* 28(13):2271-2281.

Woodbury-Smith, and Scherer SW. (2018) Progress in the Genetics of Autism Spectrum Disorder. *Developmental Medicine & Child Neurology* 60(5):445–51.

Wozniak DJ, Kajdacsy-Balla A, Macias V, Ball-Kell S, Zenner ML, Bie W, and Tyner AL. (2017) PTEN Is a Protein Phosphatase That Targets Active PTK6 and Inhibits PTK6 Oncogenic Signaling in Prostate Cancer. *Nature Communications* 8(1):1508.

Xiang AH, Wang X, Martinez MP, Walthall JC, Curry ES, Page K, Buchanan TA, Coleman KJ, and Getahun D. (2015) Association of Maternal Diabetes with Autism in Offspring. *JAMA* 313(14):1425.

Xu W, Yang Z, Xie C, Zhu Y, Shu X, Zhang Z, Li N, Chai N, Zhang S, Wu K, Nie Y, and Lu N. (2018) PTEN Lipid Phosphatase Inactivation Links the Hippo and PI3K/Akt Pathways to Induce Gastric Tumorigenesis. *Journal of Experimental & Clinical Cancer Research* 37(1):198.

Xue Y, Chen Y, Ayub Q, Huang N, Ball EV, Mort M, Phillips AD, Shaw K, Stenson PD, Cooper DN, Tyler-Smith C, and 1000 Genomes Project Consortium. (2012) Deleterious- and Disease-Allele Prevalence in Healthy Individuals: Insights from Current Predictions, Mutation Databases, and Population-Scale Resequencing." *American Journal of Human Genetics* 91(6):1022–32.

Yang C, Li J, Wu Q, Yang X, Huang AY, Zhang J, Ye AY, Dou Y, Yan L, Zhou W, Kong L, Wang M, Ai C, Yang D, and Wei L. (2018) AutismKB 2.0: A Knowledgebase for the Genetic Evidence of Autism Spectrum Disorder. *Database* 1-8.

Yang Z and Klionsky DJ. (2009) An Overview of the Molecular Mechanism of Autophagy. In: Levine B, Yoshimori T, and Deretic V(eds) Autophagy in Infection and Immunity. Current Topics in Microbiology and Immunology. Volume 335 Springer, Berlin Heidelberg.

Yehia L, Ngeow J, and Eng C. (2019) PTEN-Opathies: From Biological Insights to Evidence-Based Precision Medicine. *Journal of Clinical Investigation* 129(2):452–64.

Yi S, Lin S, Li Y, Zhao W, Mills GB, Sahni N. (2017) Functional variomics and network perturbation: connecting genotype to phenotype in cancer. *Nat Rev Genet* 18(7):395-410.

Young BP and Loewen CJR. (2013) Balony: a software package for analysis of data generated by synthetic genetic array experiments. *BMC Bioinformatics* 14:354

Zhang XC, Piccini A, Myers MP, Van Aelst L, and Tonks NK. (2012) Functional Analysis of the Protein Phosphatase Activity of PTEN. *Biochemical Journal* 444 (3):457–64.

Zoghbi HY. "MeCP2 Dysfunction in Humans and Mice. *Journal of Child Neurology* 20(9): 736-740.

Zori RT, Marsh DJ, Graham GE, Marliss EB, and Eng C. (1998) Germline PTEN Mutation in a Family with Cowden Syndrome and Bannayan-Riley-Ruvalcaba Syndrome. *American Journal of Medical Genetics* 80(4):399–40

Appendix: Plasmids acquired and created

Below is a full list of all plasmids acquired and created for all model systems including yeast, HEK293, fly, worm, and rat. Plasmids were made for experimentation in all five model systems. *Drosophila*, HEK293 and worm utilized pDEST vectors while a pCAG-GFP construct was obtained from Addgene (catalog# 11150) for rat hippocampal culture experiments. The GFP from this plasmid was excised and replaced with human *PTEN* with four synonymous mutations to prevent shRNA binding (c.573G>T, c.579G>A, c.585T>C and c.591G>A) and three N-terminal human influenza hemagglutinin (HA) tags. Creation of all 97 variants was done in the pENTR clone using site directed mutagenesis (SDM). SDM was also used to create the 22 variants in the shRNA-resistant pCAG rat vector. In brief, SDM was performed using forward and reverse primer sequences (up to 60 nucleic acids long) which were designed to overlap the desired mutation. Polymerase chain reactions (PCRs) were completed using QuikChange Site-Directed Mutagenesis and Pfu polymerase (Agilent, catalog# 600390). Each variant was sequence validated. Once validated, variants were moved into assay-specific vectors using Gateway cloning. For yeast assays, PTEN was amplified from the pENTR vector with overhangs for the pEGH1 (Mitchell et al., 1993) yeast vector.

The destination vector used for *Drosophila* experiments was a custom *10xUAS-attB* vector (Bischof et al., 2013) which could then be integrated into the fly genome with protein expression induced via a cell-specific GAL4 promoter. For *C. elegans*, a pDEST-*aex-3p* destination vector for pan-neuronal expression (obtained from Dr. Hidehito Kuroyanagi, Tokyo Medical and Dental University) was used to generate the *aex-3p::PTEN::unc-54 UTR* rescue construct. The HEK293 destination vector (pCAG-sfGFP-P2A-mtagRFPt) was used for HEK293

cell experiments (mtagRFPt and super fluorescent GFP(sfGFP) were gifts from Dr. Brian Chen, McGill University).

Plasmids Acquired

Description	Vector	Origin
PTEN cDNA	pCMV-PTEN	Addgene (28298)
Entry vector	pCR8/GW/TOPO (pENTR)	Thermo Scientific (K250020)
Yeast vector	pEGH1	Mitchell DA, Marshall TK, and Deschenes RJ. (1993) Vectors for the inducible overexpression of glutathione S-transferase fusion protein in yeast. *Yeast* 9(7):715-722.
HEK293 destination vector	pCAG-sfGFP-P2A-mtagRFPt	mtagRFP-T and sfGFP presents from Dr. Brian Chen (McGill University)
C. elegans destination vector	pDEST-*aex-3p*	obtained from Dr. Hidehito Kuroyanagi (Tokyo Medical and Dental University)
Drosophila destination vector	pUASg-attB	obtained from Dr. Douglas Allan (UBC)
Rat hippocampal culture vector	pCAG-GFP	Addgene (11150)

PTEN Plasmids Created	Entry Vector	Yeast Vector	HEK293 vector	*Drosophila* Vector	C. elegans Vector	Rat Vector
Empty Vector	pENTR	pEGH1	pCAG-mtagRFPt-P2A-sfGFP	pUASg-attB	pDEST-aex-3p	pCAG-GFP
WT-PTEN	pENTR-PTEN	pEGH1-PTEN	pCAG-mtagRFPt-P2A-PTEN	pUASg-attB-PTEN	pDEST-aex-3p-PTEN	pCAG-rPTEN-HA
4A	pENTR-4A	pEGH1-4A	pCAG-mtagRFPt--P2A-GFP-4A	pUASg-attB-4A		
A121E	pENTR-A121E	pEGH1-A121E	pCAG-mtagRFPt--P2A-GFP-A121E	pUASg-attB-A121E		
A121P	pENTR-A121P	pEGH1-A121P	pCAG-mtagRFPt--P2A-GFP-A121P	pUASg-attB-A121P		
A126D	pENTR-A126D	pEGH1-A126D	pCAG-mtagRFPt--P2A-GFP-A126D	pUASg-attB-A126D		
A126P	pENTR-A126P	pEGH1-A126P	pCAG-mtagRFPt--P2A-GFP-A126P	pUASg-attB-A126P		
A151P	pENTR-A151P	pEGH1-A151P	pCAG-mtagRFPt--P2A-GFP-A151P	pUASg-attB-A151P		
A309S	pENTR-A309S	pEGH1-A309S	pCAG-mtagRFPt--P2A-GFP-A309S	pUASg-attB-A309S		
A34D	pENTR-A34D	pEGH1-A34D	pCAG-mtagRFPt--P2A-GFP-A34D	pUASg-attB-A34D		
A34P	pENTR-A34P	pEGH1-A34P	pCAG-mtagRFPt--P2A-GFP-A34P	pUASg-attB-A34P		
A79T	pENTR-A79T	pEGH1-A79T	pCAG-mtagRFPt--P2A-GFP-A79T	pUASg-attB-A79T	pDEST-aex-3p-A79T	pCAG-rA79T-3xHA
C124R	pENTR-C124R	pEGH1-C124R	pCAG-mtagRFPt--P2A-GFP-C124R	pUASg-attB-C124R		

PTEN Plasmids Created	Entry Vector	Yeast Vector	HEK293 vector	*Drosophila* Vector	C. elegans Vector	Rat Vector
C124S	pENTR-C124S	pEGH1-C124S	pCAG-mtagRFPt--P2A-GFP-C124S	pUASg-attB-C124S	pDEST-aex-3p-C124S	pCAG-rC124S-3xHA
C124S-4A	pENTR-C124S-4A	pEGH1-C124S-4A	pCAG-mtagRFPt--P2A-GFP-C124S-4A	pUASg-attB-C124S-4A		
C211W	pENTR-C211W	pEGH1-C211W	pCAG-mtagRFPt--P2A-GFP-C211W	pUASg-attB-C211W		
D107G	pENTR-D107G	pEGH1-D107G	pCAG-mtagRFPt--P2A-GFP-D107G	pUASg-attB-D107G		
D107V	pENTR-D107V	pEGH1-D107V	pCAG-mtagRFPt--P2A-GFP-D107V	pUASg-attB-D107V		
D22E	pENTR-D22E	pEGH1-D22E	pCAG-mtagRFPt--P2A-GFP-D22E	pUASg-attB-D22E		
D252G	pENTR-D252G	pEGH1-D252G	pCAG-mtagRFPt--P2A-GFP-D252G	pUASg-attB-D252G		
D268E	pENTR-D268E	pEGH1-D268E	pCAG-mtagRFPt--P2A-GFP-D268E	pUASg-attB-D268E	pDEST-aex-3p-D268E	pCAG-rD268E-3xHA
D326N	pENTR-D326N	pEGH1-D326N	pCAG-mtagRFPt--P2A-GFP-D326N	pUASg-attB-D326N	pDEST-aex-3p-D326N	pCAG-rD326N-3xHA
D92N	pENTR-D92N	pEGH1-D92N	pCAG-mtagRFPt--P2A-GFP-D92N	pUASg-attB-D92N	pDEST-aex-3p-D92N	pCAG-rD92N-3xHA
E157G	pENTR-E157G	pEGH1-E157G	pCAG-mtagRFPt--P2A-GFP-E157G	pUASg-attB-E157G		
E256K	pENTR-E256K	pEGH1-E256K	pCAG-mtagRFPt--P2A-GFP-E256K	pUASg-attB-E256K		
E307Q	pENTR-E307Q	pEGH1-E307Q	pCAG-mtagRFPt--P2A-GFP-E307Q	pUASg-attB-E307Q		
E73K	pENTR-E73K	pEGH1-E73K	pCAG-mtagRFPt--P2A-GFP-E73K	pUASg-attB-E73K		
F241S	pENTR-F241S	pEGH1-F241S	pCAG-mtagRFPt--P2A-GFP-F241S	pUASg-attB-F241S		
F279I	pENTR-F279I	pEGH1-F279I	pCAG-mtagRFPt--P2A-GFP-F279I	pUASg-attB-F279I		
F279L	pENTR-F279L	pEGH1-F279L	pCAG-mtagRFPt--P2A-GFP-F279L	pUASg-attB-F279L		
F56C	pENTR-F56C	pEGH1-F56C	pCAG-mtagRFPt--P2A-GFP-F56C	pUASg-attB-F56C		
G127R	pENTR-G127R	pEGH1-G127R	pCAG-mtagRFPt--P2A-GFP-G127R	pUASg-attB-G127R		
G129E	pENTR-G129E	pEGH1-G129E	pCAG-mtagRFPt--P2A-GFP-G129E	pUASg-attB-G129E		
G129R	pENTR-G129R	pEGH1-G129R	pCAG-mtagRFPt--P2A-GFP-G129R	pUASg-attB-G129R		
G132D	pENTR-G132D	pEGH1-G132D	pCAG-mtagRFPt--P2A-GFP-G132D	pUASg-attB-G132D	pDEST-aex-3p-G132D	pCAG-rG132D-3xHA
G36E	pENTR-G36E	pEGH1-G36E	pCAG-mtagRFPt--P2A-GFP-G36E	pUASg-attB-G36E		
G44D	pENTR-G44D	pEGH1-G44D	pCAG-mtagRFPt--P2A-GFP-G44D	pUASg-attB-G44D	pDEST-aex-3p-G44D	pCAG-rG44D-3xHA
H118P	pENTR-H118P	pEGH1-H118P	pCAG-mtagRFPt--P2A-GFP-H118P	pUASg-attB-H118P		
H123Q	pENTR-H123Q	pEGH1-H123Q	pCAG-mtagRFPt--P2A-GFP-H123Q	pUASg-attB-H123Q	pDEST-aex-3p-H123Q	pCAG-rH123Q-3xHA

PTEN Plasmids Created	Entry Vector	Yeast Vector	HEK293 vector	*Drosophila* Vector	C. elegans Vector	Rat Vector
H123Y	pENTR-H123Y	pEGH1-H123Y	pCAG-mtagRFPt--P2A-GFP-H123Y	pUASg-attB-H123Y		
H93Q	pENTR-H93Q	pEGH1-H93Q	pCAG-mtagRFPt--P2A-GFP-H93Q	pUASg-attB-H93Q		
H93R	pENTR-H93R	pEGH1-H93R	pCAG-mtagRFPt--P2A-GFP-H93R	pUASg-attB-H93R	pDEST-aex-3p-H93R	pCAG-rH93R-3xHA
H93Y	pENTR-H93Y	pEGH1-H93Y	pCAG-mtagRFPt--P2A-GFP-H93Y	pUASg-attB-H93Y		
I101T	pENTR-I101T	pEGH1-I101T	pCAG-mtagRFPt--P2A-GFP-I101T	pUASg-attB-I101T	pDEST-aex-3p-I101T	pCAG-rI101T-3xHA
I135T	pENTR-I135T	pEGH1-I135T	pCAG-mtagRFPt--P2A-GFP-I135T	pUASg-attB-I135T		
I135V	pENTR-I135V	pEGH1-I135V	pCAG-mtagRFPt--P2A-GFP-I135V	pUASg-attB-I135V		
I203V	pENTR-I203V	pEGH1-I203V	pCAG-mtagRFPt--P2A-GFP-I203V	pUASg-attB-I203V		
I400V	pENTR-I400V	pEGH1-I400V	pCAG-mtagRFPt--P2A-GFP-I400V	pUASg-attB-I400V		
K322E	pENTR-K322E	pEGH1-K322E	pCAG-mtagRFPt--P2A-GFP-K322E	pUASg-attB-K322E		
K330E	pENTR-K330E	pEGH1-K330E	pCAG-mtagRFPt--P2A-GFP-K330E	pUASg-attB-K330E		
K342N	pENTR-K342N	pEGH1-K342N	pCAG-mtagRFPt--P2A-GFP-K342N	pUASg-attB-K342N		
K402N	pENTR-K402N	pEGH1-K402N	pCAG-mtagRFPt--P2A-GFP-K402N	pUASg-attB-K402N		
K6E	pENTR-K6E	pEGH1-K6E	pCAG-mtagRFPt--P2A-GFP-K6E	pUASg-attB-K6E	pDEST-aex-3p-K6E	pCAG-rK6E-3xHA
K6I	pENTR-K6I	pEGH1-K6I	pCAG-mtagRFPt--P2A-GFP-K6I	pUASg-attB-K6I	pDEST-aex-3p-K6I	pCAG-rK6I-3xHA
L295V	pENTR-L295V	pEGH1-L295V	pCAG-mtagRFPt--P2A-GFP-L295V	pUASg-attB-L295V		
L345V	pENTR-L345V	pEGH1-L345V	pCAG-mtagRFPt--P2A-GFP-L345V	pUASg-attB-L345V		
L70V	pENTR-L70V	pEGH1-L70V	pCAG-mtagRFPt--P2A-GFP-L70V	pUASg-attB-L70V		
M134I	pENTR-M134I	pEGH1-M134I	pCAG-mtagRFPt--P2A-GFP-M134I	pUASg-attB-M134I		
M134T	pENTR-M134T	pEGH1-M134T	pCAG-mtagRFPt--P2A-GFP-M134T	pUASg-attB-M134T		
M198I	pENTR-M198I	pEGH1-M198I	pCAG-mtagRFPt--P2A-GFP-M198I	pUASg-attB-M198I		
M1I	pENTR-M1I	pEGH1-M1I	pCAG-mtagRFPt--P2A-GFP-M1I	pUASg-attB-M1I		
M35V	pENTR-M35V	pEGH1-M35V	pCAG-mtagRFPt--P2A-GFP-M35V	pUASg-attB-M35V		
N117S	pENTR-N117S	pEGH1-N117S	pCAG-mtagRFPt--P2A-GFP-N117S	pUASg-attB-N117S		
N12T	pENTR-N12T	pEGH1-N12T	pCAG-mtagRFPt--P2A-GFP-N12T	pUASg-attB-N12T		
N228S	pENTR-N228S	pEGH1-N228S	pCAG-mtagRFPt--P2A-GFP-N228S	pUASg-attB-N228S		
N262S	pENTR-N262S	pEGH1-N262S	pCAG-mtagRFPt--P2A-GFP-N262S	pUASg-attB-N262S		

PTEN Plasmids Created	Entry Vector	Yeast Vector	HEK293 vector	*Drosophila* Vector	C. elegans Vector	Rat Vector
N276S	pENTR-N276S	pEGH1-N276S	pCAG-mtagRFPt--P2A-GFP-N276S	pUASg-attB-N276S		
N340D	pENTR-N340D	pEGH1-N340D	pCAG-mtagRFPt--P2A-GFP-N340D	pUASg-attB-N340D		
N340H	pENTR-N340H	pEGH1-N340H	pCAG-mtagRFPt--P2A-GFP-N340H	pUASg-attB-N340H		
N356D	pENTR-N356D	pEGH1-N356D	pCAG-mtagRFPt--P2A-GFP-N356D	pUASg-attB-N356D		
N356H	pENTR-N356H	pEGH1-N356H	pCAG-mtagRFPt--P2A-GFP-N356H	pUASg-attB-N356H		
P246L	pENTR-P246L	pEGH1-P246L	pCAG-mtagRFPt--P2A-GFP-P246L	pUASg-attB-P246L		
P354Q	pENTR-P354Q	pEGH1-P354Q	pCAG-mtagRFPt--P2A-GFP-P354Q	pUASg-attB-P354Q	pDEST-aex-3p-P354Q	pCAG-rP354Q-3xHA
P357S	pENTR-P357S	pEGH1-P357S	pCAG-mtagRFPt--P2A-GFP-P357S	pUASg-attB-P357S		
P38H	pENTR-P38H	pEGH1-P38H	pCAG-mtagRFPt--P2A-GFP-P38H	pUASg-attB-P38H	pDEST-aex-3p-P38H	pCAG-rP38H-3xHA
Q171E	pENTR-Q171E	pEGH1-Q171E	pCAG-mtagRFPt--P2A-GFP-Q171E	pUASg-attB-Q171E	pDEST-aex-3p-Q171E	pCAG-rQ171E-3xHA
Q298E	pENTR-Q298E	pEGH1-Q298E	pCAG-mtagRFPt--P2A-GFP-Q298E	pUASg-attB-Q298E		
Q396R	pENTR-Q396R	pEGH1-Q396R	pCAG-mtagRFPt--P2A-GFP-Q396R	pUASg-attB-Q396R		
R130L	pENTR-R130L	pEGH1-R130L	pCAG-mtagRFPt--P2A-GFP-R130L	pUASg-attB-R130L	pDEST-aex-3p-R130L	pCAG-rR130L-3xHA
R130P	pENTR-R130P	pEGH1-R130P	pCAG-mtagRFPt--P2A-GFP-R130P	pUASg-attB-R130P		
R130Q	pENTR-R130Q	pEGH1-R130Q	pCAG-mtagRFPt--P2A-GFP-R130Q	pUASg-attB-R130Q		
R130X	pENTR-R130X	pEGH1-R130X	pCAG-mtagRFPt--P2A-GFP-R130X	pUASg-attB-R130X		
R14G	pENTR-R14G	pEGH1-R14G	pCAG-mtagRFPt--P2A-GFP-R14G	pUASg-attB-R14G		
R15S	pENTR-R15S	pEGH1-R15S	pCAG-mtagRFPt--P2A-GFP-R15S	pUASg-attB-R15S		
R173H	pENTR-R173H	pEGH1-R173H	pCAG-mtagRFPt--P2A-GFP-R173H	pUASg-attB-R173H		
R173P	pENTR-R173P	pEGH1-R173P	pCAG-mtagRFPt--P2A-GFP-R173P	pUASg-attB-R173P		
R335X	pENTR-R335X	pEGH1-R335X	pCAG-mtagRFPt--P2A-GFP-R335X	pUASg-attB-R335X		
R47K	pENTR-R47K	pEGH1-R47K	pCAG-mtagRFPt--P2A-GFP-R47K	pUASg-attB-R47K		
R47W	pENTR-R47W	pEGH1-R47W	pCAG-mtagRFPt--P2A-GFP-R47W	pUASg-attB-R47W		
S229T	pENTR-S229T	pEGH1-S229T	pCAG-mtagRFPt--P2A-GFP-S229T	pUASg-attB-S229T		
T131I	pENTR-T131I	pEGH1-T131I	pCAG-mtagRFPt--P2A-GFP-T131I	pUASg-attB-T131I	pDEST-aex-3p-T131I	pCAG-rT131I-3xHA
T167N	pENTR-T167N	pEGH1-T167N	pCAG-mtagRFPt--P2A-GFP-T167N	pUASg-attB-T167N	pDEST-aex-3p-T167N	pCAG-rT167N-3xHA

PTEN Plasmids Created	Entry Vector	Yeast Vector	HEK293 vector	*Drosophila* Vector	C. elegans Vector	Rat Vector
T202I	pENTR-T202I	pEGH1-T202I	pCAG-mtagRFPt--P2A-GFP-T202I	pUASg-attB-T202I		
T348S	pENTR-T348S	pEGH1-T348S	pCAG-mtagRFPt--P2A-GFP-T348S	pUASg-attB-T348S		
T363N	pENTR-T363N	pEGH1-T363N	pCAG-mtagRFPt--P2A-GFP-T363N	pUASg-attB-T363N		
T78A	pENTR-T78A	pEGH1-T78A	pCAG-mtagRFPt--P2A-GFP-T78A	pUASg-attB-T78A		
V133I	pENTR-V133I	pEGH1-V133I	pCAG-mtagRFPt--P2A-GFP-V133I	pUASg-attB-V133I		
V255A	pENTR-V255A	pEGH1-V255A	pCAG-mtagRFPt--P2A-GFP-V255A	pUASg-attB-V255A		
W274L	pENTR-W274L	pEGH1-W274L	pCAG-mtagRFPt--P2A-GFP-W274L	pUASg-attB-W274L		
Y138L	pENTR-Y138L	pEGH1-Y138L	pCAG-mtagRFPt--P2A-GFP-Y138L	pUASg-attB-Y138L		
Y155C	pENTR-Y155C	pEGH1-Y155C	pCAG-mtagRFPt--P2A-GFP-Y155C	pUASg-attB-Y155C		
Y155H	pENTR-Y155H	pEGH1-Y155H	pCAG-mtagRFPt--P2A-GFP-Y155H	pUASg-attB-Y155H		
Y155N	pENTR-Y155N	pEGH1-Y155N	pCAG-mtagRFPt--P2A-GFP-Y155N	pUASg-attB-Y155N		
Y176C	pENTR-Y176C	pEGH1-Y176C	pCAG-mtagRFPt--P2A-GFP-Y176C	pUASg-attB-Y176C	pDEST-aex-3p-Y176C	pCAG-rY176C-3xHA
Y180H	pENTR-Y180H	pEGH1-Y180H	pCAG-mtagRFPt--P2A-GFP-Y180H	pUASg-attB-Y180H		
Y27C	pENTR-Y27C	pEGH1-Y27C	pCAG-mtagRFPt--P2A-GFP-Y27C	pUASg-attB-Y27C		
Y346F	pENTR-Y346F	pEGH1-Y346F	pCAG-mtagRFPt--P2A-GFP-Y346F	pUASg-attB-Y346F		
Y65C	pENTR-Y65C	pEGH1-Y65C	pCAG-mtagRFPt--P2A-GFP-Y65C	pUASg-attB-Y65C		
Y68H	pENTR-Y68H	pEGH1-Y68H	pCAG-mtagRFPt--P2A-GFP-Y68H	pUASg-attB-Y68H		
Y68N	pENTR-Y68N	pEGH1-Y68N	pCAG-mtagRFPt--P2A-GFP-Y68N	pUASg-attB-Y68N		

Table A.1: Plasmids acquired and created

A full list of all plasmids acquired and created for all model systems including yeast, HEK293, fly, worm, and rat.