

THE SHARED SIGNALING PATHWAYS OF AUTISM-RISK GENES AND THEIR
DISRUPTION BY GENETIC VARIANTS

INVESTIGATING THE CONVERGENT DISEASE-RELEVANT MECHANISMS IN
AUTISM SPECTRUM DISORDER

By

NADEEM MURTAZA, B.Sc.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the
Requirements for the Degree Doctor of Philosophy

McMaster University

© Copyright Nadeem Murtaza, July 2022

PhD Thesis – Nadeem Murtaza; McMaster University – Biochemistry and Biomedical Sciences

DOCTOR OF PHILOSOPHY (2022)

McMASTER UNIVERSITY

Biochemistry and Biomedical Sciences

Hamilton, ON

TITLE: Investigating the convergent disease-related mechanisms in autism spectrum disorder

AUTHOR: Nadeem Murtaza, B.Sc.

SUPERVISOR: Dr. Karun Singh

NUMBER OF PAGES: xxi, 298

LAY ABSTRACT

Autism spectrum disorder (ASD) is a group of brain disorders that affect more than 1% of children. Genetic variants are thought to cause ASD pathology, however there are currently hundreds of genes that have not been studied. We studied how disruption of one of those genes, *TAOK2*, alters brain development in mice and identified *TAOK2* variants in multiple children with ASD. We then used BioID to find the shared disease-related mechanisms between multiple ASD-risk genes, and found that mitochondrial function and activity were connected to many of these genes. We showed that BioID can be used to study the effect of mutations in multiple ASD-risk genes simultaneously. Last, we could group children with ASD with similar behavior test scores based on the shared mechanisms of ASD-risk genes. Together our findings could be used to advance the development of robust treatments or new diagnostic tools for ASD.

ABSTRACT

Autism spectrum disorder (ASD) encompasses a broad range of neurodevelopmental disorders, with two core symptoms: deficits in social communication, and restrictive interests and repetitive behaviors. Genetics is thought to play a large role in ASD and currently there are hundreds of associated genes. We first studied the thousand and one amino acid kinase gene (*TAOK2*), which plays an important role in neurodevelopment. We found that loss of *TAOK2* causes deficits in neuron development and activity, leading to morphological changes in various mouse brain regions and ASD-related behaviors. We studied the impact of *de novo* mutations identified in *TAOK2*, which caused aberrant neuron dendritic arborization and formation of synapses. To elucidate how *TAOK2* regulates neuron development we used a proximity-labeling proteomics technique (BioID) to identify its protein-protein interaction (PPI) network. We applied this same methodology to a total of 41 ASD-risk genes and observed multiple convergent biological processes, including the less-studied mitochondrial and metabolic pathways. ASD-risk genes, including *TAOK2*, associated with mitochondrial proteins were found to have altered cellular respiration. The shared ASD-risk gene PPI network enriched for other ASD-risk genes and was used to group genes based on their shared PPI networks. These gene groups showed correlation between the clinical behavior scores of individuals that had mutations within the distinct gene groups. Lastly, we identified changes in the PPI networks of multiple ASD-risk genes through BioID, which we validated with various functional assays. In summary, we developed a proximity-labeling proteomics method that identified multiple convergent

biological pathways associated with ASD. Studying the function of TAOK2 revealed multiple disease-relevant pathologies associated with the disorder, however proximity labeling has the potential to categorize multiple ASD-risk genes and elucidate their shared signaling pathways, which together, can advance the development of robust treatments for ASD.

ACKNOWLEDGMENTS

First, I would like to thank my supervisor, Dr. Karun Singh, who has given me more guidance than I ever expected since I began my graduate studies. He gave me a lot of room to grow into the researcher and person I am today, and I have learned a lot since the day I started in the lab. I will carry forward the lessons he taught me for the rest of my life. No words can express my gratitude all of his help and the trust that he put in me over the years. My passion for academia has only grown while working under his supervision.

I would like to thank my supervisory committee members, Drs. Brad Doble, Angela Scott, and Laurie Doering, who were always encouraging and gave me insights that were very helpful during my PhD. And a special thanks to Brad, whose eye to detail always kept me trying harder. Also, I want to thank Dr. Yu Lu, whose help and guidance in proteomics will never be forgotten.

I also want to thank all the members of the lab, including those who have come and gone over the years. You were all great to be around and always willing to help. You made my time in the lab enjoyable and fun, which I will miss greatly. A special thanks to Vickie Kwan for always being the voice of reason I needed, and Annie (Anran) Cheng, my certified food taster who was always willing to offer a helping hand.

I would like to express my sincere thanks to my family for their love and support during my PhD. My parents, Ashiya and Aziz, who always supported me regardless of what I wanted to do. My younger siblings, Shafeel and Shaliya, who have grown up faster than I expected and for whom I try to be a better role model. My aunt and uncle, Shareeza and Shafraz, and their kids, Mikhail and Liyana, who I can always rely on and

that make me feel right at home whenever I go back. And Amma, who always took care of me when I was young and is one of the reasons I am where I am today.

I would also like to express my extreme gratitude to the many people that made my PhD life exciting and enjoyable after moving away from my home in BC (in no particular order): Victor Gordon and Alexandria Afonso for being the friends and support I needed throughout my PhD and who were always right there with me. The Ewen House: Diogo Poças, Yusuke Shimabukuro, Peter Gysbers, Wesley Ford, Sam Peter, Tyler Meadows, Craig Hutton, Nick Jentsch, and sometimes Amay Cheam. Our crazy shenanigans were what kept me going through grad school, and without Sam and Nick I do not know how I would have survived 2020. The Cottage Crew: Janine Strehmel, Nejib Chekir, Diane Depardon, and Anthony D’Agostino, the only people with whom I could just sit back and relax. Mathieu Seyfrid for being the friend that is always up for an adventure and showed me the importance of a work/life balance, as we say “un jour, chaque jour”. Kaushal Baid and Sam Workenhe for somehow always being able to make me passionate about research on the worst days. Priscilla Hirst, Anita Phanthikane, and Arash Samiei for always being there to listen even when we are so far apart. Finally, all the members of MOPS Monday, especially Falene Chai and Armin Nasser-Ranjbar, who never let me feel like I was actually 3000 kilometers away.

Lastly, I dedicate this thesis to the late Firoz Dean, who showed me what really matters in life.

TABLE OF CONTENTS

LAY ABSTRACT	iii
ABSTRACT	iv
ACKNOWLEDGMENTS	vi
TABLE OF CONTENTS	viii
LIST OF FIGURES	xiii
LIST OF ABBREVIATIONS	xviii
CHAPTER 1: INTRODUCTION	1
1.1 AUTISM SPECTRUM DISORDER	1
1.1.1 The symptomology of ASD	2
1.1.2 Current Treatments for ASD	3
1.1.3 The etiology of ASD	4
1.1.4 Environmental stressors that increase the risk of ASD	5
1.1.5 Genetic causes and risks of ASD	6
1.1.6 Impact of de novo mutations on disease-relevant pathways.....	10
1.2 Pathophysiology of ASD	12
1.2.1 Altered brain architecture and neural circuits.....	12
1.2.2 Excitatory/Inhibitory (E/I) imbalance hypothesis of ASD.....	15

2.2 Supplemental Data.....	74
CHAPTER 3: NEURON-SPECIFIC PROTEIN NETWORK MAPPING OF AUTISM RISK GENES IDENTIFIES MITOCHONDRIAL RESPIRATION AND METABOLISM AS A CONVERGENT BIOLOGICAL PATHWAY.....	
132	
PREFACE.....	133
3.1 ABSTRACT.....	134
3.2 INTRODUCTION.....	135
3.3 METHODS AND MATERIALS.....	139
3.4 RESULTS.....	164
3.4.1 Development of a neuronal proximity-based proteomic system to identify PPI networks.....	164
3.4.2 Identification of a shared PPI network map and common pathways of 41 ASD-risk genes in mouse cortical neurons.....	169
3.4.3 The shared PPI network map identifies the tricarboxylic acid (TCA) cycle and pyruvate metabolism as a common signaling pathway in ASD.....	172
3.5 Discussion.....	177
3.6 CONCLUSION.....	182
3.7 SUPPLEMENTAL DATA.....	187

CHAPTER 4: PROXIMITY-LABELING OF AUTISM RISK GENES IDENTIFIES DISEASE RELEVANT IMPACT ON PROTEIN FUNCTION AND CLINICAL PATHOLOGIES	201
PREFACE	202
4.1 ABSTRACT	204
4.2 INTRODUCTION	205
4.3 MATERIALS AND METHODS	208
4.4 RESULTS.....	211
4.4.1 PPI networks identify differences in signaling between missense variants in ASD risk genes	211
4.4.2 The 41 ASD-risk gene PPI network map enriches for additional ASD risk genes, human disease cell types, and correlates with human behavioral phenotypes from clinical datasets	216
4.5 DISCUSSION.....	219
4.6 CONCLUSION.....	222
4.7 SUPPLEMENTAL DATA.....	227
CHAPTER 5: CONCLUSIONS AND FUTURE DIRECTIONS	233
5.1 Elucidation of TAOK2 function in neurodevelopment and using it to study ASD pathology	233

5.2 Identification of shared ASD-risk gene PPI network and convergent pathways associated with ASD using proximity-labeling proteomics.....	236
5.3 Proximity-labeling proteomics as a screening method to study the functional impact of genetic variants associated with ASD	238
5.4 Future studies using proximity-label proteomics to elucidate disease-relevant mechanisms	240
5.5 Using 3D human stem cell derived organoids to study ASD pathophysiology	242
5.6 Significance	243
REFERENCES	245

LIST OF FIGURES

CHAPTER 1

- Figure 1: Genetic variants associated with ASD.
- Figure 2: Different levels of disruption in ASD
- Figure 3: Schematic of how BirA* biotinylates interacting and proximal proteins using BioID.

CHAPTER 2

- Figure 1: *Taok2* Ko mice have altered brain morphology and brain activity
- Figure 2: *Taok2* KO mice show alterations in neurodevelopmental disorder-related mouse behavior
- Figure 3: *Taok2* KO mice have reduced dendrite growth and synaptic connectivity in the prefrontal cortex
- Figure 4: *De novo* mutations in TAOK2 impair phosphorylation at Ser181, localization in cortical neurons, and dendritic spine motility
- Figure 5: *De novo* mutations in TAOK2 impair dendrite growth and synaptic connectivity in the mouse cortex
- Figure 6: TAOK2 regulates spine function through RhoA signaling
- Table 1: Diagnostic score summary of individuals diagnosed with ASD that have *de novo* or truncating variants in TAOK2
- Figure S1: *Taok2* KO mice have altered brain morphology
- Figure S2: Abnormal neuronal and cortical positioning after loss of *Taok2*

- Figure S3: Patterns of oscillatory activity in the prelimbic subdivision of PFC and hippocampal CA1 area of neonatal *Taok2* KO mice
- Figure S4: Hyperactivity and locomotor activity controls tests in open-field and anxiety-related behavior for *Taok2* Het and KO mice
- Figure S5: Gender control in sociability and locomotor activity controls in social, working and spatial memory and fear learning paradigms for *Taok2* Het and KO mice
- Figure S6: Learning curves during training in the water maze of WT and *Taok2* Het and KO mice
- Figure S7: *Taok2* deficient mice show enhanced preference for platform position in Morris water maze
- Figure S8: *Taok2* Het and KO mice have milder reduction in dendrite growth and synaptic connectivity in the somatosensory cortex
- Figure S9: TAOK2 Ko mice have no alterations in dendrite growth and synaptic connectivity in the hippocampus
- Figure S10: No change in mIPSC frequency and amplitude in PFC and SSC neurons in *Taok2* KO mice
- Figure S11: TAOK2 is enriched in the post-synaptic density and loss of *Taok2* decreases synaptic density and stability
- Figure S12: Location of all *de novo* and rare-inherited mutations on TAOK2 and pedigree of *de novo* and truncation mutation probands

Figure S13: *TAOK2* mRNA expression analysis of 01-0337, 07-0179, and 01-0559 families

Figure S14: The effect of A135P mutation on *TAOK2* α function and localization

Figure S15: P1022* *de novo* mutation does not affect *TAOK2* auto-phosphorylation and stability and RhoA levels in LCLs compared to family members

Figure S16: P1022* elongates basal dendrite and *TAOK2* A135P and P1022* reduce dendritic integrity

Figure S17: RhoA interacts with *Taok2* in the mouse cortex and preferentially with *TAOK2* β isoform in HEK293 cells

Figure S18: The role of human-derived *TAOK2* mutations identified in ASD probands at the synapse

Table S2: Mean prefrontal-hippocampal coherence in different frequency bands as well as properties of discontinuous patterns of oscillatory activity in PFC and HC of neonatal control and *TAOK2* KO mice

Table S3: Genetic variants of *TAOK2* identified in individuals with ASD

CHAPTER 3

Figure 1: Development of a proximity-based proteomics screen to identify neuronal PPI networks for ASD-risk genes

Figure 2: Neuron-specific PPI network map identifies convergent signaling pathways between 41 ASD-risk genes

Figure 3: Regulation of cellular respiration and mitochondrial activity by ASD-risk genes

Figure S1: BioID2 of DLG4 in mouse cortical neurons

Figure S2: Comparison of ASD-risk gene PPI networks from BioID2 in HEK293 cells and mouse cortical neurons

Figure S3: Validation of BioID2 using genes localized to specific compartments in mouse cortical neurons

Figure S4: BioID2 of MECP2 in mouse cortical neurons

Figure S5: Western blots of 41 ASD-risk gene BioID2 constructs

Figure S6: Enriched pathways in the shared ASD-risk gene PPI network map identified using BioID

Figure S7: Loss of disruption of ASD-risk gene TAOK2 causes alterations in cellular respiration and mitochondrial proteins

Figure S8: Acute knockout of ASD-risk genes in mouse cortical neurons alters mitochondrial activity and cellular respiration

CHAPTER 4

Figure 1: *De novo* missense variants alter the PPI network of uncharacterized and established ASD-risk genes that correspond to functional deficits

Figure 2: Shared ASD PPI network map correlates to human brain development and disease pathology

Figure S1: *De novo* mutations in TAOK2 caused altered synaptic transmission and neuron firing

Figure S2: Enrichment of ASD-risk genes in the shared ASD PPI network map and grouping of clinical phenotypes

LIST OF ABBREVIATIONS

2D	Two dimension
2H	Two hybrid
3D	Three dimension
AD	Alzheimer's disease
ADHD	Attention-deficit/Hyperactivity disorder
AP	Affinity purification
APEX	Ascorbic acid peroxidase
ASD	Autism spectrum disorder
AVP	Arginine vasopressin pathway
BD	Bipolar Disorder
BDNF	Brain-derived neurotrophic factor
bHLH	Basic Helix-loop-helix
BioID	Proximity-dependant biotin identification
cAMP	Cyclic adenosine monophosphate
CBT	Cognitive behavior therapy
cGMP	Cyclin guanosine monophosphate
CNV	Copy number variation
Co-IP	Co-immunoprecipitation
CRISPR	Clustered regularly interspaced short palindromic repeats
DEG	Differentially expressed genes
DEP	Differentially expressed proteins

DNA	Deoxyribonucleic acid
DSM-V	Diagnostic and statistical manual of mental disorder, 5 th edition
ER	Endoplasmic reticulum
ESC	Embryonic stem cell
ETC	Electron transfer chain
fMRI	Functional magnetic resonance imaging
FXS	Fragile X syndrome
GDNF	Glial-derived neurotrophic factor
GFP	Green fluorescent protein
GOI	Gene-of-interest
GSEA	Gene set enrichment analysis
GWAS	Genome-wide association study
Het	Heterozygous knockout
HC	Hippocampus
ID	Intellectual disability
IPC	Intermediate progenitor cell
iPSC	Induced pluripotent stem cell
JNK	c-Jun n-terminal kinase
KO	Homozygous KO
LC	Liquid chromatography
LGD	Likely gene disrupting
LoF	Loss-of-function

MAPK	Mitogen-activated protein kinase
mPTP	Mitochondrial permeability transition pore
MRI	Magnetic resonance imaging
MS	Mass spectrometry
NAD	Nicotinamide adenine dinucleotide
NDD	Neurodevelopmental disorder
NPC	Neural progenitor cell
PDD-NOS	Pervasive developmental disorder – Not otherwise specified
PDE	Phosphodiesterase
PFC	Prefrontal cortex
POI	Protein-of-interest
PRS	Polygenic risk score
PSD	Post-synaptic density
PTM	Post-translation modification
PUR	Purmorphamine
RGC	Radial glial cell
RNA	Ribonucleic acid
SAG	Smoothed agonist
SSC	Somatosensory cortex
scRNA-Seq	Single cell RNA sequencing
SCZ	Schizophrenia
SHH	Sonic hedgehog

SNP	Single nucleotide polymorphism
SNV	Single nucleotide variant
TCA	Tricarboxylic acid
WES	Whole exome sequencing
WGS	Whole genome sequencing
Y2H	Yeast-two hybrid

CHAPTER 1: INTRODUCTION

1.1 AUTISM SPECTRUM DISORDER

Autism spectrum disorder (ASD) is a diagnosed condition that encompasses a broad range of neurodevelopmental disorders. ASD ranges in severity for its associated symptoms between individuals and is diagnosed as early as 18 months using comprehensive diagnostic tools, including parent interviews and scheduled observational scoring (Lord et al., 2018). The 2018 report from the National Autism Spectrum Disorder Surveillance System of Canada found that 1 in 66 children between the ages of 5-17 were diagnosed with ASD, while in the United States of America diagnosis of ASD in the same year was 1 in 44 children (Of et al., 2018; Maenner et al., 2021). However, global estimates are closer to 1% of individuals diagnosed with ASD (Baxter et al., 2015). The diagnosis of ASD is complicated and complex due to its co-occurrence with other mental and neurological disorders, which coincides with higher prevalence in mental health inpatients compared to the general population (Tromans et al., 2018). Prior to the release of the 5th edition of the Diagnostic and Statistical Manual of mental disorders (DSM-V), ASD had three distinct categories: Autistic disorder, Asperger's syndrome, and pervasive disorder – not otherwise specified (PDD-NOS) (Miles, 2011). The distinction between the three categories based on behavior symptoms were not reliable and had many overlapping features, therefore in the DSM-V, the categories were combined under the umbrella of ASD (Brentani et al., 2013).

Similar to other neurological disorders, ASD is more common in males (up to 4x higher) than in females, however this can vary from a 1:1 to 5:1 male to female ratio.

This ratio decreases with the presence of severe symptoms, such as intellectual disability (ID) (Baxter et al., 2015), suggesting that females may require a higher genetic load to show similar symptoms as males. Nevertheless, the high prevalence rate and complex symptomology make ASD one of the most common neurodevelopmental disorders (NDDs).

1.1.1 The symptomology of ASD

ASD is defined by two core symptoms: 1) deficits in social communication and reciprocal interaction and 2) the presence of restricted interests and repetitive behaviors, interests, and activities. These core symptoms are often accompanied by a large set of secondary symptoms, which have mixed co-occurrences in individuals diagnosed with ASD (Doyle and McDougle, 2012). These secondary symptoms include motor abnormalities, gastro-intestinal problems, epilepsy, intellectual disabilities, sleep disorders, language disorders, anxiety, hyper-activity, irritability, aggression, and hyper- or hypo-responsiveness to outward stimuli. To complicate matters further, at least 10 percent of ASD is co-diagnosed with other monogenic disorders, such as Fragile X syndrome, due to the number of shared symptoms between these disorders (Miles 2011). The observance of varying symptoms of ASD and co-diagnosis with other NDDs and neurological disorders further complicates the diagnosis and treatment of ASD.

1.1.2 Current Treatments for ASD

At present, the existing treatments for ASD target the associated secondary symptoms or teach coping mechanisms for anxiety-related behaviors in individuals diagnosed with the disorder (Ghanizadeh et al., 2014; Tonge et al., 2014). Existing drug treatments have been repurposed from other neurological and psychiatric disorders. Aripiprazole and Risperidone were used as atypical antipsychotics to treat schizophrenia and bipolar disorder, and are now used to treat agitation and irritability (Ghanizadeh et al., 2014). Other drugs target attention-deficit/hyperactivity disorder (ADHD) symptoms, which have high co-occurrence in individuals with ASD, including Methylphenidate, Atomoxetine, and Guanfacine. These compounds primarily target the function of dopamine receptors to modulate the release and uptake of dopamine in the brain. However, the exact mechanisms of action of these compounds have not been elucidated (Harfterkamp et al., 2012; Okazaki et al., 2019; Ventura et al., 2020). As mentioned, non-pharmacological treatments exist and focus on the behavioral aspects of the disorder. Behavioral therapies, including cognitive behavior therapy (CBT), applied behavioral analysis and pivotal response training, aim to teach children with ASD coping mechanisms against stressful situations, especially social interactions (Tonge et al., 2014). Although there are many successful cases, this is limited to a smaller portion and subset of all ASD cases. Waiting lists for these therapies can be up to 2 years, because of the sheer number of children diagnosed with ASD and the limited number of medical professionals to administer the therapy (Vismara and Rogers, 2010). Without proper treatment these children will grow up with poor prognoses due to deficits in social

interaction and their reliance on support from others, resulting in poor outcomes in education, careers and meaningful relationships (Howlin et al., 2004; Billstedt et al., 2005).

More recent advances in the treatment of ASD include a repurposed Vasopressin peptide and a novel drug labelled STP1. Vasopressin is an FDA-approved inhaled peptide that activates the arginine vasopressin (AVP) signaling pathway to increase social behaviors (Hendaus et al., 2019). Furthermore, a drug named STP1, which combines a phosphodiesterase (PDE) and NKCC1 transporter inhibitor, is currently in phase I clinical trials for the treatment of ASD. PDE regulates levels of cyclic-adenosine monophosphate (cAMP) and cyclic-guanine monophosphate (cGMP), which are consistently dysregulated in ASD mouse models (Delhaye and Bardoni, 2021). The NKCC1 transporter maintains internal chlorine ion concentration in neurons, where a high concentration of internal chloride ions prevents proper inhibition of neuron firing by inhibitory neurons and can result in seizures, a common symptom associated with ASD (Savardi et al., 2021). In summary, there are multiple types of treatments for ASD, however none target the core symptoms of the disorder. The treatments instead focus on symptoms that already vary between ASD individuals, due to a lack of underlying knowledge in the core pathophysiology and etiology of the disorder.

1.1.3 The etiology of ASD

The first theories regarding the etiology of ASD, focused on deficits in higher cognitive processing, and more specifically the individual's ability to understand the

mental states of those around them (i.e., the theory of mind) (Lai et al., 2014; Velikonja et al., 2019). Unfortunately, these theories were broad and could be applied to most neurodevelopmental disorders. To overcome the heterogeneity of ASD and the presence of symptoms shared with other neurological disorders, a more specific hypothesis for the pathophysiology of ASD was required. Current theories focus on the role of environmental stressors and genetic factors that have large contributions towards the risk of developing ASD.

1.1.4 Environmental stressors that increase the risk of ASD

Environmental and genetic causes have both been widely studied in regard to ASD pathophysiology, as well as epigenetic effects that connect environmental stresses and stimuli to genetic predispositions. Multiple environmental factors have been investigated in relation to their risk towards ASD, including higher parental age, birth trauma, maternal obesity, short intervals between pregnancies, gestational diabetes mellitus, valproate use during pregnancy, hypoxia-inducing events, and issues during pregnancy and birth (Modabbernia et al., 2017). Current hypotheses suggest that environmental factors, such as inflammation, oxidative stress, and hypoxic and ischemic damage, act as stressors, which have different effects depending on the genetic predisposition and the spatial and temporal developmental time points.

1.1.5 Genetic causes and risks of ASD

Environmental factors and stressors cannot explain a large portion of ASD cases and are affected by the genetic background of the individual. Genetics has a more significant contribution towards ASD risk, with heritability estimates from monozygotic twin studies ranging from 40 to 90%. (Gaugler et al., 2014; Colvert et al., 2015). Siblings of individuals with ASD have a 50x higher chance of also being diagnosed with the disorder, and the risk is proportional to the percentage of DNA shared with the individual diagnosed with ASD (Folstein and Rutter, 1977; Constantino et al., 2010; Risch et al., 2014; Sandin et al., 2014).

Initially, karyotyping experiments identified multiple genome regions altered in individuals with ASD (1p, 3q, 7q, 16p, and 15q) (Risch et al., 2014). Microarrays then identified smaller, more specified DNA aberrations called copy number variations (CNVs), which encompassed specific regions of the genome (Figure 1). However, these CNVs had varying penetrance, which may be explained by the presence of other ASD-associated genetic variants. For example, secondary variants have been identified in individuals with the 1q21.1, 7q11.24 and 16p11.2 CNVs and correlated to their clinical severity (Mefford et al., 2008; Girirajan et al., 2010; Qaiser et al., 2021).

The 16p11.2 deletion is the most common CNV associated with ASD; it is identified in 1% of ASD cases (Lauren A. Weiss et al., 2008; Lin et al., 2015). Fly, mouse and human induced pluripotent stem cell (iPSC)-derived neuron models of the 16p11.2 deletion have decreased brain sizes, decreased dendritic arborization and altered synaptic transmission (Blizinsky et al., 2016; Deshpande et al., 2017; Iyer et al., 2018;

Pucilowska et al., 2018). Behaviorally, 16p11.2 deletion mice present ASD-related behavior phenotypes, including deficits in vocalizations, increased anxiety-like behavior and issues with their memory (Pucilowska et al., 2018). 16p11.2, along with 15q11.2 and 15q11.13, are a few CNV regions that have shown dosage effects between heterozygous knockout (Het) and homozygous knockout (KO) models, and tend to have lower rates of *de novo* loss-of-function (LoF) mutations. Although CNVs, like the 16p11.2 deletion, are commonly identified in individuals with ASD, it is difficult to pinpoint the exact genetic cause of the disease-relevant pathology. This has created the need to identify “driver genes” or more accurately genes primarily involved in disease-related pathology. For example, *UBE3A*, *SHANK3* and *RAI1* have been identified as driver genes for the 15q11.13 deletion (Angelman’s syndrome), 22q13 deletion (Phelan-McDermid syndrome) (Buiting et al., 2016; Lutz et al., 2020). In contrast, 16p11.2 CNV has multiple genes associated with deficits in brain development and connectivity, including *TAOK2*, *KCTD13*, *MAPK3*, and *MVP*. Combinatorial knockout of 16p11.2 genes in flies revealed complex phenotypes that were not simple increases in the severity of the developmental phenotypes (Iyer et al., 2018). The identification of CNVs associated with ASD, provided powerful models for the investigation of disease-related pathways, however the large number of genes and complex interactions of the genes within these regions have made it difficult to tease apart the specific contribution of each gene to disease phenotypes.

As previously mentioned, ASD is commonly diagnosed with other neurodevelopmental and neurological disorders, which is referred to as syndromic ASD. Many of these disorders are “monogenic”, referring to the fact that one gene is directly

associated with the disorder. Some examples of co-morbid monogenic disorders include, Fragile X syndrome (FXS) (caused by silencing of the *FMRI* gene through increased presence of methylated CGG repeats), Rett's syndrome (caused by mutations in the X-chromosome linked *MECP2* gene), Angelman's syndrome, and Dravet's syndrome (caused by mutations in the *SCN1A* gene) (Comery et al., 1997; Kishino et al., 1997; Amir et al., 1999; Marini et al., 2011). Investigations of the function of these genes during development uncovered multiple pathways associated to NDDs, however the specific relevance to ASD is unclear.

The largest contributor for the genetic risk of ASD are common genetic variants, which by definition are present in greater than 1% of the entire human population. Genome wide association studies (GWAS) were the first to identify these common genetic variants, which primarily consist of single nucleotide variants/polymorphisms (SNVs or SNPs) (Figure 1). Individually these common variants carry very small genetic risk for ASD, but additively they can hold approximately 50% of the total genetic liability for ASD (Gaugler et al., 2014). GWAS were intended to identify SNVs that were enriched in the ASD population, however only 5 risk loci have been found in recent years (Grove et al., 2019). The difficulty in identifying ASD-associated common genetic variants is due to the low statistical power of GWAS, which require a large population to identify high confidence SNPs (Wang et al., 2009; Gaugler et al., 2014). SNP common heritability is estimated between 17-52%, and polygenic risk scores (PRS) are used to calculate the additive risk effects of all the SNPs on a specific disease-associated trait.

The identification of SNPs highlights the range of genetic risk associated with ASD, from polygenic to monogenic risk, which are encapsulated in the PRS. Polygenic effects are caused by an accumulation of common variants (SNVs or CNVs), which subtly alter the function of multiple genes within the same or different pathways. Individuals with ASD that carry *de novo* mutations (present only in the affected child) or rare CNVs, have additional common variants, which can be correlated to the severity of their symptoms. Overall, highly penetrant rare variants have lower PRS scores, suggesting that individuals range from hundreds of variants with small effect sizes to one mutation with a very large effect size (i.e., monogenic disorders or *de novo* mutations) (Bergen et al., 2019). Rare-inherited variants fall in the middle of this scale and primarily encompass large recurrent CNVs, deletions in cis-regulatory gene elements, and recessive gene variants (Figure 1). However, recessive gene variants are rare and have largely only been identified in consanguineous families.

Finally, with the advent of cost-effective whole exome sequencing (WES) and genome sequencing (WGS), many studies focused on the identification of rare-inherited variants (present in <1% of the population) and *de novo* mutations (O’Roak et al., 2012a, 2012b; Sanders et al., 2012, 2015; Jiang et al., 2013; Iossifov et al., 2014; Kosmicki et al., 2017; Stessman et al., 2017; Yuen et al., 2017) (Figure 1). There is an increased prevalence of *de novo* mutations in individuals with ASD and they possess higher numbers of *de novo* CNVs and LoF mutations (Neale et al., 2012; Sanders et al., 2012; Ronemus et al., 2014; Koire et al., 2021). Genes with an increased rate of *de novo* mutations are consistently evolutionarily constrained and have lower rates of mutations,

suggesting that disruption of these genes is prone to natural selection (Darnell et al., 2011; Samocha et al., 2014). Rare variants and *de novo* mutations encompass up to 3-7% of risk towards ASD, however each individual mutation has a larger impact in the pathogenesis of the disorder and present an opportunity to identify gene networks also affected by common variants (Gaugler et al., 2014; Yoon et al., 2021). ASD-associated genes have at least two or more LoF or missense *de novo* mutations identified in different individuals and a higher rate than expected rate of *de novo* mutations (Koire et al., 2021). The identification of at least two *de novo* mutations in a gene significantly increases the possible causal role of the gene in a neurological disorders, and has been seen in multiple high-risk ASD associated genes, including *CHD8*, *SCN2A*, *KATNAL2* or *DYRK1A* (Samocha et al., 2014; Satterstrom et al., 2020). Furthermore, correlation between LoF *de novo* mutations and the IQ of individuals with ASD, demonstrate the impact that *de novo* mutations have on ASD pathology (Iossifov et al., 2014; Robinson et al., 2014; Koire et al., 2021).

1.1.6 Impact of de novo mutations on disease-relevant pathways

ASD-risk genes with *de novo* mutations fall into specific cellular networks, including chromatin modification and regulation, neuronal signaling and cytoskeleton organization, ion channel activity, and post-synaptic density (Pinto et al., 2010; Voineagu et al., 2011; O’Roak et al., 2012b; Iossifov et al., 2014; Chang et al., 2015; Velmeshev et al., 2019; Ramaswami et al., 2020). Numerous genome sequencing studies have identified *de novo* and rare-inherited variants in genes that broadly fall within these

pathways. At a broader level, there is separation between ASD-risk genes that are involved in gene regulation compared to those involved in neuronal communication (Satterstrom et al., 2020). The impact and enrichment of *de novo* mutations, highlights the large contribution they have in the pathology of ASD and the benefit of studying these variants to identify disrupted biological pathways.

The majority of *de novo* missense mutations primarily affect genes at the core of biological processes, haplo-insufficient genes and evolutionarily constrained genes (Chen et al., 2018). Functional screens of *de novo* variants in less complex model organisms, such as *Drosophila melanogaster*, *Saccharomyces cerevisiae*, or *Caenorhabditis elegans* have identified functional deficits in their affected genes. One study identified the functional consequences of 74 human genes and 79 patient genetic variants, by creating transgenic *Drosophila* lines expressing the wildtype and mutant ASD-risk genes (Marcogliese et al., 2022). They investigated the effects on fly viability, behavior and morphological changes to the eye and wings. Similarly, *C. elegans* was used to study the impact of genetic variants on multiple ASD-risk genes, including *CHD8*, *CUL3*, *DLG4*, *SYNGAP1*, and *PTEN* (Wong et al., 2019; Post et al., 2020). However, translating these phenotypes and deficits into human neuron phenotypes is difficult and many times impossible. More recent studies used CRISPR/Cas9 and single-cell RNA sequencing technologies and identified subsets of ASD-risk genes that are involved in neural progenitor cell (NPC) proliferation and neurogenesis in human iPSC-derived neurons and the mouse brain (Cederquist et al., 2020; Jin et al., 2020). However, these screens do not identify the specific functions of these ASD-risk genes and are biased towards genes

involved in neurogenesis. Previous studies have also used the yeast-two-hybrid (Y2H) system to identify the interaction between ASD-risk genes and their isoforms (Sakai et al., 2011; Corominas et al., 2014). Together, these models provided important information into possible pathologies caused by patient mutations; however, the human disease relevance of these functional deficits is not easily understood. The complex behavioral phenotypes of ASD are not fully recapitulated in these models, and the function of the gene may not reflect what is seen in human neurons. Chapters 3 and 4 of my thesis focuses on using our proximity-labeling proteomics method to address the need for unbiased identification of protein interactions for ASD-risk genes and how they are impacted by genetic variants.

1.2 Pathophysiology of ASD

DNA sequencing has identified hundreds of genes that are mutated in individuals with ASD. These genes have functional roles at every aspect of brain development, and this section highlights major deficits at each stage and examples of ASD-risk genes that are involved in various processes during each stage. The neuropathology of ASD takes form at three major levels: 1) the anatomical level, 2) the brain circuit level, and 3) the cellular and 4) the molecular level (Figure 2).

1.2.1 Altered brain architecture and neural circuits

At the anatomical level, there are significant changes in the brain architecture that have been measured using magnetic resonance imaging (MRI) of live brains or imaging

of post-mortem brain slices (Anagnostou and Taylor, 2011; Lord et al., 2020). Diffusion tensor imaging, which measures the movement of water through white matter tracts in the brain, identified changes in the cerebellum, cortex, and corpus callosum of children with ASD (Courchesne et al., 2001; Emerson et al., 2017; Hazlett et al., 2017). Longitudinal imaging found differences in fiber density developmental trajectories that coincided with abnormal cortical surface growth and increased brain volume (Wolff et al., 2012; Hazlett et al., 2017). More advanced techniques known as functional MRI (fMRI), have been used to study changes in activity across the brain that may not have any large anatomical differences. These include deficits in brain circuits associated with language production and comprehension and hyperactivation of the superior temporal gyrus and inferior temporal gyrus (Herringshaw et al., 2016; Emerson et al., 2017; Uddin et al., 2017a), which together may cause difficulties in social communication through auditory and visual sensory input processing.

The cerebral cortex consists of multiple layers, which form in an inside-out pattern, where the early neurons form the deep layers and the upper layers are filled with later-stage neurons (Subramanian et al., 2020). Post-mortem brain tissue from children with ASD showed deficits in cortical layering and decreased mini-column size and density, where mini-columns are anatomical structures that contain the six cortical layers throughout the cortex (Bailey et al., 1998; Buxhoeveden et al., 2006). The formation of cortical layers during development depends on two major processes: 1) NPC proliferation and differentiation and 2) neural migration.

Cortical neurons originate from NPCs, which include radial glial cells and intermediate progenitor cells located in the ventricular and subventricular zones that form around the lateral ventricles (Garcia-Forn et al., 2020). Altered proliferation of radial glial cells (RGCs) and intermediate progenitor cells (IPCs) can lead to macroscopic morphological changes in brain structure, such as microcephaly or macrocephaly. Changes in brains size can coincide with the increase or decrease of progenitor cells, respectively, eventually resulting in fewer neurons and glial cells. Disruption of the ASD-risk genes *DDX3X*, *FOXP1*, *PTEN* and *POGZ* cause an increase in the NPC population (Amiri et al., 2012; Braccioli et al., 2017; Lennox et al., 2020; Matsumura et al., 2020). Conversely, *Chd8* knockout or mutant mouse models have depleted progenitor cell populations (Durak et al., 2016). However, these mice eventually develop megalencephaly due to over-connectivity between neurons (Suetterlin et al., 2018).

Following formation of the neocortex, NPCs differentiate into neurons and begin migrating to the various layers of the cortex in an inside-out manner. Although it is difficult to determine neuron migration deficits in ASD cases through MRI, the structural abnormalities identified in these individuals suggest early deficits in neurogenesis and migration (Pan et al., 2019). The contactin-associated protein-like 2 protein (encoded by the *CNTNAP2* gene) is important for proper neural migration and loss of the protein shows observable changes in migration of upper layer neurons (Peñagarikano et al., 2011; Gdalyahu et al., 2015). NPC proliferation and neuron migration is essential for proper cortical layering and brain architecture, which is necessary for the formation of brain circuits.

1.2.2 Excitatory/Inhibitory (E/I) imbalance hypothesis of ASD

At the brain circuit level, ASD pathology involves the disruption of information processing and cognition, and this process is carried out primarily in the cerebral cortex (Belmonte et al., 2004). Atypical neural connectivity in the brain is therefore a major component in ASD and disruptions of these circuits at critical times during development are hypothesized as a major cause for many neurodevelopmental disorders (Varghese et al., 2017). The most basic neural circuits include a balance between excitatory and inhibitory (E/I) neuron transmission, and the imbalance of these inputs is theorized as a major part in the pathophysiology of ASD (Gao and Penzes, 2015). In the cortex, this balance is primarily formed between excitatory pyramidal neurons, which are approximately 80% of the cortical neurons, and GABA-ergic inhibitory neurons that migrate from the ventral ganglionic eminences. The excitatory neurons use glutamate as the neurotransmitter to pass along excitatory signals, while inhibitory neurons use GABA to reduce the firing probability of other neurons within the circuit (Roux and Buzsáki, 2015; Tremblay et al., 2016).

Both excitatory and inhibitory neurons form axonal and dendritic projections at opposite ends. Axonal projections (or axons) grow with actin-rich protrusions, called axon growth cones. Axons grow outwards and towards different brain regions using guidance or avoidance cues throughout the brain (Zang et al., 2021). Mature axons produce and propagate action potentials towards the presynaptic bouton to transfer the signals from one neuron to the next, which is essential for proper brain circuit formation and at a higher level, cognition and information processing. An important region, known

as the axon initiation segment (AIS), is a highly organized structure at close proximity to the cell body. It is at the AIS that voltage-gated sodium and potassium channels responsible for initiating the action potential are held by multiple scaffolding proteins (Leterrier, 2018; Hamdan et al., 2020). Improper axon formation and growth or disorganization of the AIS is a common phenotype observed in multiple mouse models of ASD. For example, *SCN2A* encodes a voltage-gated sodium channel, Nav1.2, which is localized to the AIS early during development and is important for initiation of the action potential (Ben-Shalom et al., 2017; Spratt et al., 2019, 2021). Mutations identified in ASD individuals or loss of *SCN2A* causes deficits in action potential initiation and firing. Similarly, the voltage-gated sodium channel Nav1.1 (encoded by *SCN1A*) is expressed primarily in inhibitory neurons, and disruption of this gene can lead to Dravet's syndrome, a severe form of epilepsy that is often diagnosed with ASD (Marini et al., 2011; Cetica et al., 2017). The scaffold proteins that hold sodium and potassium channels at the AIS, such as Ankyrin G (encoded by *ANKK3*) or Ankyrin B (encoded by *ANKK2*) are also associated with ASD and other NDDs. Loss of either Ankyrin protein causes disorganization of the AIS and impairment of action potential initiation and firing (Hedstrom et al., 2008; Leussis et al., 2013; Yang et al., 2019; Creighton et al., 2021).

1.2.3 Deficits in synaptic formation, function and plasticity

The presynaptic bouton or terminal is a specialized compartment at the end of the axon, which forms the first half of a synapse. Once the action potential reaches the presynaptic terminal, the influx of calcium at the terminal causes the release of synaptic

vesicles into the synaptic cleft, which is the space between the pre-synaptic terminal and the post-synaptic dendritic spine. These vesicles are held in a region termed the active zone and contain different types of neurotransmitters (e.g., acetylcholine, dopamine, serotonin, gamma-aminobutyric acid (GABA), glutamate and histamine), depending on the type of neuron from which they originate. The release of these vesicles relies on numerous proteins, such as the STXBP1 protein that controls synaptic vesicle release, and is involved in regulating synaptic transmission (Patzke et al., 2015). STXBP1 has multiple mutations identified in individuals with ASD (Uddin et al., 2017b). Another ASD-risk gene, *SLC6A1*, encodes the GABA transporter type 1 protein and is required for the re-uptake and recycling of GABA from the synaptic cleft after the neurotransmitter vesicle was released (Goodspeed et al., 2020). Mutations in *SLC6A1* prevent proper re-uptake of GABA and can cause unwanted inhibition of neighbouring neurons, through activation of their GABA receptors. Proper connectivity in brain circuits relies on synapse formation and function (Geschwind and Levitt, 2007).

GABA receptors are one of many potential receptors on the receiving end of the pre-synaptic terminal, which can vary depending on the type of neuron-to-neuron connection. For example, inhibitory neurons form synapses onto the cell body or dendrites of excitatory or inhibitory neurons, while excitatory-excitatory neuron synaptic connections in the cortex occur at dendritic spines (Hering and Sheng, 2001). Contactin-associated protein-like 2 (encoded by *CNTNAP2*), neuroligins (encoded by *NRXN1*, *NRXN2* and *NRXN3*) and neuroligins (encoded by *NLGN1* and *NLGN2*) are involved in holding the presynaptic and postsynaptic terminals together. Disruption in these genes

has been associated with impairments in synaptic formation and function (Peñagarikano et al., 2011; Varghese et al., 2017).

Dendritic spines are small actin-rich protrusions that form small compartments that house multiple types of receptors and secondary messenger molecules (Hering and Sheng, 2001). Closer to the membrane and exposed to the synaptic cleft, is the specialized region called the post-synaptic density (PSD). The PSD contains multiple ASD-risk genes, including scaffolding proteins (e.g., *DLG4*, *SHANK2* and *SHANK3*), glutamate receptor proteins (e.g., *GRIA1*, *GRIA2*, *GRIN2B*) and downstream signaling proteins (e.g., *SYNGAP1* and *PTEN*). The loss or mutation of these genes cause altered synaptic transmission, which is the most accepted pathological mechanism that underlies the core symptoms of ASD at a sub-cellular level (Penzes et al., 2011; Clement et al., 2012).

Dendritic spines can take multiple forms that reflect different stages of maturity, ranging from immature thin spines to mature mushroom shaped spines. However, these forms are dynamic and can change quickly depending on the activity levels of the neuron (Hotulainen and Hoogenraad, 2010). Many neuropsychiatric disorders have found atypical numbers and morphologies of dendritic spines (Penzes et al., 2011; Broek et al., 2014). Post-mortem brains from individuals with ASD showed increased or decreased in the dendritic spine density, which has also been observed in multiple genetic models of ASD (Nimchinsky et al., 2001; Hutsler et al., 2007; Schmeisser et al., 2012). The contradictory phenotypes seen in these studies, suggested that deviance from an optimal level of connectivity in different regions of the brain can cause similar deficits overall or

that there are different types of ASDs. The disruption of synaptic transmission and plasticity is a major theory behind the pathology of ASD and multiple NDDs.

1.2.4 Mitochondria in Neurogenesis and Neuronal Synaptic function

Metabolic and mitochondrial signaling is critical for brain development because of the massive energy demand in neurons. Metabolic and mitochondrial functions play important roles in the developing brain in various cell types including NPCs, neurons and glial cells (Khacho et al., 2019; Rangaraju et al., 2019; Namba et al., 2021; Zehnder et al., 2021). The mitochondrial matrix houses multiple metabolic processes important for neurogenesis and synaptic function and plasticity. These include the tricarboxylic acid (TCA) cycle, oxidative phosphorylation via the electron transfer chain, pyruvate oxidation and beta-oxidation of fatty acids.

The metabolic functions in NPCs are essential for proper development of the neocortex. The energy in neural progenitor cells is largely supplied by glycolysis ending in lactate production, glutaminolysis and fatty acid synthesis. Due to the hypoxic conditions during early neocortical development, glucose is converted into pyruvate and converted into lactate, instead of acetyl-CoA that is used in the TCA cycle. Angiogenesis and introduction of oxygen into deeper layers of the brain helps to promote neural differentiation. The shifting of lactic acid production to TCA cycle and oxidative phosphorylation via the mitochondrial electron transfer chain (ETC) results in increase differentiation of NPCs into neurons (Mao et al., 2013).

Increased lactate production from glycolysis is concurrent with increased glutaminolysis in highly proliferative cells such as NPCs. Glutaminolysis converts glutamine into alpha-ketoglutarate, an intermediate molecule in the TCA cycle. Glutaminolysis promotes partial TCA cycle function, without the use of acetyl-CoA, which is required for proliferation and maintenance of progenitor cells (Lindhurst et al., 2006; Journiac et al., 2020; Namba et al., 2020)

Mitochondrial function is essential for the neurogenesis in the brain. Increased mitochondrial fission induces neurogenesis, while mitochondrial fusion promotes proliferation of NPCs (Iwata et al., 2020). Increased mitochondrial fission coincides with increased mitochondrial cellular respiration and oxidative phosphorylation. Oxidative phosphorylation controls the ratio of nicotinamide adenine dinucleotide (NAD) to the reduced NADH form (Stein and Imai, 2012). Rises in the ratio cause differentiation of NPCs into neurons, in part through the activation of Sirtuin-1 deacetylase (Cantó et al., 2015). Sirtuin-1 deacetylates BCL6 target genes and derives neurogenesis in NPCs (Tiberi et al., 2012; Bonnefont et al., 2019). Furthermore, CRISPR-mediated induction of neuron-enriched mitochondrial proteins can increase the conversion efficiency of glial cells to neurons during direct reprogramming, highlighting the significance of mitochondrial function in neurogenesis (Russo et al., 2021).

Mitochondria have an important role in synaptic transmission and plasticity in mature neurons at both the presynaptic terminal and the post-synaptic compartment. Neurons rely on the TCA cycle and oxidative phosphorylation via the ETC in mitochondria over lactate producing-glycolysis and glutaminolysis (Harris and Attwell,

2012). Dysfunction in mitochondrial functions regarding ATP production and calcium buffering play a significant role in the pathology of neurological disorders. Mitochondrial ATP and mitochondrial calcium are implicated in short-term synaptic plasticity (Tang and Zucker, 1997; Billups and Forsythe, 2002; Guo et al., 2005; Verstreken et al., 2005). Disruption of mitochondrial fusion in axons can lead to abnormal branching and calcium buffering. Mitochondria are trafficked to and from the presynaptic terminal and they regulate the maturation or pruning of the synapse itself (Devine and Kittler, 2018). Synaptic mitochondria are smaller, more motile, and have more cristae compared to mitochondria in the cell body. Mitochondria produce the majority of cellular ATP required for processes, such as exocytosis of the readily releasable pool of vesicles encapsulating neurotransmitters or neuropeptides. Mitochondria also have a significant role in dendrite growth and formation of dendritic spines and provides ATP for the ATP-dependent processes at the synapse, including glutamate receptor signaling and protein translation (Li et al., 2004; Cherra et al., 2013; Brot et al., 2014; Rangaraju et al., 2019).

The buffering of calcium ions in the synaptic compartment regulates synaptic plasticity during high-frequency neural stimulation. The mitochondrial calcium uniporter uptakes calcium into the mitochondria, while the mitochondrial permeability transition pore (mPTP) releases calcium into the cytoplasm. Sequestering of calcium in mitochondria prevent large fluctuations in the baseline cytoplasmic calcium levels. This allows mitochondria to buffer mean firing rates, while irregular calcium release causes aberrant synaptic transmission and attenuated LTP (Levy et al., 2003; Ruggiero et al., 2021). In addition, glial metabolic activity is also important for brain development, and

disruption of mitochondrial function in glia leads to abnormalities in astrocyte maturation and synaptic function (Zehnder et al., 2021).

A key issue in the role of mitochondria in ASD pathology is whether metabolic changes are primary pathologies, or secondary to other causes. Monogenic syndromic forms of ASD, such as TSC, CDKL5-deficiency disorder, and FXS, show that specific mitochondrial and metabolic defects and may have a genetic basis (Ebrahimi-fakhari et al., 2016; Jagtap et al., 2019; Licznanski et al., 2020). *Tsc1/2* knockout mice have increased mitochondrial size and decreased number of mitochondria at the synaptic and axonal compartments. Mouse models of Angelman syndrome displays deficits in LTP in the hippocampus, which coincides with the presence of mitochondria with disorganized cristae and reduced ETC activity (Su et al., 2011). The *BCKDK* mutant model of ASD has been treated with a diet that included high consumption of branched-chain amino acids (Novarino et al., 2012). Interestingly, ketogenic diets and low-glycemic diets in epileptic individuals has been found to decrease seizure frequency (Neal et al., 2008, 2009). The changes in mitochondrial morphology and activity, suggests that developmental alterations in metabolic/mitochondrial signaling pose vulnerability for ASD. However, very few mitochondrial genes are associated with ASD itself, such as *ETFB*, *RHEB*, and *BCKDK*, which we used in our study. Identifying a possible connection between current ASD-risk genes that do not have known metabolic functions is of interest and may help develop an understanding of how metabolic dysfunctions occur or are caused in individuals with ASD. ASD gene mutations that alter

mitochondrial and metabolic function in the developing brain have the potential to affect NPC proliferation, neurogenesis, and neuron activity.

1.3 Modelling ASD and related neurodevelopmental disorders

1.3.1 Mouse models of ASD

There are many model organisms that have been used to study ASD, however due to the focus of my thesis I will only provide background in on the use of mouse and human models to study autism and neurodevelopmental disorders (White, 2016).

Rodent models, such as mouse and rat models, have multiple advantages over current human models, including the ability to study changes in behavior and *in vivo* neural circuits. Although human models can be used to study multiple aspects of neurodevelopment, they can not yet recapitulate direct neural circuits between two brain regions or model behavior. Simultaneously, mouse models do not fully recapitulate the brain development in humans (Nestler and Hyman, 2010). For example, a major difference in the human brain is the gyrification of the cortex, which is not present in the mouse brain (Semple et al., 2013). The developmental speed of the mouse brain leads to discrepancies in the presence of cell populations that exist for longer periods of time in the human brain, thus resulting in difficulties in translating findings between the two models. Differences between mouse and human genetic models have caused confusion and uncertainty in the field regarding the impact of some ASD-risk genes (Zhao and Bhattacharyya, 2018). Mice are social animals and have specific behaviors that can be categorized as social interactions and repetitive behaviors and interests (Crawley, 2007).

A few mouse models have shown deficits in both of these core symptomatic behaviors, such as *Cntnap2*, *Shank2*, *Shank3*, and *Ube3A* mouse models (Peça et al., 2011; Peñagarikano et al., 2011; Smith et al., 2011; Schmeisser et al., 2012; Gdalyahu et al., 2015). However, multiple mouse models of ASD-risk genes have shown altered levels of neural progenitor proliferation, neural migration, synaptic connectivity, and synaptic plasticity (Chen et al., 2015). In addition, some non-genetic mouse models that utilize exposure to valproic acid or viral infections have also shown ASD-related mouse behaviors (Zerbo et al., 2013; Wagner et al., 2021). Mouse models are an essential tool to study ASD-risk genes from the molecular level to the behavioral level, which can be than in part be translated to human behavior and neuropathology through inference.

1.3.2 Differentiation of human iPSCs into neurons

Prior to the use of human stem cells, the study of human cortical neurons was limited to post mortem brain tissue. However, post mortem tissues were often from adult patients and did not reflect changes in early neurodevelopment. Furthermore, post mortem tissues have already begun the process of cell death, which can introduce many artifacts (Lewis, 2002). Human pluripotent embryonic stem cells (ESCs) or iPSCs allow the study of early neurodevelopmental time points in NPCs and neurons. The technology to reprogramme human fibroblast cells from skin biopsies or peripheral blood mononuclear cells from blood into iPSCs was first established by Takahashi et al. (2007). Before then, hESCs had been differentiated into multiple brain cell types (Thomson,

1998). The advantages of using iPSCs are that it preserves the human genetic background and can be used to study the impact of patient mutations.

Current methods can create many different types of 2D *in vitro* neurons from ESCs and iPSCs, including excitatory cortical, GABA-ergic, dopaminergic, cholinergic, serotonergic, and motor neurons (Dimos et al., 2008; Bissonnette et al., 2011; Boyer et al., 2012; Shi et al., 2012b; Liu et al., 2013; Lu et al., 2016). Glial cells have also been created, including astrocyte progenitors, and microglia (Douvaras et al., 2014, 2017; Muffat et al., 2016; Krencik et al., 2017). It should be noted that I have only cited the first published protocol for each sub-type, as novel and modified protocols for these cell types are rapidly being generated. The efficiencies of differentiation can vary due to the variability in reprogramming and genetic background in iPSCs.

There are two major methods of neural differentiation: 1) stepwise differentiation using timed extrinsic small inhibitors and growth factors and 2) direct differentiation by expressing lineage-specifying transcription factors. However, combinations of these two processes have been used to increase differentiation efficiency and cell maturation (Nehme et al., 2018). Generation of two-dimensional (2D) *in vitro* brain cells, both neuronal and glial cell types using stepwise differentiation generally have low yield, reproducibility, and require a longer time to reach specific maturation than direct differentiation methods.

In the step-wise differentiation process, neuron sub-types are generated through formation of embryoid bodies using iPSCs or ESCs. Dual-SMAD inhibition then pushes these embryoid bodies towards the neuroectoderm lineage, using SB431542 which

inhibits TGF-beta, activin and nodal signaling for mesendoderm differentiation, and BMP inhibitors (Noggin, Dorsomorphin, or LDN) to prevent trophectoderm and ectoderm differentiation (Chavali et al., 2020). The resulting embryoid bodies possess NPCs in organized formations called neural rosettes. The NPCs are then differentiated into excitatory neurons using neurotrophic growth factors (BDNF, GDNF, or NT3) and retinoic acid signaling (Vitamin A)(Chambers et al., 2009; Shi et al., 2012a). With the addition of different canonical WNT pathway modulators (CHIR99021, XAV939, dorsomorphin) dorsal neuron types are generated, while activation of the sonic hedgehog pathway (SHH recombinant proteins, purmorphamine (PUR), and smoothed agonist (SAG)) push the NPCs towards ventral neuron types (Hulme et al., 2022). Finally, use of GSK3 inhibitors to activate WNT signaling at specific time points can push embryoid bodies to produce neural crest cells and peripheral nervous system neurons. Due to the complex role of the WNT signaling and SHH signaling pathways in the brain development, step wise differentiation results in the production of mixed cell types, resulting in high variability between experiments.

To counteract the variability introduced through variable differentiation of stem cells and more importantly the neural progenitor cells, direct differentiation bypasses this step and directly differentiates iPSCs/ESCs or fibroblasts into specific neuron types. The neurogenin family (NGN1, NGN2, and NGN3) of basic-helix-loop-helix (bHLH) transcription factors are strong pro-neural transcription factors that inhibit differentiation of progenitor cells into the glial or inhibitory neuron fates (Hulme et al., 2022). Specifically, NGN2 negatively regulates PAX6 and increases expression of neuron

specific transcription factors, such as NEUROD1 and TBR1/2, while simultaneously decreasing cell cycle regulation factors inducing cell cycle exit. NGN2 also inhibits expression of the MASH1 transcription factor, which prevents differentiation into ventral cell types, such as GABA-ergic interneurons. Lentiviral infection of NGN2 expression cassettes, ectopically express NGN2 in iPSCs and directly differentiates them into iNeurons (Zhang et al., 2013). iNeurons are primarily glutamatergic/dopaminergic and express upper cortical layer markers, the presynaptic protein Synapsin1, and AMPA receptors within 3 weeks of differentiation. However, iNeurons do not express NMDA receptors until much later, highlighting some discrepancies in their development compared to step-wise differentiated neurons. However, methods have been used to counteract the generation of deep brain region neurons, including sorting of CAMKII α -expressing neurons to obtain pure glutamatergic neuron cultures or co-expression of NGN2 with EMX1, a transcription factor expressed in pyramidal cortical neurons, which results in 100% glutamatergic neurons (Nehme et al., 2018; Ang et al., 2020).

The reduced variability between different iPSC lines or differentiation batches drastically raises the benefits of using iNeurons for high-throughput screens. High-throughput screen methods using iNeurons have identified Tau reducing compounds, genes important for neuronal survival, and compounds that alter neurite outgrowth (Wang et al., 2017a). iNeurons provide a quick and robust method of studying excitatory neuron function at a higher throughput than traditional step-wise differentiation methods and has potential therapeutic uses.

1.3.3 Using iPSC-derived neurons to study NDDs and ASD

iPSCs have been used to study many neurological disorders, including Rett syndrome, FXS, Timothy syndrome, William syndrome, Prader-Willi syndrome, Angelman syndrome, Schizophrenia (SCZ), and Bipolar Disorder (BD) (Chamberlain et al., 2010; Marchetto et al., 2010; Ananiev et al., 2011; Brennand et al., 2011; Paşca et al., 2011; Sheridan et al., 2011; Cheung et al., 2011; Krey et al., 2013; Tian et al., 2014; Doers et al., 2014; Mertens et al., 2015; Nageshappa et al., 2015; Djuric et al., 2015; Khattak et al., 2015; Chailangkarn et al., 2016). Some studies have also used patient or isogenic iPSC lines to study idiopathic ASD, ASD-linked genes (*CHD8*, *SHANK3*) or CNVs (*22q13*, *16p11.2*) (Shcheglovitov et al., 2013; Wang et al., 2015, 2017b; Yi et al., 2016; Deshpande et al., 2017). Fewer studies have used NGN2-induced iNeurons due to their more recent development (Zhang 2013). However, some have been used to study neurological disorders and psychiatric disorders, including Alzheimer's disease (AD), ASD, Down syndrome, FXS, and SCZ to study deficits in excitatory neuron morphology, function and/or connectivity (Lin et al., 2018; Xu et al., 2018; Deneault et al., 2019; Meyer et al., 2019; Graef et al., 2020; Hirata et al., 2020). The NGN2 method combined with CRISPR/Cas9 have also been used to study *C9orf72* in FTD/ALS iPSCs. X-linked dystonia parkinsonism (XDP) has been rescued using CRISPR/Cas9 in iPSC derived neurons to normalize TAF-1 expression (Jovičić et al., 2015; Aneichyk et al., 2018). In addition, a study from our lab used isogenic iPSC lines with knockouts of various ASD-risk genes and identified common alterations in neural activity and connectivity (Deneault et al., 2018).

iPSC-derived neurons provide a quick and efficient method to study neurodevelopmental disorders, and combined with CRISPR/Cas gene editing, can be used to investigate the potential impact of patient mutations. In Chapter 3 and 4, I used CRISPR/Cas9 to study the impact of a *de novo* missense mutation on the function of TAOK2 in neuron function and activity.

1.4 The function of TAOK2 in brain development

1.4.1 Thousand and one amino acid kinase 2 (TAOK2)

TAOK2 is a mitogen-activated protein kinase (MAPK) and is part of the Sterile 20-like kinase family, which are generally downstream of G-coupled protein receptors and upstream of other MAPKs (Delpire, 2009). Although two other orthologues of TAOK2 exist, TAOK1 and TAOK3, with greater than 80% homology, only TAOK1 and TAOK2 are expressed in the brain. The TAOK2 protein consists of an N-terminal kinase domain, a serine-rich MEK binding domain, and either a C-terminal domain with 2-3 coiled-coils in the beta isoform or leucine-rich repeats in the alpha isoform. There are two major known isoforms of TAOK2, the longest 1235aa alpha isoform and the 1049aa beta isoform. The alpha isoform is expressed from early in development and is more associated with microtubule dynamics, while the beta isoform is expressed later in development during early synaptogenesis and has been found to regulate synapse formation (Chapter 2 Supplementary Figure 18). Both contain the N-terminal kinase domain, but differ in their C-terminal regulatory domain. Serine-181 is located in the kinase domain, and phosphorylation is essential for TAOK2 kinase activity, while, the C-

terminal domain of the beta isoform was found to act as a negative-regulator of the kinase domain (Chen et al., 1999; Zhou et al., 2004) (Chapter 2 Figure 4)

TAOK2 is involved in MAPK signaling and cytoskeleton dynamics (Hu et al., 2021). TAOK2 is an upstream activator of the stress-related MAPK signaling pathways, comprised of the p38/MAPK14 and JNK/SAPK cascades (Hutchison et al., 1998; Chen et al., 1999; Moore et al., 2000; Chen and Cobb, 2001). The stress-related MAPK signaling pathways are activated through multiple types of receptors and are essential for cell proliferation, differentiation, migration and apoptosis (Takeda and Ichijio, 2002; Nishina et al., 2004). TAOK2 targets the stress responsive MAPK pathway, interacting with both the c-Jun n-terminal kinase (JNK) and p38 MAPKs. TAOK2 also functions to regulate both the actin and microtubule cytoskeleton (Moore et al., 2000; Mitsopoulos et al., 2003), which functionally distinguishes it from TAOK1 (Hu et al., 2021). During cell stress, TAOK2 α can activate the JNK pathway leading to the cleavage of TAOK2 α from the microtubule cytoskeleton by caspase 9 and the nuclear localization of its kinase domain, resulting in apoptosis (Zihni et al., 2006). A recent study also found that TAOK2 is required for proper tethering of the endoplasmic reticulum (ER) to the microtubules (Nourbakhsh et al., 2021). The MAPK signaling and cytoskeleton dynamics regulated by TAOK2 are two core functions that translate into its control over neuron function and activity.

1.4.2 Neuron development and synaptic functions of TAOK2

TAOK2 is heavily involved in brain development and neurodevelopmental disorders. The TAOK2 gene is located in the 16p11.2 region, which is deleted in 1% of ASD patients (Lauren A. Weiss et al., 2008). TAOK2 α and β isoforms were found to differentially regulate neuron migration and synaptic transmission, respectively. Our collaborator, Dr. Calderon de Anda, originally found that TAOK2 α interacts with Neuropilin 1 and Semaphorin 3a to activate the JNK pathway, and disruption of this pathway resulted in decreased basal dendrite growth in mice (Calderon de Anda et al., 2012). At the synaptic level, TAOK2 β binds to the N-cadherin/arcadlin (rat ortholog of protocadherin) dimer and activates p38 via phosphorylation of MEK3. N-cadherin is a synaptic adhesion protein, and its endocytosis is important for dendritic spine and synapse reorganization (Arikkath and Reichardt, 2008). p38 then phosphorylates TAO2 β at Ser1038, which induces the endocytosis of the cadherins and overall number of dendritic spines (Yasuda et al., 2007). TAOK2 was also found to regulate synaptic development through myosin Va localization and PSD95 stabilization. Phosphorylation of TAOK2 via MST3 is important for proper myosin Va localization and synapse development (Ultanir et al., 2014). While, phosphorylation of Septin7 by TAOK2 stabilizes PSD95, which is required for dendritic spine maturation (Yadav et al., 2017).

In addition, TAOK2 is a category 2 gene in the SFARI database, a widely accepted database of ASD-associated genes, created through review of genomic sequencing data and individual gene studies. Category 2 is the second highest category for non-syndromic ASD and is based on rigorous testing and replication of results from

multiple independent sources. Our collaborator, Dr. Scherer (SickKids) identified multiple genetic variants through whole-exome and -genome sequencing studies of more than 5000 individuals (at that time). Of these mutations three were *de novo* variants, A135P, P1022*, and a third splice site variant (c.563+12_563+15 deletion), and two rare-inherited missense mutations, A335V and H781R. The identification of these mutations was a major part of my first aim, which was to study the impact of disrupting *TAOK2* on neurodevelopment and how it related to ASD pathology. Together, the previous studies on *TAOK2*, combined with our publication, implicated *TAOK2* as a novel ASD-risk gene and highlighted its important function in proper brain development.

1.5 Current challenges in studying ASD

1.5.1 Hypothesis-based vs discovery-based studies in ASD

Hypothesis-based studies focus on expanding on what is currently known, while discovery-based studies use broad techniques to identify novel avenues for investigation. Hypothesis-based studies focus on elucidating the function of a single gene. These studies are generally observationally driven, such that the results from previous experiments lead to the next question. This type of investigation has led to much of our understanding in many ASD-risk genes. However, these are time and resource intensive, and some labs can study a single gene for decades with novel questions being generated regularly. The large number of sequencing studies that have been conducted in the last decade have identified 100s of genes and created a bottleneck in understanding the function of all these ASD-risk genes. The time required to study each gene has restricted most studies to

focus on genes considered “low-hanging fruit”, as in genes that have strong phenotypes. For example, *FMRP*, *CHD8*, *SCN2A*, and *SYNGAP1* are genes that are highly associated with ASD and studies have found significant roles for each during development. However, the focus on these genes has created a void where less-known genes with subtle phenotypes are not strongly studied or easily funded. Furthermore, neurodevelopmental disorders share a wide variety of symptoms and risk genes, creating a complex issue in studying single genes as models of specific disorders. The overlap of these genes between disorders and molecular pathways have suggested that they fall within convergent pathways, however studying shared pathways between multiple genes is difficult and many times infeasible. This bottle neck has led to a shift in the field towards using discovery-based studies to obtain large datasets through ‘-omics’ techniques, such as RNA/DNA sequencing, mass spectrometry, and high-throughput screens and assays.

1.5.2 Transcriptomics studies in ASD

RNA sequencing identifies and quantifies the number of mRNA transcripts in a sample through sequencing of cDNA libraries prepared from the sample itself. RNA is extracted from tissue or cell pellets and purified for mRNAs with poly(A) tails, to avoid detection of ribosomal RNAs and microRNAs. The isolated mRNA is then reverse transcribed into cDNA, and short DNA adaptors and barcodes are ligated onto the ends. The barcodes allow the combining of samples so that they can be run simultaneously through the sequencing machine. The number of reads obtained per sample, depends on the sensitivity and throughput of the machine. The number of read pairs correlates with

the number of reads per sequenced region, which is important for reliable quantification. The reads are then mapped to a reference genome or transcriptome and the number of reads per gene normalized to the total reads are counted. Bioinformatic tools are then used to determine differentially expressed genes (DEGs) between separate samples. Following the identification of DEGs, pathway analysis is used to identify gene networks or gene sets that are dysregulated in one condition compared to the control. There are two main types of analysis, gene set enrichment analysis (GSEA) or gene over-representation analysis (Reimand et al., 2019). GSEA uses the fold change difference and significance to estimate shifts in the expression of a gene within a gene set and can be used to determine up or downregulation of pathways (Subramanian et al., 2005). Conversely, gene over-representation analysis uses a set of chosen DEGs to determine if certain pathways are enriched within the gene set and potentially affected in the tissue or cells (Raudvere et al., 2019).

RNA-sequencing of whole tissue or cell cultures, known as bulk RNA-Seq, has identified multiple regions of the brain have been implicated in the pathology of ASD. Combining genome and epigenome sequencing experiments with transcriptome data have found large effects in the cortex of ASD cases (Voineagu et al., 2011; He et al., 2019). More specifically, mid-fetal deep cortical projection neurons and glutamatergic projection neurons had significantly altered gene expression profiles and enrichment of ASD-risk genes compared to unaffected controls (Parikshak et al., 2013; Willsey et al., 2013; Ramaswami et al., 2020). The loss of transcriptomic differences between different brain regions was seen when looking at the transcriptome of frontal and temporal cortices

of individuals with ASD (Parikshak et al., 2013). Furthermore, bulk RNA-seq of multiple mouse models and human stem cell-derived neuron models, have identified many ASD-associated pathways that are commonly disrupted in multiple models (Quesnel-vallières et al., 2019; Griesi-Oliveira et al., 2021)

Single cell RNA-sequencing (scRNA-seq) is similar to bulk RNA-seq, however the method first separates individual cells in small oil beads and sequences each cell separately. Usually, 5000-10000 cells per sample are sequenced and an average of 3000-4000 transcripts are identified per cell, as opposed to the whole transcriptome. However, scRNA-seq studies have identified multiple gene co-expression modules or cell types that are altered in genetic models of ASD and individuals with the disorder.

ScRNA-seq of post-mortem brains from individuals with ASD and their matched controls identified DEGs in layer 2/3 and 4 cortical neurons, VIP interneurons, branched grey matter localized astrocytes, and microglia, highlighting their importance in ASD pathology (Velmeshev et al., 2019). Cortical excitatory and inhibitory neurons also had the highest percentage of DEGs in individuals with ASD and epilepsy, which highlights the correlation of clinical behavior and the role of specific cell types. Another study from the same lab, described differences in the developmental trajectory of different brain regions and cell-types using scRNA-seq. Together, these studies underscore the importance of developmental time points and spatial location in the brain in regards to the pathophysiology of ASD. Another lab combined *in vivo* CRISPR/Cas9 knockout of 35 ASD-risk genes in embryonic mice with single cell RNA-sequencing to identify common disrupted cell-types; this method was termed *in vivo* Perturb-Seq (Jin et al., 2020).

Similar to previous studies, they found that gene expression modules associated with cortical projection neurons were altered, as well as genes associated with inhibitory neurons and multiple glial cell types in the mouse brains at postnatal day 7. ASD is a neurodevelopmental disorder and therefore specific ASD-risk genes most likely have larger roles and impacts at different developmental times and areas in the brain.

1.5.3 Proteomic studies in ASD

During my PhD, I have published a review on the emerging proteomic techniques that are used to study neurodevelopmental disorders and ASD. I have left the detailed background regarding mass spectrometry outside of this thesis and focus on the use of proteomics to study neurological disorders (Murtaza et al., 2020). In brief, the core for proteomic studies is the use of the mass spectrometer to identify changes in protein abundance, post translational modification (PTMs) and protein interactions. The two major types of quantitative mass spectrometry techniques are top-down and bottom-up proteomics. Top-down proteomics is used to analyze simple protein mixtures and has been used extensively to identify biomarkers for ASD through blood (Taurines et al., 2010; Cortelazzo et al., 2013; Ni et al., 2018). Bottom-up proteomics, or sometimes referred to as ‘shotgun’ proteomics, is similar to RNA sequencing in that it identifies and maps peptides to proteins and is used to quantify the abundance of those proteins (Silva et al., 2006). To prepare samples for mass spectrometry, tissue or cells are lysed and proteins are digested, with peptidases such as Trypsin, into peptides which are then separated by size and charge by liquid chromatography (LC). The peptides are sprayed

into the mass spectrometer (MS), at which point the peptides can be run through 1 or 2 (MS or MS/MS) tandem analyzers. MS systems measure the mass to charge ratio (m/z) of each peptide, while MS/MS systems fragment select ions to be analyzed again with the benefit of increased quantification and identification accuracy (Thompson et al., 2003). However, the trade-off is that less peptides and therefore proteins are often identified using LC-MS/MS systems.

Initial studies using shotgun proteomics used post-mortem brain tissues and identified specific proteomic signatures that were altered in individuals with ASD (Abraham et al., 2019). Interestingly, this work found that the proteomic signatures between different regions of the brain were unchanged in ASD, which was the opposite of what was seen through RNA sequencing (Voineagu et al., 2011; Velmeshev et al., 2019). Although mRNA stability, degradation, and localization play a large role in these differences, it can also be impacted by translation rate, protein stability, protein degradation, protein folding, and post-translational modification. However, similar to transcriptomics studies, changes in proteins related to synaptic function, cytoskeleton organization, organelle transport, and glutamatergic transmission were observed. Work from our lab also found that the changes in the transcriptome and proteomics of *Dixdc1* KO mice did not show any correlation (Kwan et al., 2021). This highlights that changes in the transcriptome between ASD and control individuals may not be followed by changes in the proteome. Selected reaction monitoring MS (SRM-MS) of post-mortem PFC and CB tissues from individuals with ASD also found changes in the levels of myelination proteins and increases in synaptic vesicle regulation and energy metabolism

proteins (Broek et al., 2014). Few studies have used proteomics to study changes in protein levels or modifications of ASD-risk genes, as most focus has been at the transcriptome level. One example is measuring changes in the PSD composition in *Shank3* mutant mice, which also saw region specific changes in the proteome (Reim et al., 2017). Another more recent study used RNA-sequencing and mass spectrometry on organoids generated from iPSCs with the 16p11.2 deletion, and identified multiple pathways associated with ASD (Urresti et al., 2021). Transcriptomics and proteomics are two major methods that can be used to study large changes in the cell, however they cannot distinguish compensational changes that occur overtime in genetic models or identify which proteins or genes are directly associated with each other.

1.5.4 Proximity-labeling proteomics to identify protein-protein interactions (PPIs)

Protein interactions are essential for their function and disruption of these interactions can cause significant downstream effects in biological processes and signaling pathways. Traditional methods utilized techniques to identify direct or indirect physical interactors, such as 2-hybrid (2H) screens, co-immunoprecipitation (Co-IP), or affinity-purification. Newer yeast-2H screens utilize the overexpression of a bait and prey proteins fused to one half of a split reporter protein (such as splitGFP). When the two recombinant proteins interact, it allows the reconstitution of the full reporter, which can then fluoresce. This version of the screen requires the use of prey proteins that are already known and can only confirm the interaction; it cannot identify proteins that are not put in the screen. However, traditional Y2H systems be used for large scale screens through the

use of a library. This system used the GAL4 transcription factor DNA-binding domain and GAL4 activation domain, which interacts with RNA polymerase (Causier and Davies, 2002). Constructs containing the DNA-binding domain are in-frame with a POI, while libraries are created of ORFs in-frame with the activation domain. When a protein from the library interacts with the POI it allows the transcription and expression of the reporter gene. The most common reporter gene is the lacZ reporter gene, which expresses the β -galactosidase that can be used for colorimetric assays with X-gal. HIS3 or LEU2 are often used as a co-reporter, as they are necessary to generate the leucine and histidine amino acids which are exempt from the growth media. Only yeast cells that have interaction between the POI and library protein will express these genes and allow the cells to survive in the selection growth media. Cells that survive and/or are positive for LacZ expression is then isolated and the cDNA is sequenced to identify the protein from the original library. Although less-biased, these screens are primarily done in yeast or mammalian cell lines through transfection or transformation of the constructs, which is not easily accomplished in primary mouse neurons, and therefore does not always reflect the relevant biological context. Y2H has already been used to identify the protein-protein interaction networks (PPI networks) of hundreds of ASD-risk genes and their isoforms (Sakai et al., 2011; Corominas et al., 2014). However, PPI networks identified in yeast cells may not identify relevant information for many of these proteins, especially proteins in neuron specific structures, such as the synapse.

An alternative to 2H screens is Co-IP or affinity-purification (AP) coupled with mass spectrometry to identify the interacting proteins of endogenous proteins in the cell

type of interest. Co-IP requires the existence of a good antibody against the protein of interest (or insertion of a tag that can be pulled-down). It also requires larger number of cells and can be strongly affected by the washing conditions, resulting in the identification of only strong, robust interactions. Another disadvantage is that Co-IP typically requires antibodies bound to beads, which may inadvertently bind proteins and increases the noise of background proteins identified through mass spectrometry. Therefore, Co-IP is also prone to identify false positive proteins that are not directly or indirectly (as a protein complex) with the POI. Co-IP and AP has been used to identify the interacting partners of ASD-risk genes, like *DYRK1A* and *DLG4* (*PSD95*) (Fernández et al., 2009; Brown et al., 2018; Guard et al., 2019; Lautz et al., 2021).

Proximity-labeling methods coupled with mass spectrometry, can allow the identification of weak, transient protein interactors along with the strong, robust interactors, with the addition of proteins in close proximity. There are two major proximity-labeling methods, that use either engineered ascorbate peroxidase (APEX) or the promiscuous BirA* biotin ligase (BioID) that utilize biotin-phenol or biotin, respectively, to label lysine residues of nearby proteins. BirA* will biotinylate any protein in a 10nm range and reaches peak biotinylation activity at 18-24 hours after the addition of biotin (Li et al., 2017). The Roux lab originally developed the BioID2 system, a method used to identify protein-protein interactions and proximal proteins for the gene of interest in an endogenous environment (Roux et al., 2012, 2013; In et al., 2016). BioID2 performs under the same premise as its predecessor, BioID, which is a promiscuous bacterial biotin ligase (BirA*). Importantly, BioID2 is smaller and has a

higher affinity for biotin compared to BioID, making it more efficient and easier to use in our screen.

APEX and BioID have primarily been used to identify the proteins in cellular compartments (organelles and stress granules) or specific signaling pathways (GTPases and GPCRs) in mammalian cells lines such as HEK293 cells (Filamin-a et al., 2017; Kim et al., 2017; Branon et al., 2018; Chou et al., 2018; Liu et al., 2018; Markmiller et al., 2018; Regulator et al., 2018; Youn et al., 2018). Only two studies have used APEX on cultured cortical neurons, focusing on the identification of proteins in the synaptic cleft using a few genes or to identify the RNA localized to different cellular compartments (Loh et al., 2016; Chung et al., 2017). These publications required a large number of neurons (>90 million cells), possibly due to the use of APEX, which has very quick activation dynamics and therefore requires many infected neurons to counter the endogenous biotin background. Two studies accomplished BioID of PSD95 (found in the excitatory post-synapse) and Gephyrin (found in inhibitory post-synapse) in the mouse brain (Uezu et al., 2016; Spence et al., 2019). However, *in vivo* BioID requires the use of Adenovirus-associated virus (AAV) that has a smaller packaging size limitation preventing the study of larger proteins that are commonly expressed in the synapse and associated with ASD. Therefore, the possible number of proteins that can be studied *in vivo* is drastically reduced due to the AAV packaging limits. A major issue in studying ASD is that not all genes can be studied, ASD has a large contribution of rare inherited and common variants that have small effects but together can lead to ASD, and because of the polygenic nature of ASD studies that try to investigate how they are linked are

severely underpowered. In Chapter 3, we developed a mouse cortical neuron specific BioID system that can be used to identify the PPI network of multiple ASD-risk genes with minimal number of neurons and in a relevant disease cell-type.

1.6 Thesis Objectives

For this thesis I have divided my main objective to investigate the function of ASD-risk genes and their convergent disease-relevant biological processes into three aims, which coincide with Chapters 2, 3, and 4.

To accomplish the first aim in Chapter 2, we characterized multiple aspects of neurodevelopment with a *Taok2* KO model, including behavior, gross brain morphology and architecture, neural circuit activity, dendritic growth and dendritic spine formation and function. Using behavioral tests, MRI, *in vivo* and *in situ* electrophysiology, and confocal microscopy, we determined that TAOK2 acts through activation of RhoA to regulate dendritic spine motility and overall spine formation. Through collaboration with Dr. Scherer at SickKids, we identified multiple *de novo* variants that altered the kinase function and stability of TAOK2 and impaired its role in dendritic arborization and dendritic spine formation. The results of this aim were published in *Molecular Psychiatry* in 2019. The majority of the work was equally done by myself and Dr. Melanie Richter, and we shared co-first authorship.

The second aim of the project, in Chapter 3, stemmed from the results of Aim 1, which highlighted the association of TAOK2 with pathologies related to ASD. In this aim, we modified a proximity-labeling proteomics system (BioID) to identify the PPI

network of ASD-risk genes in mouse cortical neurons. Originally, this aim was focused on identifying the PPI network of TAOK2, we then widened our focus to identify the PPI network of 41 ASD-risk genes with a range of functions to elucidate convergent disease mechanisms, seen in the Taok2 KO mouse, and associated with ASD. Using BioID2 we identified PPI networks of 41 ASD-risk genes and multiple convergent pathways enriched in their shared PPI network. These included known pathways related to synaptic transmission and formation, as well as the less-studied mitochondrial activity and metabolic pathways. TAOK2 was found to be associated with mitochondria proteins, and validation with RNA-sequencing and mass spectrometry of the mouse brain found that mitochondrial proteins were reduced at both the mRNA and protein level. Metabolic seahorse assays combined with CRISPR/Cas9 gene knockout system were used to show deficits in mitochondrial cellular respiration in *TAOK2* KO neurons and other ASD-risk gene KOs (*SYNGAP1*, *ETFB*, and *RHEB*). In this aim, we also used CRISPR/Cas9 gene editing to create isogenic iPSC lines that have complete loss of TAOK2 or the heterozygous A135P *de novo* mutation. We validated the synaptic formation and transmission deficits seen in the mouse model in Aim 1 and replicated the deficits in cellular respiration identified in Aim 2, highlighting the similarities of TAOK2 disrupted pathways in mouse and human models. Overall, in aim two we identified mitochondrial function and activity as a convergent biological process between multiple ASD-risk genes, including *TAOK2*.

The final aim, found in Chapter 4, focused on using BioID to determine the impact of patient mutations on ASD-risk gene interaction networks and the ability to use

PPI networks to group individuals based on the connectivity of these genes. We studied the effect of *de novo* mutations on three ASD-risk genes, *TAOK2*, *GRIA1*, and *PPP2R5D*. *TAOK2* and *GRIA1* are known synaptic proteins, while *PPP2R5D* was found to localize to the synapse through the BioID screen in Aim 2. We found that *TAOK2* A135P and *GRIA1* R208H PPI networks had reduced enrichment of synaptic proteins, which coincided with reduced synaptic transmission seen in human iPSC-derived neurons and mouse cortical neurons, respectively. *PPP2R5D* E198K and E420K both had novel interactions with Golgi-apparatus proteins and caused atypical Golgi morphology and localization in mouse cortical neurons. The shared ASD-risk gene PPI network identified in Aim 2, had a significant enrichment of 112 other ASD-risk genes and DEGs found in individuals with ASD. To determine the relevance of this enrichment to ASD clinical phenotypes, we focused on grouping ASD-risk genes based on their shared PPI networks. We identified three groups of genes based on their connectivity and found that these groups correlated to the adaptive behavior and socialization clinical test scores in ASD patients.

The overall objective of my PhD thesis was to investigate the convergent pathways between ASD-risk genes associated with the pathophysiology of ASD. This objective was divided into three major aims: 1) Determine the function of *TAOK2* in brain development and how its disruption in ASD causes disease-related pathologies 2) Identify the shared PPI networks and convergent biological pathways of *TAOK2* and other ASD-risk genes, and 3) Develop an unbiased screening method to study the impact of patient mutations on disease-relevant PPI networks and mechanisms.

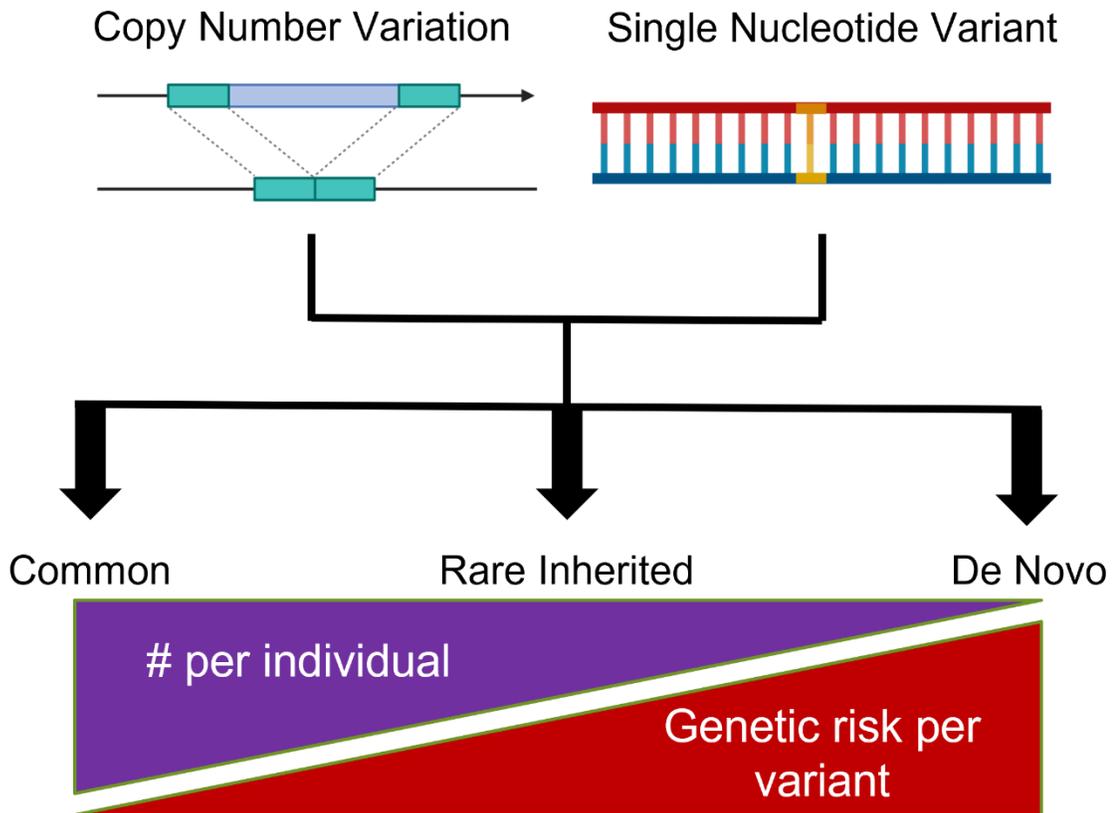


Figure 1. Genetic variants associated with ASD. Schematic shows the two major types of genetic variants identified in individuals with ASD and the types of inheritance of these variants. Common variants are found in greater than one percent of the population and individually have the least risk per variant, however cumulatively can add up to large total risk for ASD. Rare inherited variants are in less than one percent of the population and have medium risk for ASD. Common and rare inherited variants are passed down through inheritance. De novo variants are only identified in the affected individual and were due to germline mutation in the parents. These variants have the highest risk for ASD per variant, and have a very high evolutionary selection pressure.

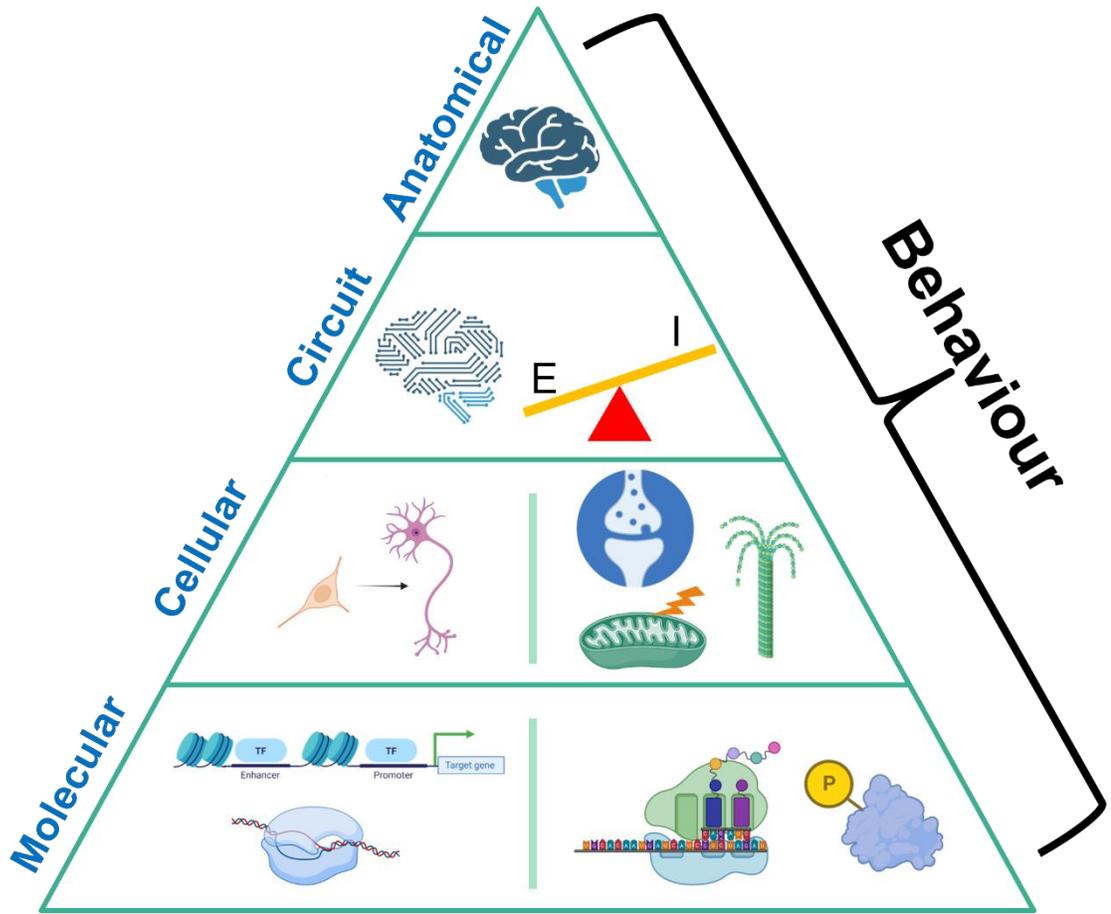


Figure 2. Different levels of disruption in ASD. Disruption of molecular mechanisms, such as gene expression, protein translation, and protein function can cause disruptions at the cellular and subcellular level, including mitochondria function, microtubule dynamics, synaptic function and neuron differentiation. These changes can lead to altered neural circuits across the brain, including atypical excitatory and inhibitory neuron activities that skew the E to I ratio. Finally, these can cause large anatomical changes in specific regions of the brain or across the whole brain. Disruption at any level can lead to the ASD-related behaviors seen in individuals with ASD or mouse models of ASD.

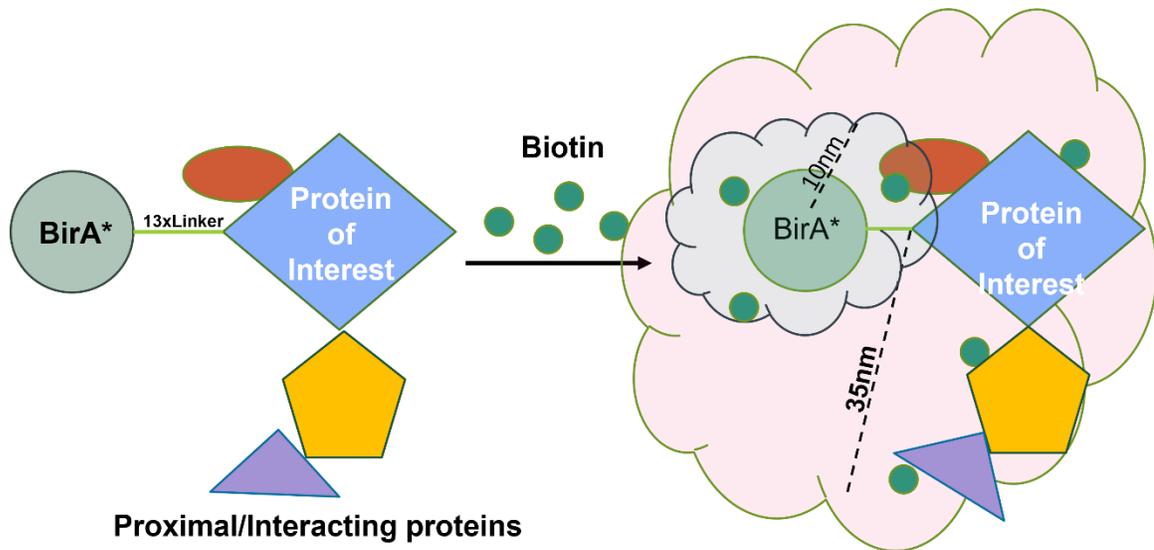


Figure 3. Schematic of how BirA* biotinylates interacting and proximal proteins using BioID.

CHAPTER 2: ALTERED TAOK2 ACTIVITY CAUSES AUTISM-RELATED NEURODEVELOPMENTAL AND COGNITIVE ABNORMALITIES THROUGH RHOA SIGNALING

Melanie Richter^{1,13}, Nadeem Murtaza^{2,3,13}, Robin Scharrenberg¹, Sean H. White^{2,3}, Ole Johanns¹, Susan Walker^{4,5}, Ryan K. Yuen^{4,5}, Birgit Schwanke¹, Bianca Bedürftig¹, Melad Henis¹, Sarah Scharf⁶, Vanessa Kraus⁶, Ronja Dörk⁶, Jakob Hellmann⁶, Zsuzsa Lindenmaier^{7,8}, Jacob Ellegood⁷, Henrike Hartung^{9,10}, Vickie Kwan^{2,3}, Jan Sedlacik¹¹, Jens Fiehler¹¹, Michaela Schweizer¹², Jason P. Lerch^{7,8}, Ileana L. Hanganu-Opatz⁹, Fabio Morellini⁶, Stephen W. Scherer^{4,5*}, Karun K. Singh^{2,3*} and Froylan Calderon de Anda^{1*}

¹Center for Molecular Neurobiology Hamburg (ZMNH), Research Group Neuronal Development, University Medical Centre Hamburg-Eppendorf, Hamburg, Germany

²Stem Cell and Cancer Research Institute, Faculty of Health Sciences

³Department of Biochemistry and Biomedical Sciences, Michael G. DeGroot School of Medicine, Faculty of Health Sciences, McMaster University, Hamilton, Ontario, Canada

⁴The Centre for Applied Genomics and Program in Genetics and Genome Biology, The Hospital for Sick Children, Peter Gilgan Centre for Research and Learning, Toronto, Ontario, Canada

⁵Department of Molecular Genetics and McLaughlin Centre, University of Toronto, Toronto, Ontario, Canada.

⁶Center for Molecular Neurobiology Hamburg (ZMNH), Research Group Behavioral Biology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany.

⁷Mouse Imaging Center, The Hospital for Sick Children

⁸Department of Medical Biophysics, University of Toronto

⁹Developmental Neurophysiology, Institute of Neuroanatomy, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

¹⁰Current address: Laboratory of Neurobiology, Department of Biosciences, University of Helsinki, Helsinki, Finland

¹¹Department of Neuroradiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

¹²Center for Molecular Neurobiology Hamburg (ZMNH), Core Facility Morphology and Electron microscopy, University Medical Centre Hamburg-Eppendorf, Hamburg, Germany

¹³Co-first authors

*Co-Senior and Co-Corresponding authors: Froylan Calderon de Anda

(froylan.calderon@zmnh.uni-hamburg.de), Karun K. Singh (singhk2@mcmaster.ca) and

Stephen W. Scherer (stephen.scherer@sickkids.ca)

PREFACE

This chapter focuses on investigating the function of the TAOK2 gene in brain development and how its disruption can cause deficits associated with ASD pathology.

This study was published in Molecular Psychiatry on February 21 2018, with myself as

co-first author with Dr. Melanie Richter. Due to space limitations of this thesis, the paper has been kept in its original formatting. References for the following chapter are self-contained within the manuscript and may be repeated in the thesis references.

Prior to this project, only a few studies had identified the role of TAOK2 in neurodevelopment (Calderon de Anda et al., 2012; Ultanir et al., 2014; Yadav et al., 2017), and was primarily known to function through stress-related MAPK signaling. The study was conducted using Taok2 KO mouse model, using both in vivo and in vitro assays. I worked on all aspects of this project, and my major contribution focused on investigating the impact of *de novo* missense and rare-inherited variants on TAOK2 kinase function and its effect on neurite growth and synaptic formation. This study provided a backdrop to study ASD pathologies and was a starting point to use proximity-labeling proteomics to study *TAOK2* and other ASD-risk genes, as described in Chapter 3 and Chapter 4.

Dr. Melanie Richter and Nadeem Murtaza share co-first authorship and equally contributed to the manuscript and performed all experiments with assistance. Dr. Karun Singh, Dr. Froylan Calderon de Anda, Dr. Melanie Richter and I designed the study and all experiments. I harvested and prepared mouse brains for MRI experiments. Zsuzsa Lindenmaier, Jacob Ellegood, and Dr. Jason P. Lerch performed MRI of the mouse brains. Jan Sedlacik and Dr. Jens Fiehler performed all diffusor tensor imaging on live mice. Sarah Scharf, Vanessa Kraus, Ronja Dork, Jacok Hellman and Dr. Fabio Morellini performed all mouse behavior assays. Henrike Hartung and Dr. Illeana L. Hanganu-Opatz performed all in vivo mouse electrophysiology experiments. Sean H. White conducted in

situ mouse brain slice electrophysiology experiments. Dr. Melanie Richter performed neuron morphology imaging experiments in in utero electroporated mice and mouse cultures, with assistance from Robin Scharrenberg, Ole Johanns, Birgit Schwanke, Bianca Bedürftig, Melad Henis. Susan Walker, Dr. Ryan K.C. Yuen, and Dr. Stephen W. Scherer provided information on the TAOK2 variants and generated the lymphoblastoid cells lines (LCLSs) and I performed all experiments with the LCLs. I generated TAOK2 variant constructs with assistance from Vickie Kwan. I performed all experiments with HEK293 cells and mouse neurons transfected with the TAOK2 constructs. Data was analyzed and figured were prepared by Dr. Melanie Richter and myself.

2.1 Published Manuscript

Molecular Psychiatry (2019) 24:1329–1350
<https://doi.org/10.1038/s41380-018-0025-5>

ARTICLE



Altered TAOK2 activity causes autism-related neurodevelopmental and cognitive abnormalities through RhoA signaling

Melanie Richter¹ · Nadeem Murtaza^{2,3} · Robin Scharrenberg¹ · Sean H. White^{2,3} · Ole Johanns¹ · Susan Walker^{4,5} · Ryan K. C. Yuen^{4,5} · Birgit Schwanke¹ · Bianca Bedürftig¹ · Melad Henis^{1,6} · Sarah Scharf⁷ · Vanessa Kraus⁷ · Ronja Dörk⁷ · Jakob Hellmann⁷ · Zsuzsa Lindenmaier^{8,9} · Jacob Ellegood⁸ · Henrike Hartung^{10,13} · Vickie Kwan^{2,3} · Jan Sedlacik¹¹ · Jens Fiehler¹¹ · Michaela Schweizer¹² · Jason P. Lerch^{8,9} · Ileana L. Hanganu-Opatz¹⁰ · Fabio Morellini⁷ · Stephen W. Scherer^{4,5} · Karun K. Singh^{2,3} · Froylan Calderon de Anda¹

Received: 12 May 2017 / Revised: 1 December 2017 / Accepted: 6 December 2017 / Published online: 21 February 2018
© The Author(s) 2018. This article is published with open access

Abstract

Atypical brain connectivity is a major contributor to the pathophysiology of neurodevelopmental disorders (NDDs) including autism spectrum disorders (ASDs). *TAOK2* is one of several genes in the 16p11.2 microdeletion region, but whether it contributes to NDDs is unknown. We performed behavioral analysis on *Taok2* heterozygous (Het) and knockout (KO) mice and found gene dosage-dependent impairments in cognition, anxiety, and social interaction. *Taok2* Het and KO mice also have dosage-dependent abnormalities in brain size and neural connectivity in multiple regions, deficits in cortical layering, dendrite and synapse formation, and reduced excitatory neurotransmission. Whole-genome and -exome sequencing of ASD families identified three de novo mutations in *TAOK2* and functional analysis in mice and human cells revealed that all the mutations impair protein stability, but they differentially impact kinase activity, dendrite growth, and spine/synapse development. Mechanistically, loss of *Taok2* activity causes a reduction in RhoA activation, and pharmacological enhancement of RhoA activity rescues synaptic phenotypes. Together, these data provide evidence that *TAOK2* is a neurodevelopmental disorder risk gene and identify RhoA signaling as a mediator of *TAOK2*-dependent synaptic development.

Introduction

Thousand and one amino-acid kinase 2 (*TAOK2*) is a family member of the mammalian sterile 20 (STE20)-like kinases and is implicated in neurodevelopmental disorders (NDDs) [1–4]. *TAOK2* is located in the autism spectrum

disorder (ASD) and schizophrenia-associated 16p11.2 chromosomal deletion region [5–8] and is associated with other neurodevelopmental phenotypes [9]. *TAOK2* is also present in the 16p11.2 duplication syndrome, which has distinct and reciprocal neurological phenotypes [9]. The 16p11.2 region also harbors *KCTD13*, *MAPK3*, and *SEZ6L2*, which may contribute to neurological phenotypes [7, 10–13]. Additional evidence comes from a genome-wide association study for psychosis that identified a significant single-nucleotide polymorphism in *TAOK2* [14]. Finally, *TAOK2* mRNA is a target of fragile X mental retardation protein (FMRP) [15]. Despite these suggestive studies, there is no direct evidence using mouse models or human cell models that genetic alterations in *TAOK2* cause NDDs.

Studies of ASD have uncovered many genes and signaling pathways, and one of the leading hypotheses implicates altered synapse formation and plasticity. Large genetic sequencing studies have revealed that mutations in different synaptic genes can cause ASD or specific NDDs [16–23]. Affected synaptic pathways include adhesion molecules [24, 25], regulatory translational proteins [26–30], ion channels [18, 23, 31], and cytoskeleton mediators [32–36].

These authors contributed equally: Melanie Richter, Nadeem Murtaza.

Co-Senior Authors: Stephen W. Scherer, Karun K. Singh, Froylan Calderon de Anda.

Electronic supplementary material The online version of this article (<https://doi.org/10.1038/s41380-018-0025-5>) contains supplementary material, which is available to authorized users.

- ✉ Stephen W. Scherer
stephen.scherer@sickkids.ca
- ✉ Karun K. Singh
singhk2@mcmaster.ca
- ✉ Froylan Calderon de Anda
froylan.calderon@znmh.uni-hamburg.de

Extended author information available on the last page of the article

SPRINGER NATURE

Pharmacological treatment ameliorated some of these genetic forms of ASD providing the framework to identify novel therapies [37–41].

There are two isoforms of murine *Taok2*, α and β , but isoform-specific roles in the brain remain poorly studied. Most STE20s activate one or more of the mitogen-activated protein kinases (MAPKs), including c-Jun N-terminal kinase (JNK) and p38, although their biological responses do not always require MAPKs [42]. TAOK2 can regulate microtubule dynamics and organization [43] and autophosphorylation of TAOK2 is important to initiate kinase activity [44]. Reduced *Taok2* expression revealed a decrease in axonal growth and basal dendrite formation of excitatory neurons in the mouse cortex [45]. TAOK2 is also necessary during activity-dependent synapse growth [46]. Furthermore, phosphorylation of TAOK2 by Mammalian STE20-like kinase 3 (MST3) regulates synapse development through an association with Myosin Va [47] or its activation of Septin7 kinase, which stabilizes PSD95 [48]. These studies indicate that TAOK2 plays an emerging role in dendritic arborization and synapse maturation, suggesting that changes in TAOK2 activity may lead to abnormal synaptic connectivity and behavioral phenotypes. However, previous studies utilized incomplete knockdown methods to study TAOK2, leaving in question the importance of TAOK2 in vivo during neural circuit function and behavior.

In this study, we combined a heterozygous (Het) and knockout (KO) mouse model with genetic sequencing of ASD subjects to dissect the role of TAOK2 in neural circuit development (summarized in Supplementary Fig. 18b). Analysis of *Taok2* Het and KO mice revealed several gene dosage-dependent impairments in behavior, whole-brain connectivity, cortical layering, neuronal morphology, and synaptic function in cortical excitatory neurons similar to other models [49–53]. We also report novel missense and truncating loss-of-function (LOF) mutations in *TAOK2* in ASD subjects that differentially impair TAOK2 function and disrupt dendrite formation and synapse structure. Further examination showed a *Taok2*-dependent reduction in RhoA activity mediating synaptic defects, which can be rescued pharmacologically by enhancing RhoA activity. Taken together, our study reveals that *Taok2* is critical for neural circuit formation and function in an animal model, and the analysis of TAOK2 human mutations provide insight into the etiology of NDDs.

Materials and methods

Animals

C57BL6/J Taok2 KO (*Taok2* *-/-*) mice were generated and described by Kapfhamer et al. [54]. Animals were housed at

the Central Animal Facilities at McMaster University and University Medical Center Hamburg-Eppendorf, Hamburg. All procedures received the approval of the Animal Research Ethics Board (AREB) and the Institutional Animal Care and Use committee of the City of Hamburg, Germany (G48/13 and G43/16 acc. to the Animal Care Act, §8 from 18 May 2006). See supplementary methods for further detail.

Magnetic resonance imaging (MRI) of live mice

Live MRI was performed using a dedicated 7 Tesla small animal MRI (ClinScan, Bruker, Ettlingen, Germany) with a mouse head four element phased array receiver surface coil and a linear polarized rat body transmit coil. Mouse imaging was done at the Neuroradiology/University Medical Center Hamburg-Eppendorf based on previous publications [55–57]. See supplementary methods for further detail.

In vivo electrophysiology

Mouse pups were initially anesthetized with isoflurane, fixed into the stereotaxic apparatus, and local anesthetic was administered. Multi-site electrodes (NeuroNexus, Ann Arbor, MI) were inserted to the skull surface into the prefrontal cortex (PFC) and hippocampus (HC). Two silver wires were inserted into cerebellum and served as ground and reference electrodes. Simultaneous recordings of local field potential (LFP) and multi-unit activity were performed from the prelimbic subdivision of the PFC and the CA1 area of the intermediate HC as described [58]. See supplementary methods for further detail.

In situ electrophysiology

Coronal brain slices (400 μ m) were prepared and miniature excitatory postsynaptic currents were recorded as previously described [35]. For miniature inhibitory postsynaptic currents recordings, 1 μ M tetrodotoxin and 1 mM kynurenic acid was used in bathing medium. See supplementary methods for further detail.

Behavior analysis

Behavioral experiments were performed with 10- to 18-week-old mice during the dark cycle in a room illuminated with dim red light. Tests started and ended at least 2 h after light offset and 3 h before light onset, respectively. Tracks representing the position of the mice were created and analyzed with EthoVision (Noldus, Wageningen, The Netherlands) [59]. Manual scoring of behavior was performed by a trained experimenter blinded to the

genotype of the mice using *The Observer* software (Noldus). See supplementary methods for further detail on the behavioral tests we used that are described previously [60–62].

Electron microscopy

Coronal vibratome sections of the cingulate cortex (cg1 and cg2) and the prelimbic cortex (PL) of the PFC, the primary somatosensory regions S1HL, S1FL, S1BF, and the intermediate HC were collected and prepared for electron microscopy as described [63]. Semithin sections (0.5 μ m) were prepared for light microscopy mounted on glass slides and stained for 1 min with 1% Toluidine blue. Ultrathin sections (60 nm) were examined in an EM902 (Zeiss, Munich, Germany). Pictures were taken with a Mega-ViewIII digital camera (A. Tröndle, Moorenweis, Germany).

Analysis of dendritic and spine morphology and spine motility

Sholl analysis was conducted using the semi-automatized Simple Neurite Tracer plug-in on Fiji (ImageJ) and analyzed using the Sholl analysis plug-in. To quantify dendrite spine morphology, image stacks were uploaded to Fiji and a semi-automatized protocol [64] was used for categorization based on spine-head width and spine length. For spine motility assays, cultures were imaged at $\times 63$ magnification and recorded at a 2-s frame rate for 5 min. Raw images were uploaded to Fiji and analysis was done by the Dendritic Filopodia Motility Analyzer [65]. For further detailed description of analyses, please refer to supplementary methods.

Sample collection and whole-genome sequencing of ASD families

We obtained informed consents, or waivers of consent, which were approved by the Western Institutional Review Board, Montreal Children's Hospital—McGill University Health Centre Research Ethics Board, McMaster University—Hamilton Integrated Research Ethics Board, Eastern Health Research Ethics Board, Holland Bloorview Research Ethics Board, and the Hospital for Sick Children Research Ethics Board. Whole-genome sequencing was performed as previously described in [66].

Biochemical assays for protein expression, activity, and interaction

For detailed description of the analyses, please refer to supplementary methods.

Statistical analysis

Data are expressed as mean \pm s.e.m. Minimums of three mice per condition, or three mouse litters for in vitro culture experiments, were used for statistical analysis. We used the Student's unpaired *t*-test, Wilcoxon signed-rank pair test, one-sample *t*-test, one-way analysis of variance (ANOVA), two-way ANOVA, mixed three-way ANOVA and post hoc Tukey, Dunnett, and Bonferroni tests in GraphPad Prism 7 statistical software for statistical analyses. Dunnett's test was utilized in all cases to compare all conditions with the control conditions, except when it was necessary to compare multiple conditions, in which case the Tukey test was utilized. Bonferroni's test was used for all behavior tests due to sample size and differences in variance. The *p*-values in the figure legends are from the specified tests, and *p* < 0.05 was considered statistically significant.

Data availability

Whole-genome data sets were generated during and analyzed during the current study are available at the MSSNG repository, <http://research.mss.ng>.

Results

Taok2 KO mice display brain morphological and behavioral abnormalities

We analyzed the anatomy of *Taok2* KO brains using MRI on fixed 8- to 10-week-old mouse brains and examined 182 independent regions. The absolute brain volume of *Taok2* KO mice was significantly enlarged compared with WT mice (Figs. 1a, b) derived from absolute and relative volumetric increases in the hindbrain, midbrain, hypothalamus, thalamus, cerebellum, and HC (Supplementary Fig. 1i and Supplementary Table 1) but a relative decrease in the somatosensory cortex (Figs. 1a, b and Supplementary Table 1) suggesting that the increase in brain volume is caused primarily by these regions. We also found significant decreases in the relative brain volumes of the corpus callosum, many cortical regions, the anterior commissure, and the olfactory bulbs (Supplementary Fig. 1j and Supplementary Table 1). *Taok2* Het mice also showed significant increases in brain volume, but not as dramatic as KO mice, consistent with a gene dosage effect (Fig. 1b). Furthermore, *Taok2* Het mice show trends (*p*-value < 0.05, but false discovery rate (FDR)-adjusted *p*-value > 0.05) similar to *Taok2* KO mice in absolute or relative volumes of regions such as the midbrain, thalamus, hypothalamus, and hindbrain regions, suggesting these regions are strongly affected

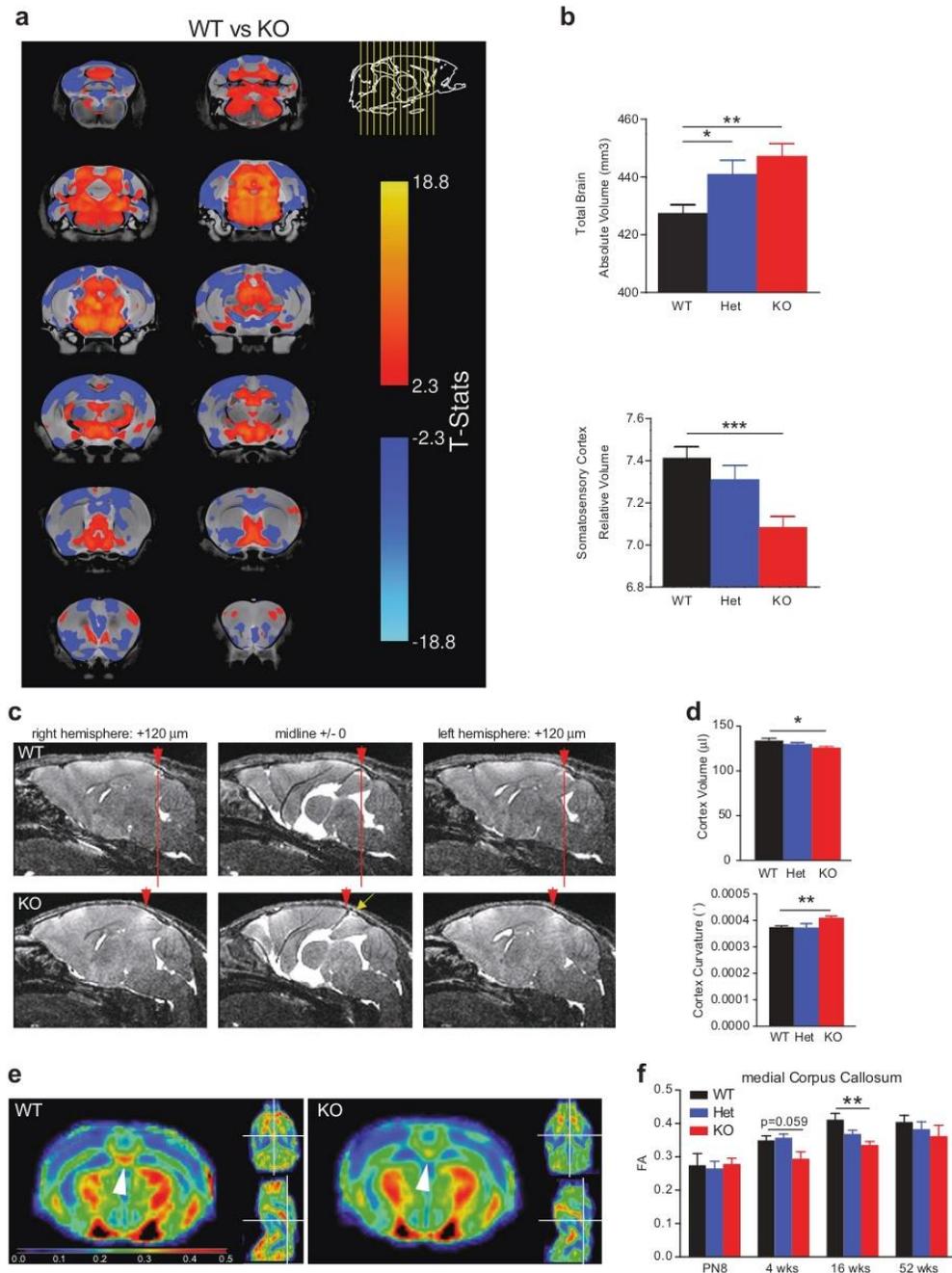
by loss of *Taok2* (Supplementary Fig. 1f, i and Supplementary Table 1). The same differences and trends were seen when comparing genotypes within each sex (Supplementary Figure 1g-j, and Supplementary Table 1). Additional live MRI on P28 mice (Figs. 1c, d) and volumetric analysis of whole non-fixed brain tissue (Supplementary Fig. 1a-e) confirmed that cortices of *Taok2* KO mice are smaller than WT and *Taok2* Het mice. Finally, a longitudinal study on the same mice using diffusion tensor imaging (DTI) to map fiber tracks uncovered a regional delay of the development of neuronal tracks such as the corpus callosum in *Taok2* KO mice across multiple time points (Figs. 1e, f and Supplementary Fig. 1k-o).

Given the changes of cortical brain size, we asked whether alterations in cortical layering may contribute to these effects. Frontal, medial, and dorsal coronal brain sections were collected for the analysis. We immunostained for the upper cortex marker *Cux-1* and the lower cortex marker *Ctip2* to analyze laminar organization including thickness and density of cortical layers. Our results show that the thickness of the *Ctip2* layer is not changed in *Taok2* KO cortices (Supplementary Fig. 2b). However, the medial and dorsal region of the cortex showed decreases in the thickness of the *Cux-1*-positive layer (Supplementary Fig. 2c, d), together with an overall reduced cortex thickness in the dorsal-caudal region of the *Taok2* KO cortices (Supplementary Fig. 2a). Importantly, density of *Cux-1*-positive cells is not altered in the *Taok2* KO cortex compared with WT littermates (Supplementary Fig. 2e). Detailed examination of the cellular distribution of *Cux-1*+ cells uncovered a redistribution of cells in KO cortices. Specifically, we found that *Cux-1*+ cells clustered more in the superficial portion of the upper cortical plate, especially in the medial-caudal and dorsal cortex (Supplementary Fig. 2f, g). Given that *Cux-1*+ cell density is not affected in KO cortices, these results suggest that neuronal migration defects might be the cellular substrate of these cytoarchitectural abnormalities. We directly examined neuronal migration in *Taok2* Het and KO mice and found defects in the migration of neurons born at embryonic day 15/16 that produce cortical layers 2/3 (data not shown).

To test if the anatomical changes relate to the neural communication defects, we examined the synchrony between oscillatory patterns of electrical activity in the prelimbic subdivision (PL) of the PFC and HC. Previous studies have shown that hippocampal–PFC connectivity is altered during anxiety, spatial learning, and memory-related tasks in rodents [67–70]. In vivo extracellular recordings of the LFP and multiple unit activity were conducted in postnatal day P8–10 mice because this is the period of maximal drive from HC to PL and is critical for prelimbic-

hippocampal network maturation [58, 71] (Supplementary Fig. 3a-f). The oscillatory events were similar in the PFC and slightly decreased in the HC of *Taok2* KO mice; however, the duration, amplitude, and power in theta (4–12 Hz), beta (12–30 Hz), and gamma (30–100 Hz) frequency ranges were significantly augmented (Supplementary Fig. 3g-n and Supplementary Table 2) [72, 73]. The coherence within the beta band was also significantly increased in *Taok2* KO mice suggesting alterations in HC and PFC connectivity, and alterations in long-range functional connectivity (Supplementary Fig. 3o, p).

Next, we performed behavioral testing on *Taok2* Het and KO mice. We assessed novelty-induced exploration and anxiety in the open field test, where *Taok2* KO mice traveled longer distances (effect of genotype: $F_{2,77} = 12.89$; $p < 0.001$) and further away from the walls (effect of genotype: $F_{2,77} = 8.38$; $p < 0.001$, Figs 2a, b) during the 30-min trial, with increased time spent in the center and decreased time spent in the border (Supplementary Fig. 4c, d). The analyses of 5-min time bins revealed that locomotion of *Taok2* KO was enhanced only at time points 15–30 min compared with WT mice (effect of the interaction “genotype \times time bin”: $F_{10,385} = 2.91$; $p = 0.002$, Supplementary Fig. 4a,b), suggesting that the enhanced locomotion of *Taok2* KO mice was due to impaired short-term habituation. We further analyzed anxiety-related behavior in the elevated plus maze test. *Taok2* KO mice showed a significant increase in the time spent in the open arms of the elevated plus maze (effect of genotype $F_{2,56} = 5.16$; $p = 0.009$, Fig. 2c), with no differences in the number of entries or time spent in the closed arm or time spent in the center (Supplementary Fig. 4e-g). Social behavior was assessed in a social preference paradigm: We found *Taok2* KO mice spent less time sniffing an unfamiliar sex-matched mouse instead of an unfamiliar object and compared with the WT littermates indicating reduced social drive (effect of the interaction genotype \times stimulus: $F_{2,63} = 4.96$; $p = 0.001$, Fig. 2d and Supplementary Fig. 5a, b). No difference in distance moved was detected in the social preference test (Supplementary Fig. 5c). Working memory was assessed in the Y-maze for spontaneous alternation. *Taok2* KO mice performed less alternations than WT mice (effect of genotype: $F_{2,70} = 7.26$; $p = 0.001$; Fig. 2e). The impaired performance of the *Taok2* KO mice does not seem to be caused by their tendency to move faster than WT mice as the average transition time between arms did not differ from WT mice (Supplementary Fig. 5e) and percentage of alternations did not correlate with the average transition time at the individual level (Supplementary Fig. 5f). We also analyzed short-term spatial memory in the object recognition paradigm and found that *Taok2* KO mice showed a reduced preference for the displaced object when compared with WT mice (effect of genotype: $F_{2,64} = 9.05$;



$p < 0.001$; Fig. 2f), with no difference in the total distance moved (Supplementary Fig. 5d). Long-term storage and retrieval of emotionally relevant information was also examined using the contextual fear-conditioning paradigm.

Taok2 Het and KO male mice spent less time immobile compared with WT male littermates (effect of the interaction "genotype \times sex": $F_{2,45} = 7.21$; $p = 0.002$; Fig. 2g), with no difference in the time spent immobile during the

◀ **Fig. 1** *Taok2* KO mice have altered brain morphology and brain activity. **a** Voxel-wise analysis highlighting significant differences in relative volume (images show the lowest threshold of 5% false discovery rate (FDR) for *Taok2* KO mice) throughout the brain between the WT and *Taok2* KO mice. T-statistic of 2.3–18.8 indicates decreasing false discovery rate, where 2.3 = 5% FDR and positive or negative T-stat indicates positive or negative change compared with WT brain. **b** Top: *Taok2* KO mice have increased absolute brain volume compared with WT mice (WT = 16, Het = 13, KO = 23 mice from three different cohorts, statistics by linear model; WT vs. KO $p = 0.0015$, WT vs. Het $p = 0.0223$). Bottom: the relative volume of the somatosensory cortex is reduced in *Taok2* KO mice (WT = 16, Het = 13, KO = 23 mice from three different cohorts, statistics by linear model corrected for multiple comparisons using FDR; WT vs. KO $p = 0.0002$, WT vs. Het $p = 0.5569$). **c** *Taok2* KO mice show shortening and increased curvature of the cortex at 4 weeks of age in vivo. Red arrowheads indicate dorsal end of the cortex, red line indicates shortening of the *Taok2* KO cortex, and yellow arrow indicates gap between cortex and colliculi. **d** Top: decreased cortex volume (μl) in *Taok2* KO mice (WT = 7, Het = 12, KO = 11 mice from three different cohorts; one-way ANOVA, post hoc Dunnett's test; $F_{2,27} = 4.369$, $p = 0.0027$; WT vs. KO $p = 0.0129$). Bottom: increased curvature of the cortex ($^{\circ}$ degree $\times e^{-0.006}$) in *Taok2* KO mice brains (WT = 7, Het = 12, KO = 11 from three different cohorts; one-way ANOVA, post hoc Dunnett's test; $F_{2,27} = 3.142$, $p = 0.0593$; unpaired t -test; WT vs. KO $p = 0.0025$). **e** Representative diffusion tensor images of WT and *Taok2* KO mouse brains. White arrowhead indicates reduced fiber density in medial corpus callosum region in *Taok2* KO mice. Blue to red indicates increased fractional anisotropy. **f** Reduced fiber track density measured by fractional anisotropy (FA) in the medial corpus callosum of *Taok2* KO mice brains. (PN8: WT = 6, Het = 10, KO = 7; 4 weeks: WT = 7, Het = 12, KO = 10; 16 weeks: WT = 7, Het = 15, KO = 11; 52 weeks: WT = 5, Het = 6, KO = 6 mice from three cohorts; one-way ANOVA, post hoc Dunnett's test; PN8: $F_{2,20} = 0.08947$, $p = 0.9148$; 4 weeks: $F_{2,36} = 4.832$, $p = 0.0164$, WT vs. KO $p = 0.0598$; 16 weeks: $F_{2,30} = 7.241$, $p = 0.0027$, WT vs. KO $p = 0.0013$; 52 weeks: $F_{2,14} = 0.6358$, $p = 0.5441$). * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$. Values are mean \pm s.e.m.

baseline and conditioning trial (Supplementary Fig. 5g, h) indicating that loss of *Taok2* impairs consolidation or retention of emotional memories specifically in male mice. Finally, the water maze test was used to assess long-term spatial memory during a transfer trial performed 24 h after the last training session. *Taok2* Het and KO mice showed no difference in the learning curves, with no difference in total distance swum, mean velocity, and time spent at the border (Supplementary Fig. 6). Whereas all genotypes showed a preference for the target quadrant (Fig. 2i), *Taok2* KO mice searched more consistently for the platform at its former position than WT mice as indicated by the reduced mean minimal distance to platform (effect of genotype: $F_{2,59} = 8.52$; $p < 0.001$; Fig. 2h) and enhanced time spent at the platform position and platform crossings (Supplementary Fig. 7a-c). Although this seems paradoxical, other ASD mouse models have displayed reduced times to find the hidden platform indicating enhanced spatial learning phenotypes [74, 75]. These findings suggest that loss of *Taok2*

results in behavior alterations related to cognition, anxiety, and social interaction.

Taok2 KO mice display dendritic morphology and synaptic functional deficits

To complement the behavioral and in vivo electrophysiology experiments, we assessed neural morphology in vivo using Golgi-Cox staining on 3-week-old WT and *Taok2* Het and KO brains. Analysis of prefrontal *Taok2* Het and KO neurons revealed a significantly reduced basal dendrite complexity and length compared with WT neurons, with only minor reductions in the apical dendrites (Figs. 3a-c). Examination of somatosensory neurons showed a small and partially significant difference in dendritic arborization in *Taok2* KO and WT cells (Supplementary Fig. 8a-c), but not in hippocampal neurons (Supplementary Fig. 9a-c).

Next, we analyzed whether *Taok2* regulates synapse development in vivo and found that *Taok2* KO neurons had significant changes in spine distribution (Figs. 3d, e) and a reduction in the total number of basal dendrite spines per PFC neuron (Fig. 3f) but not in the somatosensory cortex or the HC (Supplementary Fig. 8d-f and Supplementary Fig. 9d-f). Furthermore, all three regions showed no difference in the number of apical dendrite spines (Fig. 3f, Supplementary Fig. 8f and Supplementary Fig. 9f). *Taok2* Het neurons showed a gene dosage-dependent difference in spine distribution and significant reduction in total basal dendrite spine numbers in PFC neurons only (Figs. 3e-f). In addition, our semi-automatized analysis of spine morphology [64] revealed shorter spine length and head width in *Taok2* KO neurons and shorter head width in *Taok2* Het neurons (Fig. 3h). Concurrently, the number of thin and stubby (immature) spines significantly increased concomitantly with a reduction in mushroom-shaped (mature) spines in *Taok2* KO neurons (Figs. 3g, i). Interestingly, *Taok2* Het neurons showed a significant increase in long thin (immature spines) and decrease in mushroom-shaped spines, but no changes in stubby spines (Fig. 3i). This phenotype was also detected in the somatosensory cortex (Supplementary Fig. 8g, h) showing an increase in thin spines, but not in the HC (Supplementary Fig. 9g, h). Electron microscopy analysis of *Taok2* KO brains revealed that the number of synapses formed onto the dendrite shaft instead of the postsynaptic spine heads was increased in prefrontal and somatosensory cortical neurons (Figs. 3j, k and Supplementary Fig. 8i), but not in the HC (Supplementary Fig. 9i). Supporting these findings, we found that the level of phosphorylated *Taok2* is higher in the cortex compared with the HC (Supplementary Fig. 9j, k). Together our findings indicate that *Taok2* is a key regulator of synapse formation predominantly in the PFC, and also the somatosensory cortex in vivo.

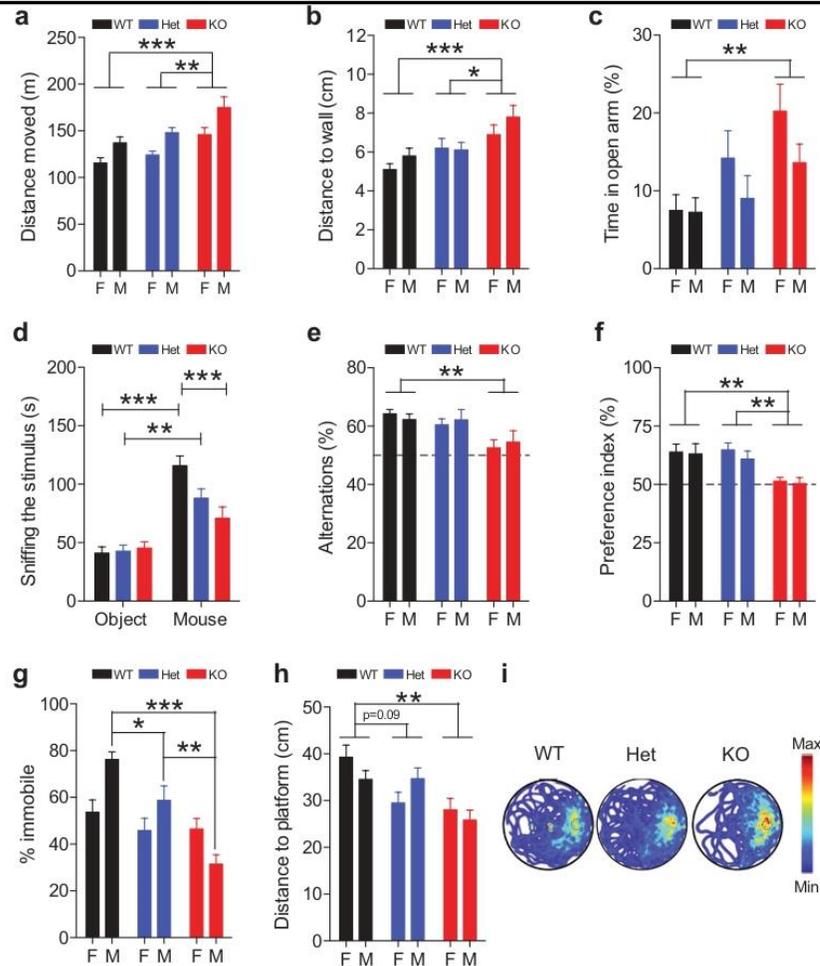


Fig. 2 *Taok2* KO mice show alterations in neurodevelopmental disorder-related mouse behavior. **a** *Taok2* KO mice travel longer distances in the open field. (WT(M) = 12, WT(F) = 16, Het(M) = 12, Het(F) = 17, KO(M) = 9, KO(F) = 17 mice from three different cohorts; two-way ANOVA, post hoc Bonferroni's test, effect of genotype $F_{2,77} = 12.89$, $p < 0.001$; WT vs. KO $p < 0.001$, Het vs. KO $p < 0.01$). **b** *Taok2* KO mice moved at longer distances from the wall in the open field test. (WT(M) = 12, WT(F) = 16, Het(M) = 12, Het(F) = 17, KO(M) = 9, KO(F) = 16 mice from three different cohorts; two-way ANOVA, post hoc Bonferroni's test, effect of genotype $F_{2,76} = 8.38$, $p < 0.001$, WT vs. KO $p < 0.001$, Het vs. KO $p < 0.05$). **c** *Taok2* KO mice spent more time in the open arm of an elevated plus maze compared with WT. (WT(M) = 10, WT(F) = 10, Het(M) = 12, Het(F) = 12, KO(M) = 9, KO(F) = 9 mice from three different cohorts; two-way ANOVA, post hoc Bonferroni's test; effect of genotype $F_{2,56} = 5.16$, $p = 0.009$; WT vs. KO $p < 0.01$). **d** *Taok2* KO mice spent less time investigating and sniffing an unfamiliar sex-matched mouse instead of an unfamiliar object compared with WT littermates. (WT(F) = 11, WT(M) = 11, Het(F) = 8, Het(M) = 15, KO(F) = 12, KO(M) = 12 mice from three different cohorts; three-way mixed ANOVA, post hoc Bonferroni; effect of the interaction "genotype \times stimulus" $F_{2,63} = 4.96$, $p = 0.001$; Mouse: WT vs. KO $p < 0.001$; Object vs. Mouse: WT $p < 0.001$, Het $p < 0.01$). **e** *Taok2* KO did not perform more alternations than chance and performed fewer alternations than WT

littermates. (WT(M) = 11, WT(F) = 14, Het(M) = 13, Het(F) = 15, KO(M) = 8, KO(F) = 15 mice from three different cohorts; two-way ANOVA, post hoc Bonferroni's test, effect of genotype $F_{2,70} = 7.26$, $p = 0.001$; WT vs. KO $p < 0.01$). **f** *Taok2* KO mice have a lower preference index than WT mice for a displaced object. (WT(M) = 11, WT(F) = 11, Het(M) = 15, Het(F) = 11, KO(M) = 10, KO(F) = 12 from three different cohorts; two-way ANOVA, post hoc Bonferroni's test, effect of genotype $F_{2,64} = 9.05$, $p < 0.001$; WT vs. KO $p < 0.01$, Het vs. KO $p < 0.01$). **g** *Taok2* Het and KO mice spent less time immobile during recall trial for contextual fear conditioning. (WT(M) = 8, WT(F) = 9, Het(M) = 9, Het(F) = 9, KO(M) = 7, KO(F) = 9 from three different cohorts; three-way mixed ANOVA, post hoc Bonferroni; effect of the interaction "genotype \times sex" $F_{2,45} = 7.21$, $p = 0.002$; WT vs. KO $p < 0.001$, Het vs. KO $p < 0.01$, WT vs. Het $p < 0.05$). **h** *Taok2* KO showed reduced mean minimal distance to the platform during the recall trial of the water maze test. (WT(M) = 10, WT(F) = 12, Het(M) = 11, Het(F) = 12, KO(M) = 8, KO(F) = 12 mice from three different cohorts; two-way ANOVA, post hoc Bonferroni's test, effect of genotype $F_{2,59} = 8.52$, $p < 0.001$; WT vs. Het $p = 0.09$, WT vs. KO $p < 0.01$). **i** Heat maps showing *Taok2* KO male mice search for platform more at rear position compared with WT and Het male mice. Blue to red indicates increased probability of a mouse being present. ns > 0.05 , * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$. Values are mean \pm s.e.m.

Next, we tested synaptic transmission using in situ electrophysiology. We performed whole-cell patch-clamp recordings of acute brain slices (P21–28) and found a significant reduction in the mean frequency of miniature excitatory postsynaptic currents (mEPSCs), but not the mean amplitude, in the PFC (Fig. 3l–n) and the somatosensory cortex (Supplementary Fig. 8j–k). *Taok2* Het brain slices also showed a strong reduction in the frequency of mEPSCs (Fig. 3m and Supplementary Fig. 8k). We also measured miniature inhibitory postsynaptic currents but found no differences (Supplementary Fig. 10). These data indicate that *Taok2* Het and KO mice display multiple abnormalities in neuronal morphology and synaptic function in cortical excitatory neurons.

Taok2 is present in the postsynaptic density and regulates synapse formation in vitro

We also assessed whether Taok2 is expressed at the postsynaptic density (PSD). We observed that Taok2 colocalizes with the postsynaptic protein SynGAP (Supplementary Fig. 11a) and, when isolated in synaptosomes from P28 mouse brains, we found both, phosphorylated and non-phosphorylated Taok2 in the PSD95-positive fraction (Supplementary Fig. 11b). We also examined synapse formation in vitro using WT and *Taok2* KO cortical neuron cultures and determined that *Taok2* KO neurons display a significant reduction in spines and a reduction in synaptic SynGAP-positive puncta that colocalize with phalloidin-stained (F-actin enriched) dendritic spines (Supplementary Fig. 11c–e). Acute downregulation of *Taok2* using short hairpin RNA (shRNA) in WT cultured cortical neurons further showed a dosage-dependent decrease in dendritic spine density (Supplementary Fig. 11f, g) and an increase in mobility of dendritic spines by tracking farnesylated (membrane-bound)-green fluorescent protein (GFP)-stained spines and Lifeact-GFP stained actin-enriched protrusions (Supplementary Fig. 11h, i). These in vitro data indicate that TAOK2 regulates spine and synaptic deficits in a cell-autonomous manner by affecting actin dynamics and stability.

Identification of de novo and LOF mutations in TAOK2 from ASD cohorts via whole-genome sequencing

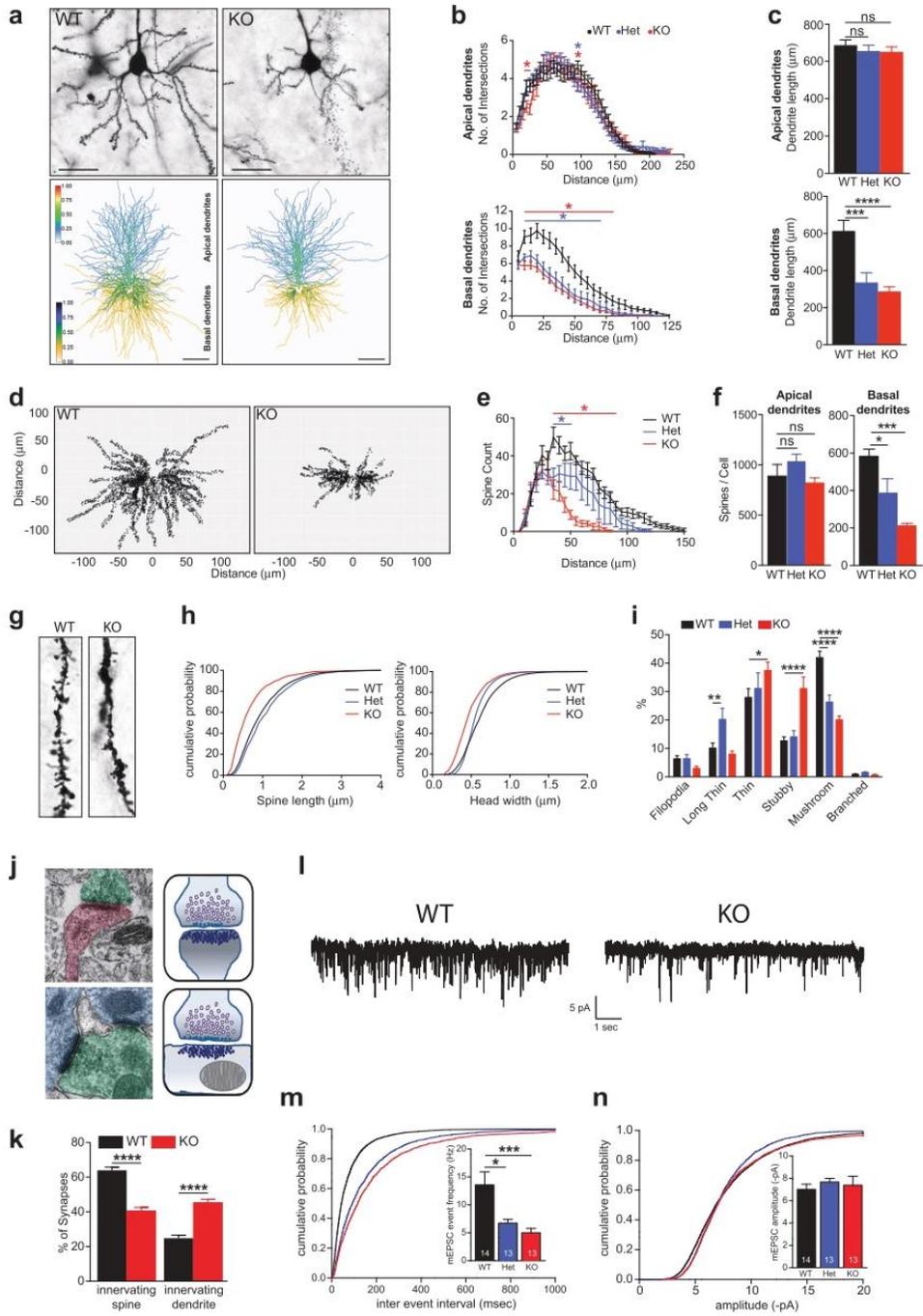
Analysis of *TAOK2* from the Exome Aggregation Consortium (ExAC) reveals it is highly constrained for LOF mutations ($pLi = 1$; the highest possible score) and missense mutations ($z = 3.54$) [76], suggesting that mutations in *TAOK2* could have deleterious effects. We used whole-genome and -exome sequencing of over 2600 families with ASD to detect de novo or inherited genetic variants [77].

We detected 24 different variants in *TAOK2* (Supplementary Table 3 and Supplementary Fig. 12a), which were confirmed by Sanger sequencing (data not shown). We focused on six genetic alterations from unrelated families (Table 1, Fig. 4a and Supplementary Fig. 12b), of which three are known to be de novo; a missense mutation in the kinase domain (A135P), a C-terminal frameshift deletion resulting in truncation (P1022*; this proband also has a frameshift mutation in *CHD8*) and a de novo splice site variant (c.563 + 12_563 + 15 del) predicted to cause intron 7 retention (Supplementary Fig. 13c). These de novo mutations in *TAOK2* are not present in the ExAC database [20]. We also identified three additional protein-truncating variants in *TAOK2* (Table 1 and Supplementary Fig. 12a), where one is a rare-inherited variant and two have unknown genetic inheritance. All six probands have been diagnosed clinically with autism with their diagnostic test scores summarized in Table 1. Given the importance of *de novo* mutations, they were functionally assessed. Droplet digital PCR confirmed that the mutations are germline (data not shown). We then tested the gene expression of *TAOK2* in lymphoblastoid cell lines (LCLs) of patients and found no significant changes for all isoforms (Supplementary Fig. 13a, b).

Next, we tested whether intron 7 was retained on the de novo splice site mutation (c.563 + 12_563 + 15 del). We found a low-level retention (approximately 1%) in all family members and unrelated wild-type *TAOK2* individuals. Interestingly, this proband showed a significantly higher level of intron retention (approximately 13%), which introduced a premature stop codon (Supplementary Fig. 13c). Although the impact of this remains unknown and requires further testing, it indicates the splice site mutation is detrimental to normal TAOK2 splicing.

De novo mutations in TAOK2 alter kinase activity, protein stability, neuronal localization, and dendritic spine motility

We functionally assessed the de novo mutations and two rare-inherited variants and tested whether they impaired auto-phosphorylation at serine 181 in both α and β isoforms [44]. We studied the A135P mutation (present in both isoforms) and the P1022* mutation (present only in the β isoform) along with the A335V and H781R rare-inherited mutations (present in both or in the β isoform only, respectively) (Fig. 4a). We expressed WT TAOK2 α , TAOK2 β and the respective mutants in HEK293 cells and analyzed TAOK2 phosphorylation and expression. We found a striking reduction in TAOK2 auto-phosphorylation caused by the A135P mutation in both isoforms, identifying this as a kinase dead mutation (Figs. 4b, c and Supplementary Fig. 14a, b), whereas the P1022* mutation did not



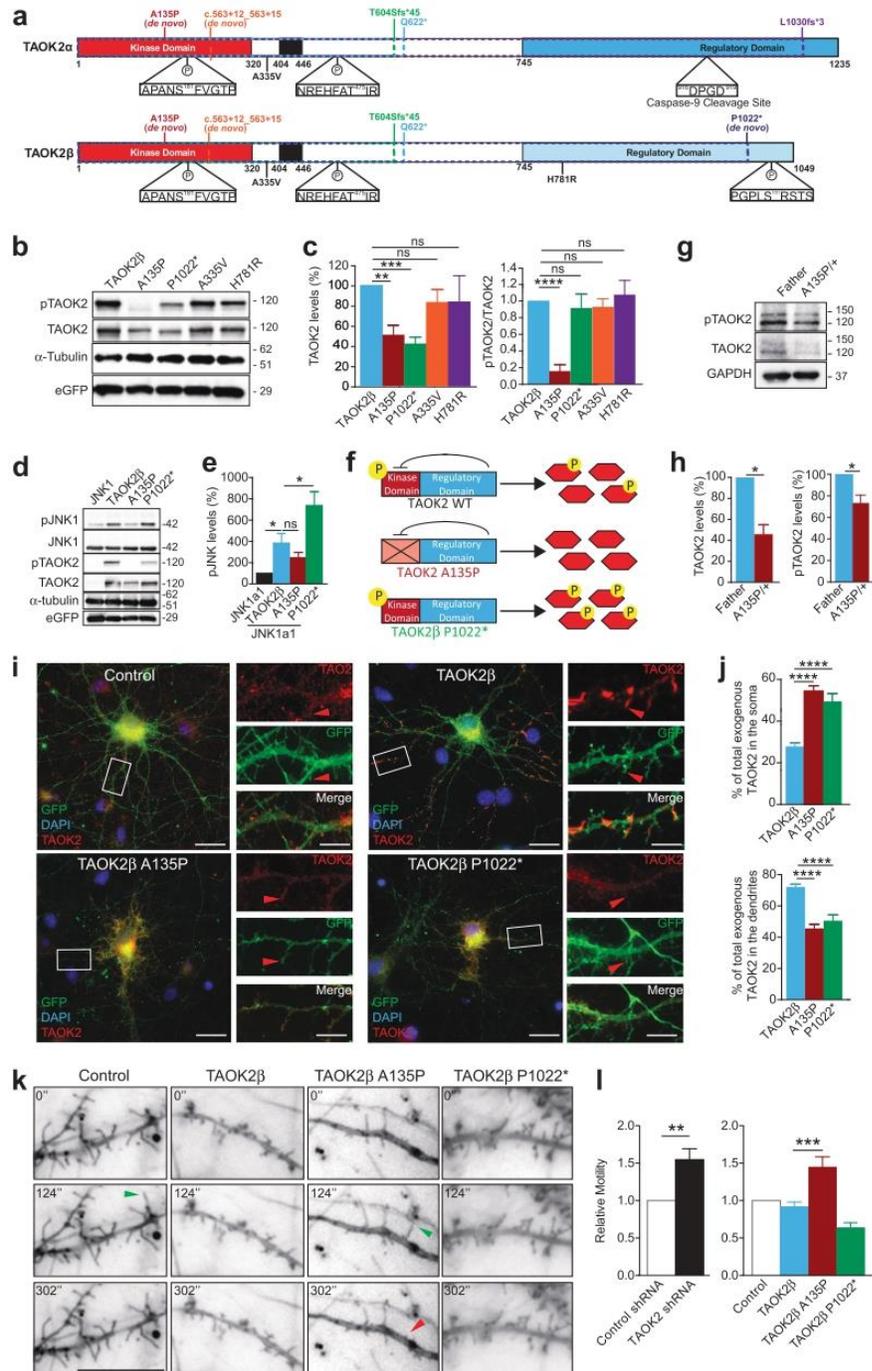
◀ **Fig. 3** *Taok2* KO mice have reduced dendrite growth and synaptic connectivity in the prefrontal cortex. **a** Top: Golgi-stained PFC neurons from P21 WT, *Taok2* Het and KO mice. Scale bars represent 20 μ m. Bottom: dendritic heat maps of superimposed neuron tracings for each condition. Blue to red (apical) and yellow to blue (basal) indicates increased probability of dendrite presence. Scale bars represent 30 μ m. **b** Top: no major difference in apical dendritic complexity in layer 2 PFC neurons in *Taok2* Het and KO mice (WT = 26, Het = 32, KO = 21 neurons from three different brains; two-way ANOVA, post hoc Dunnett's test; $F_{2,3496} = 3.055$, $p = 0.0472$ between genotypes; *represents ranges of significance; WT vs. Het (blue), WT vs. KO (red); see supplemental statistics). Bottom: significantly reduced basal dendritic complexity in layer 2 PFC neurons in *Taok2* Het and KO mice (WT = 19, Het = 18 and KO = 18 neurons from three different brains; two-way ANOVA, post hoc Dunnett's test; $F_{2,1716} = 128.7$, $p < 0.0001$ between genotypes; *represents ranges of significance; WT vs. Het (blue), WT vs. KO (red); see supplemental statistics). **c** Top: no change in apical dendrite length (μ m) in *Taok2* KO PFC neurons (WT = 27, Het = 34, and KO = 22 neurons from three different mice brains; one-way ANOVA, post hoc Dunnett's test; $F_{2,80} = 0.3346$, $p = 0.7167$). Bottom: reduced basal dendrite length (μ m) in *Taok2* KO PFC neurons (WT = 22, Het = 17, and KO = 18 neurons from three different mice brains; one-way ANOVA, post hoc Dunnett's test; $F_{2,54} = 12.47$, $p < 0.0001$; WT vs. Het $p = 0.0007$, WT vs. KO $p = 0.0001$). **d** Dendritic spine distribution maps of Golgi-stained WT and *Taok2* KO PFC neurons. **e** *Taok2* KO PFC neurons show reduced number of distal dendritic spines (WT = 6, Het = 6, and KO = 6 neurons from three different brains; two-way ANOVA, post hoc Dunnett's test; $F_{2,465} = 89.35$, $p < 0.0001$ between genotypes; *represents ranges of significance; WT vs. Het (blue), WT vs. KO (red); see supplemental statistics). **f** Left: *Taok2* KO PFC neurons show no difference in total apical dendritic spines per cell (WT = 6, Het = 6, KO = 6 neurons from three different mice brains; One-way ANOVA, post hoc Dunnett's test; $F_{2,15} = 1.766$, $p = 0.2048$). Right: *Taok2* KO PFC neurons show decreased number of total basal dendritic spines per cell (WT = 6, Het = 6, KO = 6 neurons from three different mice brains; One-way ANOVA, post hoc Dunnett's test; $F_{2,15} = 13.76$, $p = 0.0004$; WT vs. Het $p = 0.0265$, WT vs. KO $p = 0.0002$). **g** Images of dendritic spines on P21 WT and *Taok2* KO PFC neuron dendrites. **h** Cumulative probability histograms show shift toward reduced dendritic spine lengths (left) *Taok2* KO PFC neurons and reduced head widths (right) in *Taok2* Het and KO PFC neurons. **i** *Taok2* KO PFC neurons have a significant increase in thin and stubby shaped spines and reduction in mushroom-like spines compared with WT PFC neurons, whereas *Taok2* Het PFC neurons have an increase in long thin shaped spines and reduction in mushroom-like spines (WT = 3504, Het = 2317, KO = 1262 spines from six cells per condition from three different brains; two-way ANOVA, post hoc Dunnett's test; $F_{2,90} = 0.0008294$, $p = 0.9992$ between genotypes; WT vs. Het filopodia: $p = 0.9997$, long thin: $p = 0.0088$, thin: $p = 0.5599$, stubby: $p = 0.8934$, mushroom: $p < 0.0001$, and branched: $p = 0.9765$; WT vs. KO filopodia: $p = 0.5111$, long thin: $p = 0.7565$, thin: $p = 0.0141$, stubby: $p < 0.0001$, mushroom: $p < 0.0001$, and branched: $p = 0.9943$). **j** Representative images of synapses innervating spines (top) and synapses innervating dendrites (bottom) and illustrations of each (right). **k** *Taok2* KO neurons imaged by electron microscopy show decreased percentage of synapses on postsynaptic spines (left) and increased synapses on dendrites (right) (WT = 159 and KO = 120 neurons; innervating spine: unpaired t -test; $t(277) = 8.814$, $p < 0.0001$; innervating dendrite: $t(277) = 8.262$, $p < 0.0001$). **l** Representative traces of mEPSC spikes from WT and *Taok2* KO PFC neurons. Scale: 5pA vs. 1 s. **m** Longer inter-event intervals in *Taok2* Het and KO PFC neurons shown on a cumulative probability histogram. Inside: reduced mean mEPSC event frequency in *Taok2* Het and KO PFC neurons (WT = 14, Het = 13 and KO = 13 neurons from three different mice brains; Kruskal-Wallis ANOVA, post hoc Dunn's test; H-value = 14.47, $p = 0.0007$;

WT vs. Het $p = 0.0212$, WT vs. KO $p = 0.0008$). **n** No change in mEPSC amplitude in *Taok2* WT, Het and KO PFC neurons shown on a cumulative probability histogram. Inside: mean mEPSC amplitude of *Taok2* WT, Het and KO neurons (WT = 14, Het = 13 and KO = 13 neurons from three different mice brains; kruskal-walis ANOVA, post hoc Dunn's test; H-value = 3.871, $p = 0.1444$; WT vs. Het $p = 0.1857$, WT vs. KO $p > 0.9999$). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ and **** $p < 0.0001$. Values are mean \pm s.e.m.

alter auto-phosphorylation (Figs. 4b,c). Interestingly, both mutations, A135P and P1022*, reduce protein expression of TAOK2 suggesting impaired protein stability (Figs. 4b, c). The A335V and H781R rare-inherited variants did not alter protein expression or auto-phosphorylation, suggesting that they have minimal impact on protein function and expression (Figs. 4b, c and Supplementary Fig. 14a, b). Therefore, further experiments were done using only the de novo mutations. Next, we analyzed the effect of these mutations on one of the known downstream targets of TAOK2, JNK1 [45, 78]. TAOK2 α phosphorylates JNK1 more than the β isoform, and accordingly we found that TAOK2 α bearing the A135P mutation significantly decreased JNK1 phosphorylation, whereas the A135P mutation in the β isoform showed no significant effect (Figs. 4d, e and Supplementary Fig. 14c, d). The P1022* mutation (C-terminal deletion), however, significantly increased phosphorylation of JNK1 compared with WT TAOK2 (Figs. 4d, e). This is consistent with previous reports identifying the C-terminus of TAOK2 as a negative regulator of TAOK2 kinase activity, predicting dysregulated (overactive) kinase activity [78, 79]. In summary, the A135P mutation impairs auto-phosphorylation and TAOK2 kinase function (LOF, analogous to the TAOK2 KO), whereas the P1022 mutation enhances TAOK2 kinase function (gain-of-function (GOF), analogous to an overexpression of TAOK2 (Fig. 4f).

Finally, we assessed the impact of the mutations in patient-derived LCLs, which endogenously express the mutation. We found that the proband harboring the A135P mutation had significantly reduced pTAOK2 and TAOK2 protein levels compared with LCLs from their unaffected father, which is consistent with our HEK293 cell data (Figs. 4g, h). Consistently, the P1022* mutation caused no change on pTAOK2, but no difference in TAOK2 expression was observed, possibly due to the heterozygosity of the patient (Supplementary Fig. 15a, b). Our data indicate that patient-derived de novo mutations in TAOK2 significantly impact TAOK2 function by decreasing (LOF) or increasing (GOF) intrinsic kinase activity toward downstream targets (Fig. 4f).

We next analyzed whether the TAOK2 mutations impact cellular distribution in primary neurons. We expressed WT and mutated α and β isoforms of TAOK2 in cortical neurons and analysis of GFP co-transfected DIV14 cortical neurons showed that WT TAOK2 α is primarily localized to the



dendritic shaft (co-stained with tubulin), but not present in actin-rich dendritic protrusions (co-stained with rhodamine-labeled phalloidin) (Supplementary Fig. 14e). Additionally,

the A135P mutation does not affect TAOK2 α localization (Supplementary Fig. 14f). On the other hand, we observed that WT TAOK2 β is expressed in dendrites with strong

◀ **Fig. 4** De novo mutations in TAOK2 impair phosphorylation at Ser181, localization in cortical neurons, and dendritic spine motility. **a** Diagram of TAOK2 α and TAOK2 β isoforms and location of *de novo* (A135P, P1022* and 563 + 12_563 + 15), truncating mutations (T604Sfs*45, Q622*, L1030fs*3), and rare-inherited variants (A335V and H781R). Different protein domains are represented by colored boxes (kinase domain: red, MEK binding domain: black, regulatory domains: blue (α) and light blue (β)). TAOK2 α has two phosphorylation sites (ser181 and thr475) and caspase-9 cleavage site (⁹¹⁶DPGD⁹¹⁹). TAOK2 β has three known phosphorylation sites (ser181, thr475, and ser1031). **b** Western blot of HEK293 cell lysates 48 h post transfection with TAOK2 β and β variants (A135P, P1022*, A335V, and H781R). **c** TAOK2 β A135P shows reduced protein expression and ser181 phosphorylation, TAOK2 β P1022* shows only reduced expression, and TAOK2 β A335V and H781R have no effect compared with TAOK2 β ($n = 6-8$ western blots; one-sample *t*-test; TAOK2 levels: A135P $t(5) = 5.303$, $p = 0.0032$; P1022* $t(5) = 8.903$, $p = 0.0003$; A335V $t(5) = 1.342$, $p = 0.2373$, H781R* $t(5) = 0.6381$, $p = 0.5515$; pTAOK2 levels: A135P $t(7) = 10.93$, $p < 0.0001$; P1022* $t(7) = 0.4992$, $p = 0.6330$; A335V $t(7) = 0.7243$, $p = 0.4924$; H781R $t(7) = 0.3526$, $p = 0.7348$). **d** Western blot of HEK293 cell lysates 48 h post transfection with JNK1a1 only or with TAOK2 β , TAOK2 β A135P and TAOK2 β P1022*. **e** TAOK2 β P1022* significantly increases phosphorylation of JNK1a1 in HEK293 cells compared with TAOK2 β ($n = 7$ western blots; one-way ANOVA, post hoc Dunnett's test; $F_{2, 18} = 6.88$, $p = 0.0060$; TAOK2 β vs. P1022* $p = 0.0342$, TAOK2 β vs. A135P $p = 0.5087$; JNK1a1 only (set to 100%) vs. TAOK2 β ; one-sample *t*-test, $t(6) = 3.167$, $p = 0.0194$). **f** Schematic showing impairment of ser181 auto-phosphorylation by the A135P mutation resulting in reduced kinase activity on downstream targets, whereas the P1022* mutation causes increased kinase activity of TAOK2. **g** Western blot of LCLs from the A135P proband and the unaffected father. **h** The A135P proband has reduced TAOK2 and pTAOK2 levels compared with the unaffected father ($n = 5$ western blots; one-sample *t*-test; TAOK2: $t(4) = 4.557$, $p = 0.0104$; pTAOK2: $t(4) = 2.74$, $p = 0.0519$). **i** Images of DIV14 cortical neuron cultures transfected with only GFP (control) or TAOK2 β , TAOK2 β A135P, and TAOK2 β P1022* with GFP and immunostained against GFP (green), TAOK2 (red) and stained with DAPI (blue). Scale bars represent 10 μ m. Boxes are shown magnified (right), with scale bars representing 3 μ m. Arrowheads represent dendritic spines filled with exogenous TAOK2 β , but not TAOK2 β A135P and P1022*. **j** Percentage of total exogenous TAOK2 β A135P and P1022* is increased in the soma (top) and decreased in the dendrite and spines (bottom) compared with TAOK2 β (TAOK2 β = 47, TAOK2 β A135P = 35 and TAOK2 β P1022* = 40 neurons from three separate cultures; one-way ANOVA post hoc Dunnett's test; Soma $F_{2, 119} = 28.01$, $p < 0.0001$: TAOK2 β vs. A135P $p < 0.0001$, TAOK2 β vs. P1022* $p < 0.0001$; Dendrite: $F_{2, 119} = 28.01$: TAOK2 β vs. A135P $p < 0.0001$, TAOK2 β vs. P1022* $p < 0.0001$). **k** Snapshots of DIV14 cortical neurons transfected with TAOK2 β and β variants (A135P and P1022*) at 0 s, 124 s and 302 s. Green arrowheads indicate extending filopodia spines and red arrowhead indicates retraction of a filopodia spine. **l** Left: increased spine motility in neurons transfected with *Taok2* shRNA compared with neurons transfected with control shRNA (Control shRNA = 12 and *Taok2* shRNA = 18 neurons from three different cultures; one-sample *t*-test; $t(17) = 3.799$, $p = 0.0014$). Right: TAOK2 β A135P transfected neurons have increased spine motility compared TAOK2 β transfected neurons (Control = 10, TAOK2 β = 22, A135P = 12 and P1022* = 14 neurons from three different cultures; one-way ANOVA, post hoc Dunnett's test; $F_{3, 42} = 17.2$, $p < 0.0001$; TAOK2 β vs. A135P $p = 0.0002$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ and **** $p < 0.0001$. Values are mean \pm s.e.m.

accumulation in filopodia/spine protrusions (Fig. 4i). We found that both mutations, A135P and P1022*, significantly increased the proportion of exogenous TAOK2 in the soma while concurrently reducing exogenous TAOK2 β in dendrites, specifically in filopodia/spines (Figs. 4i, j). These results suggest that the mutations impact protein trafficking of TAOK2 β in dendrites and spines.

Next, we measured dendritic spine motility at DIV14 *in vitro*. As a measure of spine dynamics, we analyzed the distance moved by the center of mass (CoM) of individual spines over 5 min [65]. We first tested this system by analyzing neurons transfected with either control or TAOK2 shRNA and found that acute silencing of TAOK2 caused a significant increase in filopodia/spine motility [48] (Fig. 4l, Supplementary Fig. 11h, i and Supplementary Video 1). Interestingly, we observed a comparable increase in filopodia/spine motility when the TAOK2 β A135P mutation was expressed in neurons, whereas the TAOK2 β P1022* mutation shows no significant effect or a trend toward decreased motility (Figs. 4k, l and Supplementary Video 2). The lack of change in spine motility when overexpressing TAOK2 β or an overactive TAOK2 β (P1022* mutant) indicates that A135P is the dominant LOF mutation that alters filopodia motility resembling the *Taok2* KO condition.

De novo mutations in TAOK2 impair dendrite and synaptic development

We next examined the effect of the de novo mutations on neurons *in vivo*, using *in utero* electroporation at E15 to transfect neural progenitor cells that produce layer 2/3 somatosensory cortical excitatory neurons. Our analysis of P21 WT and *Taok2* Het mouse cortices revealed a significant reduction in dendritic complexity with the loss of one *Taok2* allele (Figs. 5a, b). Introduction of WT TAOK2 α/β in the *Taok2* Het background using *in utero* electroporation significantly rescued the reduction in dendritic arborization; however, it did not completely match WT mice (Figs. 5a, b). To examine the de novo mutations, we expressed either Venus-GFP alone or together with WT TAOK2 α/β , TAOK2 α/β A135P (the patient possesses the mutation in both isoforms), or TAOK2 β P1022* (the patient only expresses the mutation in the β -isoform) in *Taok2* Het mice. We found that expression of the TAOK2 α/β A135P mutation did not rescue the dendrite phenotype of *Taok2* Het mice and revealed reduced branching (Figs. 5a, b) indicating potential dominant-negative effects. Interestingly, the TAOK2 β P1022* mutation significantly enhanced distal dendrite branching (Figs. 5a, b) with elongated dendrites (Supplementary Fig. 16a) compared with WT TAOK2 α/β . Although the P1022* mutation reduces the

Table 1 Diagnostic score summary of individuals diagnosed with ASD that have de novo or truncating variants in *TAOK2*

Proband	Variant	Type	RefSeq effect	Affected isoforms	Inheritance	Polyphen score	Diagnosis	IQ	Language	Adaptive behavior	Other medical info
1-0337-003	<i>TAO2</i> A135P	HET	MISSENSE	Alpha Beta Gamma	De novo	0.97	ADI dx: autism ADOS dx: not autism spectrum	Letter-R: 123	OWLS: listening comprehension SS = 81, oral expression SS = 77, oral composite SS = 77	VABS-II: communication SS = 79, daily living SS = 75 socialization SS = 72, adaptive behavior SS = 74	
1-0559-003	<i>TAO2</i> P1022*	HET	FRAMESHIFT DELETION	Beta	De novo	N/A	ADI and ADOS dx: autism; SCQ dx: autism	WASI-II: verbal comprehension composite = 55, perceptual reasoning composite = 71	OWLS: listening comprehension SS = 40, oral expression SS = 37, socialization SS = 61, adaptive behavior SS = 64;	VABS-II: communication SS = 70, daily living SS = 66, socialization SS = 61, adaptive behavior SS = 64;	CHD8 mutation (T2050Nfs*17)
7-0179-003	<i>TAO2</i> c.563 + 12_563 +15	HET	SPLICE SITE DELETION	Alpha Beta Gamma	De novo	N/A	SCQ dx: autism			ABAS-II: GAC composite = 40, practical composite = 40, social composite = 54, conceptual composite = 49	ADHD-Inattentive (SWAN)
1-0446-003	<i>TAO2</i> L1050Wfs*3	HET	FRAMESHIFT DELETION	Alpha Beta Gamma	N/A	N/A	ADOS dx: autism	WASI-II: verbal IQ = 117, performance IQ = 84, full scale IQ = 100	OWLS: listening comprehension SS = 96, oral expression SS = 91, oral composite SS = 92; PPVT-4: SS = 99	VABS-II: communication SS = 74, daily living SS = 66, socialization SS = 61, adaptive behavior SS = 65	

Table 1 (continued)

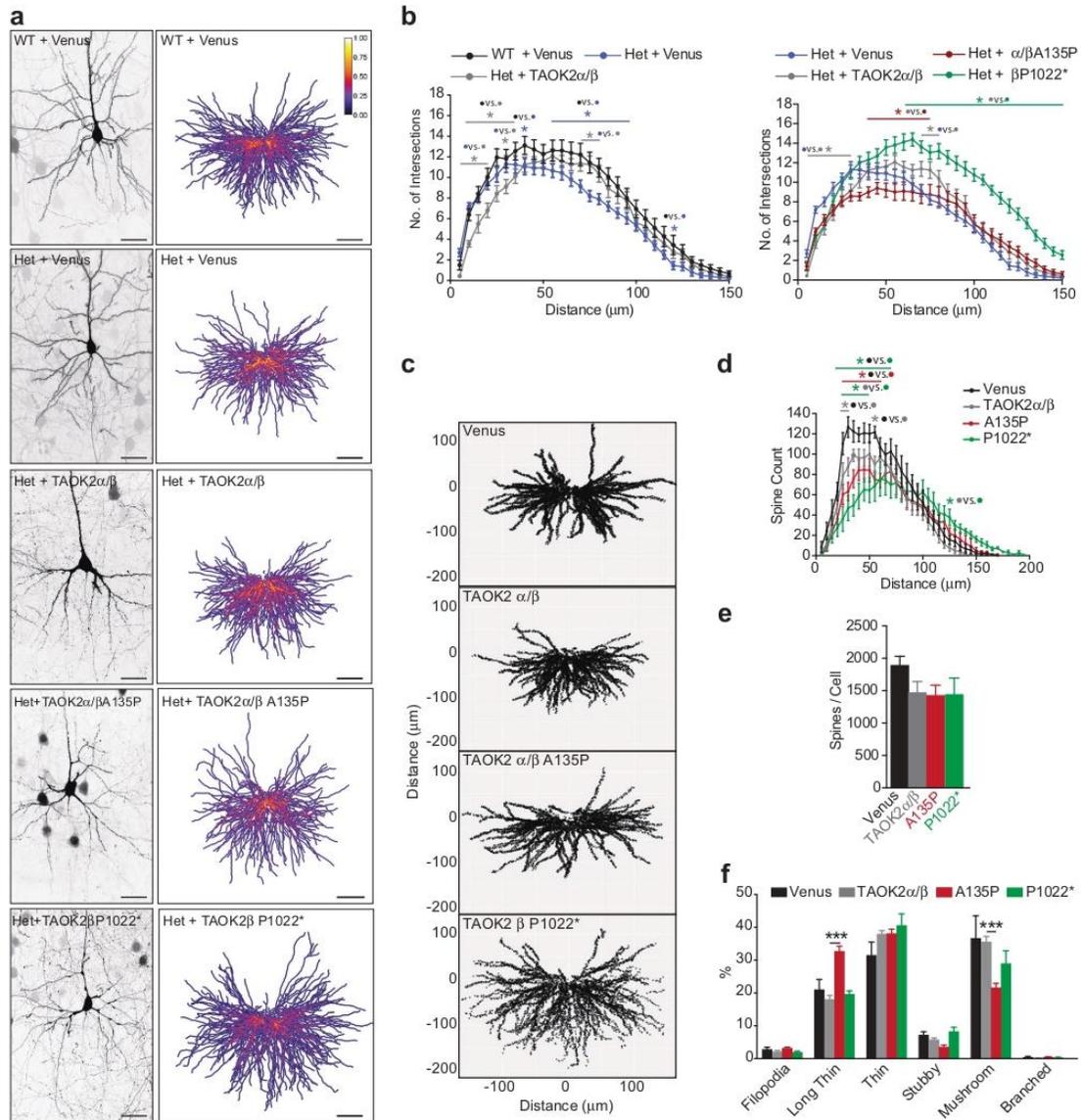
Proband	Variant	Type	RefSeq effect	Affected isoforms	Inheritance	Polyphen score	Diagnosis	IQ	Language	Adaptive behavior	Other medical info
AU4112303	TAO2 T6045fs*45	HET	NONSENSE	Alpha Beta Gamma	N/A	N/A	Clinical dx: autism; ADI dx: autism		PPVT-4: SS = 56	SRS-2 T-score (latest): social communication = 56, social motivation = 56, social awareness = 50, social cognition = 61, autistic RRb/ mannerisms = 58 total score = 57	
AU4261301	TAO2 Q622*	HET	FRAMESHIFT DELETION	Alpha Gamma	Paternal; Present in unaffected and affected sibling	N/A		N/A	N/A	N/A	N/A

level of TAOK2 expression, it also makes it highly active toward JNK1. This is consistent with our observation that the P1022* mutations elevates phosphorylation of pJNK1 (Figs. 4d, e) and previous work has shown that active JNK1 (pJNK1) increases cortical neuron dendrite branching [45]. We further examined dendrite integrity by measuring the thickness of the dendrite after a branching point. We found that while WT TAOK2 α/β increases branch thickness, the TAOK2 α/β A135P and TAOK2 β P1022* mutations had significantly reduced thickness (Supplementary Fig. 16b, c). These data demonstrate the complex effects of the de novo mutations on dendrite formation.

Finally, we analyzed the effects of the TAOK2 mutations on dendritic spine density and maturation. We found that TAOK2 α/β A135P and TAOK2 β P1022* both decreased proximal dendritic spines (Figs. 5c, d) without changing the total number of spines per neurons (Fig. 5e). Unexpectedly, TAOK2 α/β also decreased the number of proximal dendritic spines (Fig. 5d). However, as the 16p11.2 copy number variation (CNV) duplication is associated with schizophrenia, this suggests that elevated expression of TAOK2 may also be detrimental to neuron development. Regarding spine morphology, only overexpression of TAOK2 α/β A135P significantly increased the number of immature long thin spines while decreasing the number of mature, mushroom-shaped spines (Fig. 5f). This again indicates dominant negative effect (LOF) of the kinase dead TAOK2 α/β A135P mutation, whereas TAOK2 α/β and TAOK2 β P1022* mutations had no effect on dendritic spine morphology (Fig. 5f). Taken together, our results show that the human-derived mutations in TAOK2 differentially impair normal dendrite and spine development.

Taok2 regulates spine function through RhoA signaling

To understand how TAOK2 regulates spine formation beyond previously identified pathways, we asked if regulation of the actin cytoskeleton was involved given that it associates with actin-regulating proteins [47, 48]. Furthermore, the 16p11.2 CNV, which harbors *TAOK2*, was recently associated with abnormal RhoA signaling, a regulator of F-actin stability [11, 13, 80]. Therefore, we asked whether RhoA levels and activity are changed in *Taok2* KO brains. We isolated active (GTP-bound) RhoA from cortical and hippocampal KO brain lysates and observed a decrease in RhoA activity in cortical lysates compared with WT. Concurrently, we found an increase in total RhoA protein in the lysates of *Taok2* KO cortices (Figs. 6a, b). In the HC, however, the levels of activated RhoA were only marginally affected by the lack of *Taok2*. However, when we compared the overall levels of RhoA activity in cortices and hippocampi in *Taok2* WT mice, we observed a striking decline in



the amount of active RhoA GTPase in the HC (Figs. 6a, c), matching the low expression levels of phosphorylated Taok2 in the HC (Supplementary Fig. 9j, k). This suggests that *Taok2* may regulate the RhoA pathway predominantly in the cortex, but not in the HC.

To corroborate this data, we examined human LCLs from the A135P proband and also found significantly less activated RhoA and a significant increase in RhoA levels (Figs. 6d, e). Analysis of the LCLs from the P1022* family

revealed no changes in RhoA activity (Supplementary Fig. 15c, d), indicating again that the A135P and P1022* mutations have different effects, likely because they are localized in different functional domains within the TAOK2 protein.

Our data suggest that reduced RhoA activation may contribute to the dendritic spine phenotypes in the *Taok2* KO mice by reducing F-actin stability. We directly tested this by monitoring F-actin using Lifeact-GFP, which

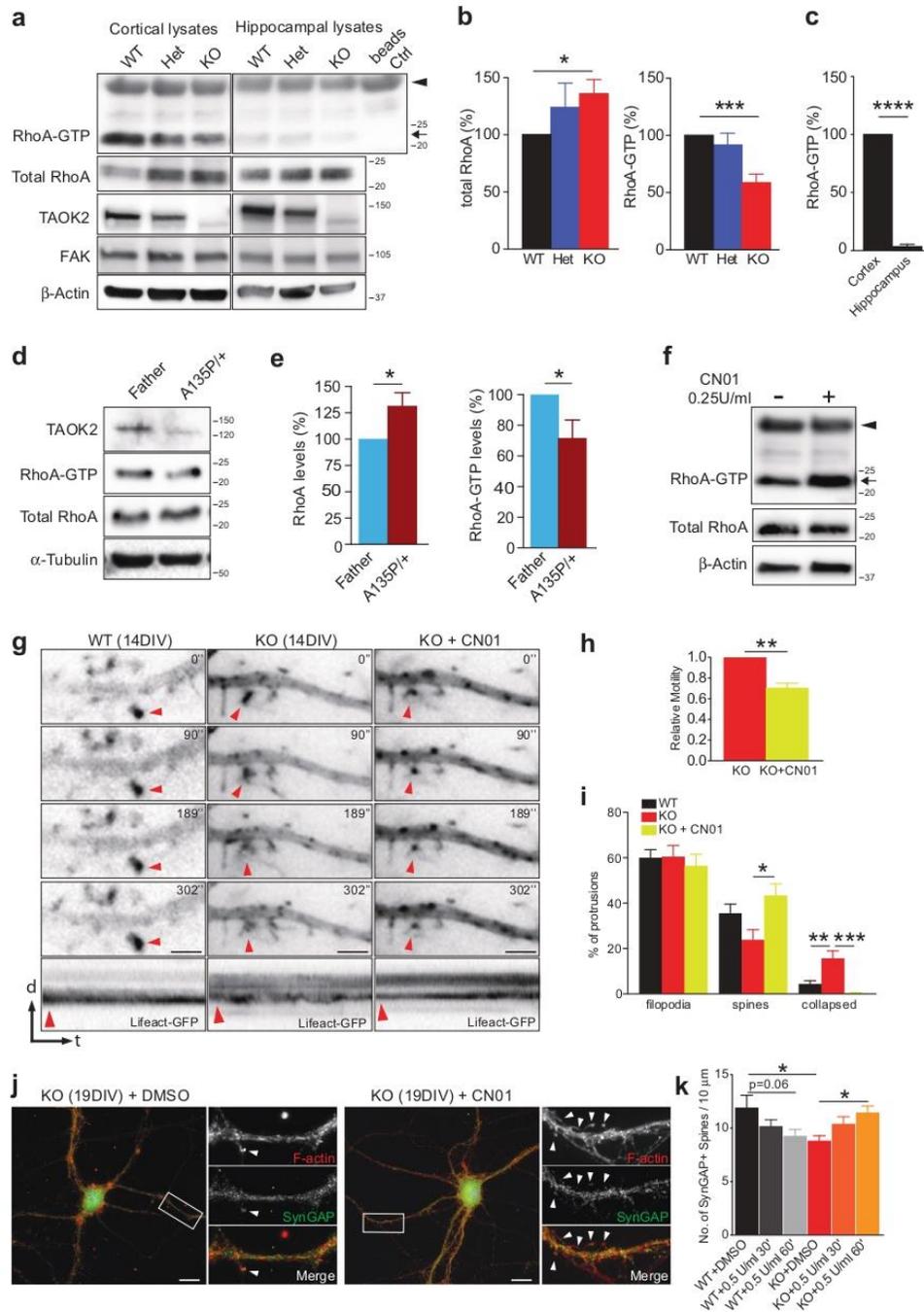
◀ **Fig. 5** De novo mutations in TAOK2 impair dendrite growth and synaptic connectivity in the mouse cortex. **a** Left: representative images of WT or Het *Taok2* cortical neurons from P21 mice in utero electroporated at E15 with GFP only (Venus, control) or Het *Taok2* cortical neurons with TAOK2 $\alpha\beta$, TAOK2 $\alpha\beta$ A135P, and TAOK2 β P1022*. Scale bars represent 20 μ m. Right: dendritic heat maps of superimposed neuron tracings for each condition. Blue to yellow indicates increased probability of dendrite presence. Scale bars represent 40 μ m. **(b)** Left: Het *Taok2* neurons show reduced dendritic complexity that can be partially rescued with overexpression of TAOK2 $\alpha\beta$ in the Het background (WT + Venus = 20 cells, Het + Venus = 18 cells, Het + TAOK2 $\alpha\beta$ = 16 cells from three different brains; two-way ANOVA, post hoc Tukey's test; $F_{2, 1989} = 42.48$, $p < 0.0001$ between genotypes; *represents ranges of significance; WT + Venus vs. Het + Venus (blue), WT + Venus vs. Het + TAOK2 $\alpha\beta$ (gray), Het + Venus vs. Het + TAOK2 $\alpha\beta$ (gray); see supplemental statistics). Right: TAOK2 $\alpha\beta$ A135P reduces proximal dendritic complexity and TAOK2 β P1022* enhances distal dendritic complexity compared with TAOK2 $\alpha\beta$ (Het + Venus = 20 cells, Het + TAOK2 $\alpha\beta$ = 16 cells, Het + TAOK2 $\alpha\beta$ A135P = 20 cells and Het + TAOK2 β P1022* = 16 cells from three different brains; two-way ANOVA, post hoc Dunnett's test; $F_{3, 2720} = 184.1$, $p < 0.0001$ between genotypes; *represents ranges of significance; TAOK2 $\alpha\beta$ vs. Venus (gray), TAOK2 $\alpha\beta$ vs. A135P (red), TAOK2 $\alpha\beta$ vs. P1022* (green); see supplemental statistics). **c** Dendritic spine distribution map of cortical neurons from WT *Taok2* cortical neurons from P21 mice in utero electroporated at E15 **d** TAOK2 $\alpha\beta$ A135P and TAOK2 β P1022* reduce the number of spines on proximal dendrites, whereas TAOK2 β P1022* also shifts spines to distal dendrites (Venus = 11414, TAOK2 $\alpha\beta$ = 8859, $\alpha\beta$ A135P = 9071 and β P1022* = 8686 spines from six neurons from three different brains per condition; two-way ANOVA, post hoc Tukey's test; $F_{3, 780} = 19.58$, $p < 0.0001$ between genotypes; *represents ranges of significance; Venus vs. TAOK2 $\alpha\beta$ (gray), TAOK2 $\alpha\beta$ vs. β P1022* (green), Venus vs. $\alpha\beta$ A135P (red), Venus vs. β P1022* (green); see supplemental statistics). **(e)** TAOK2 $\alpha\beta$ A135P and TAOK2 β P1022* show no significant change in dendritic spines per neuron (Venus = 11,414, TAOK2 $\alpha\beta$ = 8859, $\alpha\beta$ A135P = 9071 and β P1022* = 8686 spines from six neurons from three different brains per condition; spine number: one-way ANOVA, post hoc Dunnett's; $F_{3, 20} = 1.539$, $p = 0.2354$). **f** TAOK2 $\alpha\beta$ A135P significantly increases long thin spines and reduces mushroom-like spines in *Taok2* Het neurons (Venus = 11,414, TAOK2 $\alpha\beta$ = 8859, $\alpha\beta$ A135P = 9071 and β P1022* = 8686 spines from six neurons from three different brains per condition; two-way ANOVA, post hoc Dunnett's test; $F_{3, 120} = 4.36e-13$, $p > 0.9999$; Long Thin: TAOK2 $\alpha\beta$ vs. TAOK2 $\alpha\beta$ A135P $p = 0.0001$; Mushroom: TAOK2 $\alpha\beta$ vs. TAOK2 $\alpha\beta$ A135P $p = 0.0003$). * $p < 0.05$, *** $p < 0.001$. Values are mean \pm s.e.m.

is a 17-amino-acid peptide that labels F-actin with GFP without affecting its function [81, 82]. We measured the distance moved by the CoM of the Lifeact-GFP signal in individual spines [65]. We found that the F-actin signal was diffuse and unstable in filopodia/spine heads in KO neurons compared with LifeAct-GFP in WT neurons (Fig. 6g). Given this, we asked whether increasing RhoA activity in *Taok2* KO neurons could rescue the reduction in F-actin stability. We tested a commercial chemical activator (CN01) and found elevated levels of RhoA-GTP in the neuroblastoma cell line SHSY5Y (arrow, Fig. 6f). We used this compound on DIV14 *Taok2*

KO neurons and measured F-actin dynamics before and after addition. The RhoA activator significantly increased F-actin distribution toward accumulation in dendritic spine heads in *Taok2* KO cortical neuron cultures, concurrently decreasing the relative spine motility (Figs. 6g, h; relative motility normalized to Lifeact-GFP-transfected KO cells without treatment; Supplementary Video 3). The reduction in spine motility was accompanied by the elimination of spine collapse, concurrent with the formation of stable spines bearing a spine-head with actin accumulation in *Taok2* KO neurons (Fig. 6i). We then analyzed whether activating RhoA in *Taok2* KO cortical neurons could rescue the deficiency in dendritic spine density. Short-term incubation with the RhoA activator (30–60 min) increased the number of dendritic synapses (spines with SynGAP co-staining) in *Taok2* KO cells compared with dimethyl sulfoxide (DMSO) treated *Taok2* KO cells (Figs. 6j, k). Importantly, incubation of WT neurons with the RhoA activator did not show any significant increase, but did show a nonsignificant decrease in SynGAP-positive dendritic spines, highlighting an optimal range of RhoA activity is necessary for proper spine formation (Fig. 6k, WT + DMSO vs. WT + CN01(60') $p = 0.0611$). Given that increasing RhoA activity rescued the *Taok2*-dependent synaptic defects, we asked whether *Taok2* bound in a functional complex with RhoA. We immunoprecipitated *Taok2* from both a crude homogenate and a crude membrane fraction of WT mouse cortices to determine whether *Taok2* interacts with RhoA. Our results show that RhoA is in the same protein complex as *Taok2* (Supplementary Fig. 17a). Interestingly, immunoprecipitation of RhoA after overexpression of TAOK2 α and β isoforms and the de novo mutants in HEK293 cells, reveal that TAOK2 β preferentially binds RhoA compared with TAOK2 α (Supplementary Fig. 17b, c). Furthermore, the binding of TAOK2 β to RhoA is affected by the P1022* mutation, suggesting that the C-terminal domain, which differs between TAOK2 α and β , is important for RhoA binding (Supplementary Fig. 17b, c). Our results reveal that *Taok2* regulates the filopodia–spine transition through RhoA activity and the activation of RhoA is sufficient to overcome spine deficits in *Taok2* KO cultures.

Discussion

Previous studies have implicated *TAOK2* in NDDs but there has been no comprehensive study on *TAOK2* to support this. Our data provide novel evidence that *TAOK2* is directly associated with ASD pathologies, such as deficits in social interaction, enlarged brain volume, and reduced dendritic growth and spine formation, and identifies a mechanistic pathway regulating synapse function. Our



results show that Taok2 regulates RhoA activation, and loss of Taok2 leads to abnormal actin dynamics, stability and organization that may contribute to the synaptic defects

observed in *Taok2* KO mice. Our data also provide new insight into the spatial role of Taok2 in the mouse brain. Loss of Taok2 has the strongest effect on excitatory neurons

Fig. 6 TAOK2 regulates spine function through RhoA signaling. **a** Western blot of RhoA-GTP (arrow) and total RhoA in cortical and hippocampal lysates from P21 WT, Het, and KO *Taok2* mice. Arrowhead indicates background from Rhotekin-GST. **b** Increased levels of total RhoA and reduced levels of Rho-GTP (normalized to β -actin) in the *Taok2* KO mice cortex. **c** Reduced levels of RhoA-GTP (normalized to β -actin) in the hippocampus compared with the cortex in WT mice. ($n = 4$ western blots; one-sample *t*-test; Hippocampus: $t(3) = 46.03, p < 0.0001$). (Eight separate mouse cortices per condition; total RhoA: one-sample *t*-test; WT vs. KO: $t(7) = 2.905, p = 0.0228$; RhoA-GTP: one-sample *t*-test; WT vs. KO: $t(7) = 5.55, p = 0.0009$). **d** Western blot of RhoA-GTP and total RhoA in LCLs of the A135P proband and the unaffected father. **e** Increased levels of total RhoA and reduced levels of Rho-GTP in LCLs of the A135P proband compared with the unaffected father (nine separate lysates per LCL; RhoA levels (normalized to tubulin): one-sample *t*-test; $t(80) = 2.499, p = 0.0370$; RhoA-GTP: one-sample *t*-test; $t(8) = 2.399, p = 0.0433$). **f** Western blots of the SHSY-5Y neuroblastoma cell lines showing increased RhoA-GTP levels after addition of the RhoA activator (CN01) (arrow). Arrowhead indicates background from Rhotekin-GST. **g** Snapshots from time-lapse analysis (0, 90, 189, and 302 s) of dendritic filopodia/spines from DIV14 cortical neurons labeled with Lifeact-GFP from WT and KO *Taok2* mice and *Taok2* KO treated with CN01 (1 U/ml). Red arrowheads indicate actin-rich protrusions selected for kymograph analysis. Bottom: kymographs reveal diffuse and unstable spine movement in *Taok2* KO neurons that is rescued by CN01. Scale bars represent 3 μ m. **h** CN01 reduces spine motility of *Taok2* KO neurons during a 5-min period (KO = 105, KO + CN01 = 109 spines from 9 cells per condition; Wilcoxon-matched paired *t*-test, $p = 0.0039$). **i** CN01 increased percentage of stable spines and reduced the percentage of collapsed spines in *Taok2* KO neurons during a 5-min period (WT = 418 spines from 8 cells, KO = 603 spines from 9 cells and KO + CN01 = 433 spines from 7 cells; Filopodia: one-way ANOVA, post hoc Tukey's test; $F_{2,21} = 0.2249, p = 0.8005$; Spines: one-way ANOVA, post hoc Tukey's test; $F_{2,21} = 4.66, p = 0.0211$; WT vs. KO + CN01 $p = 0.4905$, KO vs. KO + CN01 $p = 0.0180$; Collapsed: one-way ANOVA, post hoc Tukey's test; $F_{2,21} = 12.5, p = 0.0003$; WT vs. KO $p = 0.0047$, WT vs. KO + CN01 $p = 0.4473$, KO vs. KO + CN01 $p = 0.0003$). **j** Images of DIV18-19 *Taok2* KO neurons with or without addition of CN01 for 30–60 min, stained with phalloidin and SynGAP. Scale bars represent 10 μ m. **k** CN01 increases the number of SynGAP + spines in *Taok2* KO neurons compared with DMSO control KO neurons (WT + DMSO = 6 cells, WT + CN01 (0.5U/ml/30') = 8 cells, WT + CN01(0.5U/ml/60') = 8 cells, KO + DMSO = 9 cells, KO + CN01 (0.5U/ml/30') = 9 cells; KO + CN01 (0.5U/ml/60') = 9 cells; one-way ANOVA, post hoc Bonferroni's test; $F_{5,43} = 3.192, p = 0.0154$; WT + DMSO vs. KO + DMSO $p = 0.0156$, WT + DMSO vs. WT + CN01(0.5U/ml/60') $p = 0.0611$, KO + DMSO vs. KO + CN01(0.5U/ml/60') $p = 0.0235$). * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$. Values are mean \pm s.e.m.

in the PFC and, by a smaller degree, in the somatosensory cortex, but shows no effect in the HC. This is consistent with a recent publication showing that knockdown of *Taok2* in hippocampal neurons reveals no electrophysiological phenotype [48].

One of our main findings is that alterations in TAOK2 activity contribute to NDDs. First, *Taok2* KO mice have several behavioral, anatomical, and synaptic deficits consistent with other ASD mouse models [49–53]. Second, we identified and characterized novel de novo mutations in *TAOK2* in human ASD cohorts revealing that the mutations

impact different signaling pathways (Supplementary Fig. 18a). Our results show that the A135P mutation reduces TAOK2 activity indicating a LOF mutation, and causes a reduction in JNK1 activation, decreased RhoA activity and reduced dendrite growth and spine maturation. Whereas the P1022* mutation is a GOF mutation, which increases TAOK2-dependent activation of JNK1, thus enhancing dendritic growth and branching. Functional differences between the mutations are expected given that auto-phosphorylation of the kinase domain is important for TAOK2 activation, whereas the C-terminus negatively regulates kinase activity [78, 79]. Therefore, either a reduction or elevation in *TAOK2* gene dosage could be pathogenic, consistent with deletions or duplications of the 16p11.2 locus, which harbors *TAOK2* and confers a risk for ASD and schizophrenia, respectively [8, 14]. This is also consistent with analysis of cortical neurons from a 16p11.2 duplication mouse that displayed excessive dendrite outgrowth, similar to the *TAOK2* P1022* mutation [83]. Although the individual with the *TAOK2* P1022* mutation also has a *CHD8* mutation, they have a more complex phenotype when comparing scores for adaptive behavior and IQ compared with the other subjects with *TAOK2* mutations, suggesting that both *TAOK2* P1022* and *CHD8* may contribute to this subject's phenotype.

Additional support for the contribution of *TAOK2* to the development of neuropathologies comes from a recent report identifying two de novo mutations in *TAOK2* in individuals with a complex developmental (and neurological) phenotype [84]. This may be relevant for 16p11.2 deletion/duplication carriers who do not develop ASD/schizophrenia, but the Het deletion/duplication of *TAOK2* may be sufficient for a NDD in these individuals. Importantly, the *TAOK2* de novo mutations we characterized are not found in the ExAC database, indicating these mutations are more likely pathogenic. *TAOK2* is also highly intolerant to LOF and missense mutations and it is known that genes within CNVs have a different mechanism of mutation compared with genes outside of CNVs [85]. Therefore, the identification of three de novo mutations and three other truncating mutations in the *TAOK2* gene is highly suggestive of its significance to NDDs including ASD.

Given the results of our study and due to *TAOK2* being localized in the 16p11.2 CNV, it is conceivable that it may contribute to disease pathophysiology. For example, we found an increase in midbrain volume in *Taok2* KO mice, similar to the 16p11.2/+ mice, however, the opposite effect is observed for total brain volume [86–88]. Interestingly, patients with the 16p11.2del CNV show measurable increases in head circumference, gray matter, and white matter in the cortex and thalamus [89–92], which coincides with the increase in total brain volume and in the thalamus and midbrain volumes in *Taok2* Het and KO mice, but not

with the 16p11.2del mice. The somatosensory cortex region has not been analyzed in 16p11.2del mice, however, a volumetric decrease in the general frontal lobe has been observed [88]. Furthermore, studies have shown 16p11.2del CNV patients have increases in the corpus callosum volume, which is dissimilar to the changes seen in the *Taok2* KO mice, indicating TAOK2 may only affect specific cortical structures in individuals with the 16p11.2 CNV [93, 94]. Interestingly, although KCTD13 is proposed to be a driver gene of head size in the 16p11.2 CNV [10], a recent study using *Kctd13* KO mice revealed no brain size changes [13], suggesting instead that TAOK2 along with other genes may regulate this phenotype. KCTD13 also regulates RhoA signaling in the brain, suggesting that elevated RhoA levels in the 16p11.2 deletion are pathogenic and caused by KCTD13 [11]. In comparison with TAOK2, cortical lysates from *Taok2* KO mouse brains and patient-derived LCLs exhibit decreased RhoA activity, indicating that TAOK2 and KCTD13 may regulate RhoA in different brain regions or at different developmental time points. Interestingly, as we found that TAOK2 binds in a functional complex with RhoA, and RhoA is a known substrate of KCTD13, this raises the possibility that TAOK2 and KCTD13 may regulate one another. Importantly, additional genes in the 16p11.2 CNV are also thought to contribute to the disease phenotype, including *MAPK3* and *SEZ6L2* [6, 7, 10, 11, 83]. Given the large 16p11.2 CNV region and the complexity of neurological phenotypes associated with it, it is likely that multiple genes play a role, similar to other CNVs [95].

In summary, we have characterized *Taok2* Het and KO mice using behavioral assays for ASD-associated phenotypes, MRI to identify gross brain morphological changes and functional studies that identified cellular deficits causing altered neural morphology, connectivity, and activity through a RhoA-dependent pathway (Supplementary Fig. 18b). In addition, we identified human-derived de novo LOF/GOF mutations and studied their impact using human and mouse in vitro and in vivo systems (Supplementary Fig. 18b). Our study defines *TAOK2* as a novel NDD risk gene and provides novel data that demonstrate how patient-derived mutations impact brain function and development. Given this, it is important to determine if other ASD or NDD cohorts possess pathogenic variants or mutations in *TAOK2* to better understand its contribution to disease.

Acknowledgements FCdA is supported by Deutsche Forschungsgemeinschaft (DFG) grant (FOR 2419; CA1495/1-1), ERANET Neuron Grant (Bundesministerium für Bildung und Forschung, BMBF, 01EW1410 ZMNH AN B1), Landesforschungsförderung Hamburg (Z-AN LF), and University Medical Center Hamburg-Eppendorf (UKE). FM is supported by DFG grant (SFB 936/project B7), Landesforschungsförderung Hamburg (Z-MO LF) and UKE. KKS is supported by grants from CIHR-ERA/NET, NSERC and

the Ontario Brain Institute (OBI). SWS is supported by grants from CIHR, Autism Speaks and OBI. MH is supported by The Egyptian Cultural Bureau, Berlin.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ling P, Lu TJ, Yuan CJ, Lai MD. Biosignaling of mammalian Ste20-related kinases. *Cell Signal*. 2008;20:1237–47.
- Record CJ, Chaikuad A, Rellos P, Das S, Pike AC, Fedorov O, et al. Structural comparison of human mammalian ste20-like kinases. *PLoS ONE*. 2010;5:e11905.
- Yusteín JT, Xia L, Kahlenburg JM, Robinson D, Templeton D, Kung HJ. Comparative studies of a new subfamily of human Ste20-like kinases: homodimerization, subcellular localization, and selective activation of MKK3 and p38. *Oncogene*. 2003;22:6129–41.
- Dan I, Watanabe NM, Kusumi A. The Ste20 group kinases as regulators of MAP kinase cascades. *Trends Cell Biol*. 2001;11:220–30.
- Barber DL. Chopped and diced: Dicer1 deletion generates myeloid dysplasia. *Blood*. 2012;119:4581–2.
- Konyukh M, Delorme R, Chaste P, Leblond C, Lemiere N, Nygren G, et al. Variations of the candidate SEZ6L2 gene on chromosome 16p11.2 in patients with autism spectrum disorders and in human populations. *PLoS ONE*. 2011;6:e17289.
- Pucilowska J, Vithayathil J, Tavares EJ, Kelly C, Karlo JC, Landreth GE. The 16p11.2 deletion mouse model of autism exhibits altered cortical progenitor proliferation and brain cytoarchitecture linked to the ERK MAPK pathway. *J Neurosci*. 2015;35:3190–200.
- Weiss LA, Shen Y, Kom JM, Arking DE, Miller DT, Fossdal R, et al. Association between microdeletion and microduplication at 16p11.2 and autism. *N Engl J Med*. 2008;358:667–75.
- Steinman KJ, Spence SJ, Ramocki MB, Proud MB, Kessler SK, Marco EJ, et al. 16p11.2 deletion and duplication: characterizing neurologic phenotypes in a large clinically ascertained cohort. *Am J Med Genet Part A*. 2016;170:2943–55.
- Golzio C, Willer J, Talkowski ME, Oh EC, Taniguchi Y, Jacquemont S, et al. KCTD13 is a major driver of mirrored neuroanatomical phenotypes of the 16p11.2 copy number variant. *Nature*. 2012;485:363–7.
- Lin GN, Corominas R, Lemmens I, Yang X, Tavernier J, Hill DE, et al. Spatiotemporal 16p11.2 protein network implicates cortical late mid-fetal brain development and KCTD13-

- Cul3-RhoA pathway in psychiatric diseases. *Neuron*. 2015;85:742–54.
12. Filges I, Sparagana S, Sargent M, Selby K, Schlade-Bartusiak K, Lueder GT, et al. Brain MRI abnormalities and spectrum of neurological and clinical findings in three patients with proximal 16p11.2 microduplication. *Am J Med Genet Part A*. 2014;164A:2003–12.
 13. Escamilla CO, Filonova I, Walker AK, Xuan ZX, Holehonnur R, Espinosa F, et al. Kctd13 deletion reduces synaptic transmission via increased RhoA. *Nature*. 2017;551:227–31.
 14. Steinberg S, de Jong S, Mattheisen M, Costas J, Demontis D, Jamain S, et al. Common variant at 16p11.2 conferring risk of psychosis. *Mol Psychiatry*. 2014;19:108–14.
 15. Damell JC, Van Driesche SJ, Zhang C, Hung KY, Mele A, Fraser CE, et al. FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell*. 2011;146:247–61.
 16. Sanders SJ, Murtha MT, Gupta AR, Murdoch JD, Raubeson MJ, Willsey AJ, et al. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature*. 2012;485:237–41.
 17. O’Roak BJ, Vives L, Girirajan S, Karakoc E, Krumm N, Coe BP, et al. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature*. 2012;485:246–50.
 18. Jiang YH, Yuen RK, Jin X, Wang M, Chen N, Wu X, et al. Detection of clinically relevant genetic variants in autism spectrum disorder by whole-genome sequencing. *Am J Human Genet*. 2013;93:249–63.
 19. Sanders SJ, He X, Willsey AJ, Ercan-Sencicek AG, Samocha KE, Ciccek AE, et al. Insights into autism spectrum disorder genomic architecture and biology from 71 risk loci. *Neuron*. 2015;87:1215–33.
 20. Kosmicki JA, Samocha KE, Howrigan DP, Sanders SJ, Slowikowski K, Lek M, et al. Refining the role of de novo protein-truncating variants in neurodevelopmental disorders by using population reference samples. *Nat Genet*. 2017;49:504–10.
 21. Stessman HA, Xiong B, Coe BP, Wang T, Hoekzema K, Fenckova M, et al. Targeted sequencing identifies 91 neurodevelopmental-disorder risk genes with autism and developmental-disability biases. *Nat Genet*. 2017;49:515–26.
 22. Iossifov I, O’Roak BJ, Sanders SJ, Ronemus M, Krumm N, Levy D, et al. The contribution of de novo coding mutations to autism spectrum disorder. *Nature*. 2014;515:216–21.
 23. O’Roak BJ, Vives L, Fu W, Egerton JD, Stanaway IB, Phelps IG, et al. Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science*. 2012;338:1619–22.
 24. Jamain S, Quach H, Betancur C, Rastam M, Colinaux C, Gillberg IC, et al. Mutations of the X-linked genes encoding neuro-ligins NLGN3 and NLGN4 are associated with autism. *Nat Genet*. 2003;34:27–9.
 25. Etherton MR, Blaiss CA, Powell CM, Sudhof TC. Mouse neurexin-1 alpha deletion causes correlated electrophysiological and behavioral changes consistent with cognitive impairments. *Proc Natl Acad Sci USA*. 2009;106:17998–8003.
 26. Berg JM, Lee C, Chen L, Galvan L, Cepeda C, Chen JY, et al. JAKMIP1, a novel regulator of neuronal translation, modulates synaptic function and autistic-like behaviors in mouse. *Neuron*. 2015;88:1173–91.
 27. Asceno M Jr, Mukherjee N, Bandaru P, Miller JB, Nusbaum JD, Corcoran DL, et al. FMRP targets distinct mRNA sequence elements to regulate protein expression. *Nature*. 2012;492:382–6.
 28. Oguro-Ando A, Rosensweig C, Heman E, Nishimura Y, Werling D, Bill BR, et al. Increased CYFIP1 dosage alters cellular and dendritic morphology and dysregulates mTOR. *Mol Psychiatry*. 2015;20:1069–78.
 29. Smith SE, Zhou YD, Zhang G, Jin Z, Stoppel DC, Anderson MP. Increased gene dosage of Ube3a results in autism traits and decreased glutamate synaptic transmission in mice. *Sci Transl Med*. 2011;3:103ra97.
 30. Kwon CH, Luikart BW, Powell CM, Zhou J, Matheny SA, Zhang W, et al. Pten regulates neuronal arborization and social interaction in mice. *Neuron*. 2006;50:377–88.
 31. Bader PL, Faizi M, Kim LH, Owen SF, Tadross MR, Alfa RW, et al. Mouse model of Timothy syndrome recapitulates triad of autistic traits. *Proc Natl Acad Sci USA*. 2011;108:15432–7.
 32. Durand CM, Betancur C, Boeckers TM, Bockmann J, Chaste P, Fauchereau F, et al. Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders. *Nat Genet*. 2007;39:25–7.
 33. Peca J, Feliciano C, Ting JT, Wang W, Wells MF, Venkatraman TN, et al. Shank3 mutant mice display autistic-like behaviours and striatal dysfunction. *Nature*. 2011;472:437–42.
 34. Schmeisser MJ, Ey E, Wegener S, Bockmann J, Stempel AV, Kuebler A, et al. Autistic-like behaviours and hyperactivity in mice lacking ProSAP1/Shank2. *Nature*. 2012;486:256–60.
 35. Kwan V, Meka DP, White SH, Hung CL, Holzapfel NT, Walker S, et al. DIXDC1 phosphorylation and control of dendritic morphology are impaired by rare genetic variants. *Cell Rep*. 2016;17:1892–904.
 36. Clement JP, Aceti M, Creson TK, Ozkan ED, Shi Y, Reish NJ, et al. Pathogenic SYNGAP1 mutations impair cognitive development by disrupting maturation of dendritic spine synapses. *Cell*. 2012;151:709–23.
 37. Bidinosti M, Botta P, Kruttner S, Proenca CC, Stoehr N, Bernhard M, et al. CLK2 inhibition ameliorates autistic features associated with SHANK3 deficiency. *Science*. 2016;351:1199–203.
 38. Marchetto MC, Belinson H, Tian Y, Freitas BC, Fu C, Vadodaria KC, et al. Altered proliferation and networks in neural cells derived from idiopathic autistic individuals. *Mol Psychiatry*. 2016;22:820–35.
 39. Gkogkas CG, Khoutorsky A, Cao R, Jafamejad SM, Prager-Khoutorsky M, Giannakas N, et al. Pharmacogenetic inhibition of eIF4E-dependent Mmp9 mRNA translation reverses fragile X syndrome-like phenotypes. *Cell Rep*. 2014;9:1742–55.
 40. Aguilar-Valles A, Matta-Camacho E, Khoutorsky A, Gkogkas C, Nader K, Lacaille JC, et al. Inhibition of group I metabotropic glutamate receptors reverses autistic-like phenotypes caused by deficiency of the translation repressor eIF4E binding protein 2. *J Neurosci*. 2015;35:11125–32.
 41. Harony-Nicolas H, Kay M, Hoffmann JD, Klein ME, Bozdagi-Gunal O, Riad M, et al. Oxytocin improves behavioral and electrophysiological deficits in a novel Shank3-deficient rat. *eLife*. 2017;6::pii: e18904.
 42. Zihni C, Mitsopoulos C, Tavares IA, Baum B, Ridley AJ, Morris JD. Prostate-derived sterile 20-like kinase 1-alpha induces apoptosis. JNK- and caspase-dependent nuclear localization is a requirement for membrane blebbing. *J Biol Chem*. 2007;282:6484–93.
 43. Moore TM, Garg R, Johnson C, Coptcoat MJ, Ridley AJ, Morris JDPSK. A novel STE20-like kinase derived from prostatic carcinoma that activates the c-Jun N-terminal kinase mitogen-activated protein kinase pathway and regulates actin cytoskeletal organization. *J Biol Chem*. 2000;275:4311–22.
 44. Zhou T, Raman M, Gao Y, Eamest S, Chen Z, Machius M, et al. Crystal structure of the TAO2 kinase domain: activation and specificity of a Ste20p MAP3K. *Structure*. 2004;12:1891–900.
 45. de Anda FC, Rosario AL, Durak O, Tran T, Graff J, Meletis K, et al. Autism spectrum disorder susceptibility gene TAO2 affects basal dendrite formation in the neocortex. *Nat Neurosci*. 2012;15:1022–31.
 46. Yasuda S, Tanaka H, Sugiura H, Okamura K, Sakaguchi T, Tran U, et al. Activity-induced protocadherin arcadlin regulates dendritic spine number by triggering N-cadherin

- endocytosis via TAO2beta and p38 MAP kinases. *Neuron*. 2007;56:456–71.
47. Ultanir SK, Yadav S, Hertz NT, Oses-Prieto JA, Claxton S, Burlingame AL, et al. MST3 kinase phosphorylates TAO1/2 to enable Myosin Va function in promoting spine synapse development. *Neuron*. 2014;84:968–82.
 48. Yadav S, Oses-Prieto JA, Peters CJ, Zhou J, Pleasure SJ, Burlingame AL, et al. TAOK2 kinase mediates PSD95 stability and dendritic spine maturation through Septin7 phosphorylation. *Neuron*. 2017;93:379–93.
 49. Etherton M, Foldy C, Sharma M, Tabuchi K, Liu X, Shamloo M, et al. Autism-linked neuroligin-3 R451C mutation differentially alters hippocampal and cortical synaptic function. *Proc Natl Acad Sci USA*. 2011;108:13764–9.
 50. Jamain S, Radyushkin K, Hammerschmidt K, Granon S, Boretius S, Varoqueaux F, et al. Reduced social interaction and ultrasonic communication in a mouse model of monogenic heritable autism. *Proc Natl Acad Sci USA*. 2008;105:1710–5.
 51. Dachtler J, Ivorra JL, Rowland TE, Lever C, Rodgers RJ, Clapcote SJ. Heterozygous deletion of alpha-neurexin I or alpha-neurexin II results in behaviors relevant to autism and schizophrenia. *Behav Neurosci*. 2015;129:765–76.
 52. Wang X, McCoy PA, Rodriguiz RM, Pan Y, Je HS, Roberts AC, et al. Synaptic dysfunction and abnormal behaviors in mice lacking major isoforms of Shank3. *Human Mol Genet*. 2011;20:3093–108.
 53. Won H, Lee HR, Gee HY, Mah W, Kim JJ, Lee J, et al. Autistic-like social behaviour in Shank2-mutant mice improved by restoring NMDA receptor function. *Nature*. 2012;486:261–5.
 54. Kapfhamer D, Taylor S, Zou ME, Lim JP, Kharazia V, Heberlein U. Taok2 controls behavioral response to ethanol in mice. *Genes Brain Behav*. 2013;12:87–97.
 55. Sedlacik J, Boelmans K, Lobel U, Holst B, Siemonsen S, Fiehler J. Reversible, irreversible and effective transverse relaxation rates in normal aging brain at 3T. *NeuroImage*. 2014;84:1032–41.
 56. Sedlacik J, Reitz M, Bolar DS, Adalsteinsson E, Schmidt NO, Fiehler J. Correlation of oxygenation and perfusion sensitive MRI with invasive micro probe measurements in healthy mice brain. *Z fur Med Phys*. 2015;25:77–85.
 57. Siemonsen S, Lobel U, Sedlacik J, Forkert ND, Mouridsen K, Ostergaard L, et al. Elevated T2-values in MRI of stroke patients shortly after symptom onset do not predict irreversible tissue infarction. *Brain J Neurol*. 2012 Jun;135(Pt 6):1981–9.
 58. Hartung H, Cichon N, De Feo V, Riemann S, Schildt S, Lindemann C, et al. From shortage to surge: a developmental switch in hippocampal-prefrontal coupling in a gene-environment model of neuropsychiatric disorders. *Cereb Cortex*. 2016;26:4265–81.
 59. Freitag S, Schachner M, Morellini F. Behavioral alterations in mice deficient for the extracellular matrix glycoprotein tenascin-R. *Behav Brain Res*. 2003;145:189–207.
 60. Fellini L, Morellini F. Mice create what-where-when hippocampus-dependent memories of unique experiences. *J Neurosci*. 2013;33:1038–43.
 61. Marguet SL, Le-Schulte VT, Merseburg A, Neu A, Eichler R, Jakovcevski I, et al. Treatment during a vulnerable developmental period rescues a genetic epilepsy. *Nat Med*. 2015;21:1436–44.
 62. Morellini F. Spatial memory tasks in rodents: what do they model? *Cell Tissue Res*. 2013;354:273–86.
 63. Koster JD, Leggewie B, Blechner C, Brandt N, Fester L, Rune G, et al. Inositol-1,4,5-trisphosphate-3-kinase-A controls morphology of hippocampal dendritic spines. *Cell Signal*. 2016;28:83–90.
 64. Risher WC, Ustunkaya T, Singh Alvarado J, Eroglu C. Rapid Golgi analysis method for efficient and unbiased classification of dendritic spines. *PLoS ONE*. 2014;9:e107591.
 65. Tamok K, Gulyas M, Bencsik N, Ferenc K, Pfizenmaier K, Hausser A, et al. A new tool for the quantitative analysis of dendritic filopodial motility. *Cytom Part A J Int Soc Anal Cytol*. 2015;87:89–96.
 66. Yuen RK, Thiruvahindrapuram B, Merico D, Walker S, Tammiemies K, Hoang N, et al. Whole-genome sequencing of quartet families with autism spectrum disorder. *Nat Med*. 2015;21:185–91.
 67. Adhikari A, Topiwala MA, Gordon JA. Single units in the medial prefrontal cortex with anxiety-related firing patterns are preferentially influenced by ventral hippocampal activity. *Neuron*. 2011;71:898–910.
 68. Spellman T, Rigotti M, Ahmari SE, Fusi S, Gogos JA, Gordon JA. Hippocampal-prefrontal input supports spatial encoding in working memory. *Nature*. 2015;522:309–14.
 69. Rajasethupathy P, Sankaran S, Marshel JH, Kim CK, Ferenczi E, Lee SY, et al. Projections from neocortex mediate top-down control of memory retrieval. *Nature*. 2015;526:653–9.
 70. Navawongse R, Eichenbaum H. Distinct pathways for rule-based retrieval and spatial mapping of memory representations in hippocampal neurons. *J Neurosci*. 2013;33:1002–13.
 71. Brockmann MD, Poschel B, Cichon N, Hanganu-Opatz IL. Coupled oscillations mediate directed interactions between prefrontal cortex and hippocampus of the neonatal rat. *Neuron*. 2011;71:332–47.
 72. Hartung H, Brockmann MD, Poschel B, De Feo V, Hanganu-Opatz IL. Thalamic and entorhinal network activity differently modulates the functional development of prefrontal-hippocampal interactions. *J Neurosci*. 2016;36:3676–90.
 73. Cichon NB, Denker M, Grun S, Hanganu-Opatz IL. Unsupervised classification of neocortical activity patterns in neonatal and juvenile rodents. *Front Neural Circuits*. 2014;8:50.
 74. Jaramillo TC, Liu S, Pettersen A, Birnbaum SG, Powell CM. Autism-related neuroligin-3 mutation alters social behavior and spatial learning. *Autism Res*. 2014;7:264–72.
 75. Hung AY, Futai K, Sala C, Valtchanoff JG, Ryu J, Woodworth MA, et al. Smaller dendritic spines, weaker synaptic transmission, but enhanced spatial learning in mice lacking Shank1. *J Neurosci*. 2008;28:1697–708.
 76. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016;536:285–91.
 77. CY RK, Merico D, Bookman M, LH J, Thiruvahindrapuram B, Patel RV, et al. Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. *Nat Neurosci*. 2017;20:602–611.
 78. Chen Z, Cobb MH. Regulation of stress-responsive mitogen-activated protein (MAP) kinase pathways by TAO2. *J Biol Chem*. 2001;276:16070–5.
 79. Chen Z, Hutchison M, Cobb MH. Isolation of the protein kinase TAO2 and identification of its mitogen-activated protein kinase/extracellular signal-regulated kinase binding domain. *J Biol Chem*. 1999;274:28803–7.
 80. Zhang H, Macara IG. The PAR-6 polarity protein regulates dendritic spine morphogenesis through p190 RhoGAP and the Rho GTPase. *Dev Cell*. 2008;14:216–26.
 81. Riedl J, Crevenna AH, Kessenbrock K, Yu JH, Neukirchen D, Bista M, et al. Lifeact: a versatile marker to visualize F-actin. *Nat Methods*. 2008;5:605–7.
 82. Riedl J, Flynn KC, Raducanu A, Gartner F, Beck G, Bosl M, et al. Lifeact mice for studying F-actin dynamics. *Nat Methods*. 2010;7:168–9.
 83. Blizinsky KD, Diaz-Castro B, Forrest MP, Schurmann B, Bach AP, Martin-de-Saavedra MD, et al. Reversal of dendritic phenotypes in 16p11.2 microduplication mouse model neurons by pharmacological targeting of a network hub. *Proc Natl Acad Sci USA*. 2016;113:8520–5.

84. McRae JF, Clayton S, Fitzgerald TW, Kaplanis J, Prigmore E, Rajan D, et al. Prevalence and architecture of de novo mutations in developmental disorders. *Nature*. 2017;542:433–8.
85. Pinto D, Delaby E, Merico D, Barbosa M, Merikangas A, Klei L, et al. Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. *Am J Human Genet*. 2014;94:677–94.
86. Portmann T, Yang M, Mao R, Panagiotakos G, Ellegood J, Dolen G, et al. Behavioral abnormalities and circuit defects in the basal ganglia of a mouse model of 16p11.2 deletion syndrome. *Cell Rep*. 2014;7:1077–92.
87. Horev G, Ellegood J, Lerch JP, Son YE, Muthuswamy L, Vogel H, et al. Dosage-dependent phenotypes in models of 16p11.2 lesions found in autism. *Proc Natl Acad Sci USA*. 2011;108:17076–81.
88. Ellegood J, Anagnostou E, Babineau BA, Crawley JN, Lin L, Genestine M, et al. Clustering autism: using neuroanatomical differences in 26 mouse models to gain insight into the heterogeneity. *Mol Psychiatry*. 2015;20:118–25.
89. Zufferey F, Sherr EH, Beckmann ND, Hanson E, Maillard AM, Hippolyte L, et al. A 600 kb deletion syndrome at 16p11.2 leads to energy imbalance and neuropsychiatric disorders. *J Med Genet*. 2012;49:660–8.
90. Shinawi M, Liu P, Kang SH, Shen J, Belmont JW, Scott DA, et al. Recurrent reciprocal 16p11.2 rearrangements associated with global developmental delay, behavioural problems, dysmorphism, epilepsy, and abnormal head size. *J Med Genet*. 2010;47:332–41.
91. Qureshi AY, Mueller S, Snyder AZ, Mukherjee P, Berman JI, Roberts TP, et al. Opposing brain differences in 16p11.2 deletion and duplication carriers. *J Neurosci*. 2014;34:11199–211.
92. Maillard AM, Ruef A, Pizzagalli F, Migliavacca E, Hippolyte L, Adaszewski S, et al. The 16p11.2 locus modulates brain structures common to autism, schizophrenia and obesity. *Mol Psychiatry*. 2015;20:140–7.
93. Chang YS, Owen JP, Pojman NJ, Thieu T, Bukshpun P, Wakahiro ML, et al. Reciprocal white matter alterations due to 16p11.2 chromosomal deletions versus duplications. *Human Brain Mapp*. 2016;37:2833–48.
94. Owen JP, Chang YS, Pojman NJ, Bukshpun P, Wakahiro ML, Marco EJ, et al. Aberrant white matter microstructure in children with 16p11.2 deletions. *J Neurosci*. 2014;34:6214–23.
95. Karayiorgou M, Simon TJ, Gogos JA. 22q11.2 Microdeletions: linking DNA structural variation to brain dysfunction and schizophrenia. *Nat Rev Neurosci*. 2010;11:402–16.

Affiliations

Melanie Richter¹ · Nadeem Murtaza^{2,3} · Robin Scharrenberg¹ · Sean H. White^{2,3} · Ole Johanns¹ · Susan Walker^{4,5} · Ryan K. C. Yuen^{4,5} · Birgit Schwanke¹ · Bianca Bedürftig¹ · Melad Henis^{1,6} · Sarah Scharf⁷ · Vanessa Kraus⁷ · Ronja Dörk⁷ · Jakob Hellmann⁷ · Zsuzsa Lindenmaier^{8,9} · Jakob Ellegood¹⁰ · Henrike Hartung^{10,13,13} · Vickie Kwan^{2,3} · Jan Sedlacik¹¹ · Jens Fiehler¹¹ · Michaela Schweizer¹² · Jason P. Lerch^{8,9} · Ileana L. Hanganu-Opatz¹⁰ · Fabio Morellini⁷ · Stephen W. Scherer^{4,5} · Karun K. Singh^{2,3} · Froylan Calderon de Anda¹

¹ Center for Molecular Neurobiology Hamburg (ZMNH), Research Group Neuronal Development, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

² Stem Cell and Cancer Research Institute, McMaster University, Hamilton, Ontario, Canada

³ Department of Biochemistry and Biomedical Sciences, Michael G. DeGroot School of Medicine, Faculty of Health Sciences, McMaster University, Hamilton, Ontario, Canada

⁴ The Centre for Applied Genomics and Program in Genetics and Genome Biology, Peter Gilgan Centre for Research and Learning, The Hospital for Sick Children, Toronto, Ontario, Canada

⁵ Department of Molecular Genetics and McLaughlin Centre, University of Toronto, Toronto, Ontario, Canada

⁶ Department of Anatomy and Histology, Faculty of Veterinary Medicine, Assiut University, Assiut, Egypt

⁷ Center for Molecular Neurobiology Hamburg (ZMNH), Research Group Behavioral Biology, University Medical Center Hamburg-

Eppendorf, Hamburg, Germany

⁸ Mouse Imaging Center, The Hospital for Sick Children, Toronto, Ontario, Canada

⁹ Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada

¹⁰ Developmental Neurophysiology, Institute of Neuroanatomy, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

¹¹ Department of Neuroradiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

¹² Center for Molecular Neurobiology Hamburg (ZMNH), Core Facility Morphology and Electronmicroscopy, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

¹³ Laboratory of Neurobiology, Department of Biosciences, University of Helsinki, Helsinki, Finland

2.2 Supplemental Data

Material & Methods

Animals

C57BL6/J Taok2 KO (*Taok2* *-/-*) mice were generated and described by Kapfhammer et al (1). Briefly, to generate a conditionally disrupted *Taok2* allele, lox P sites were introduced flanking exons 2 through 7 of the *Taok2* gene. Within intron 7, a neomycin resistance gene was inserted, flanked by lox P and *flp* sites. Upon crossing this line to mice expressing Flpe recombinase, the neomycin selectable marker was removed along with one of the lox P sites to generate the *Taok2tm1fl* allele. In the presence of Cre recombinase, the remaining lox P sites were recombined into a single lox P site, removing exons 2 through 7 and the translational start site to generate the *Taok2tm1Δ* allele. Animals were bred, genotyped and housed at the Central Animal Facility at McMaster University and University Medical Center Hamburg-Eppendorf, Hamburg. All procedures received the approval of the Animal Research Ethics Board (AREB) and the Institutional Animal Care and Use committee of the City of Hamburg, Germany (G48/13 and G43/16 acc. to the Animal Care Act, §8 from May, 18th 2006). Genotypes were identified during breeding by PCR of ear notches or tail biopsies. Two *Taok2* females were bred with one *Taok2* male per breeding cage. Both female and male *Taok2* Het and KO mice were used for experiments, only males were used for behavior analysis and dendritic and spine analysis after Golgi-Cox staining. 3-weeks old mice (P21-P24) were used after IUE and for western blot and RhoA GTPase experiments, 1- to 52-weeks old mice were used for live MRI imaging, and 8- to 10-weeks old mice were used for fixed brain MRI imaging. To obtain cortical cultures, *Taok2* Het females were timed bred with males and males were removed when a plug was observed, indicating copulation. At E16, mothers were sacrificed and brains from embryos were collected. Animals of appropriate genotype were included. No randomization was used and blinding was only used during animal behavioral analysis, EM analysis and spine motility analysis.

DNA constructs

The following constructs were used in these studies for transient expression in cell lines after liposomal transfection or for expression in neurons after Amaxa electroporation or *in utero* electroporation, respectively: Wild-type TAOK2 α (RC214297, NM_016151) and β isoforms (SC117141, NM_004783.2) were purchased (Origene, Rockville, MD). Human-derived

mutations in TAOK2 were generated by site-directed mutagenesis (Agilent, Santa Clara, CA). cDNA sequences of WT or mutated TAOK2 were re-cloned into a recombinant adeno-associated virus (rAAV)-vector under control of a chicken β -actin promoter (pCAGIG) (2). Silencing RNA shRNA ctrl and TAOK2 shRNA were used as described before (3). Palmitoylated EGFP plasmid originates from Clontech, (Mountain View, CA, 6085-1). Flag-Jnk1a1 (pCDNA3) was a gift from Roger Davis (Addgene, Cambridge, MA; #13789); LifeAct-GFP from Dr. Frank Bradke (4, 5) and GST-Rhotekin-RBD (pGEX) from Dr. Elena Pasquale (6).

Antibodies

The following antibodies were used in these studies for immunocytochemical or western blotting analysis: mouse anti α -tubulin (Abcam, Cambridge, UK; ab7291, ICC 1/500, WB 1/5.000); goat anti-TAOK2 (Santa Cruz Biotechnology, Dallas, TX; sc-47447, ICC 1/200, WB 1/1.000); rabbit anti-pTAOK2 (ser181) (Santa Cruz Biotechnology, sc-135712, WB 1/1.000; R&D Systems, Minneapolis, MN; PPS037, WB 1/1.000); mouse anti-RhoA (Santa Cruz Biotechnology, sc-418, WB 1/100); mouse anti-pJNK1/2 (Cell Signaling Technology/NEB, Danvers, MA; 9255L, WB 1/250); mouse anti pan-JNK (BD Transduction Laboratories, San Jose, CA; 610627, WB 1/500); rabbit anti-SynGAP (Synaptic Systems, Göttingen, Germany; ABR-01255, ICC 1/200, WB 1/1.000); rabbit anti-Synaptophysin-1 (Synaptic Systems, 101002, WB 1/1.000); mouse anti PSD-95 (Millipore, Billerica; MA; MABN68, WB 1/1.000); rabbit anti-CUX-1 (Santa Cruz Biotechnology, sc-13014, IHC 1/200); rat anti-Ctip2 (Abcam, ab18465, IHC 1/50); mouse anti- β -actin (Sigma, St. Louis, MI; A5316, WB 1/5.000); rabbit anti-FAK (Santa Cruz Biotechnology, sc-271195, WB, 1/5.000); mouse anti-GAPDH (Santa Cruz Biotechnology, sc-32233, WB 1/5.000) and chicken anti-GFP (2B Scientific, Upper Heyford, UK; GFP-1020, WB 1/5.000). Nuclei were visualized with Hoechst (Invitrogen, Carlsbad, CA; 33258, ICC/IHC 1/10.000) and F-actin with rhodamine-conjugated Phalloidin (Cytoskeleton, Denver, CO; PHDR1, ICC 1/300). HRP- (Dianova, Hamburg, Germany; WB 1/1.000-1/5.000) and Alexa-conjugated secondary antibodies (Life Technology, Carlsbad, CA; ICC 1/1.000).

Magnetic Resonance Imaging of fixed brains

8- to 10-weeks old mice were perfused with 4% PFA (with 2mM Prohance) in PBS. Mice were genotyped after perfusions and only mice with clear genotypes were used. Brains were left intact

in the skull, and the zygomatic bone and muscles were removed. Brains were then left in 4% PFA (with 2mM Prohance) in PBS overnight and transferred to PBS (with 0.02% Sodium Azide and 2mM Prohance) and kept at 4°C. All mouse imaging was done at the Mouse Imaging Center (The Hospital for Sick Kids). A multi-channel 7.0 MRI scanner (Varian Inc., Palo Alto, CA) was used to image the brains within skulls. Sixteen custom-built solenoid coils were used to image the brains in parallel (7). Parameters for the anatomical MRI scans are as follows: T2-weighted, 3D fast spin-echo sequence, with a cylindrical acquisition of k-space, and with a TR of 350ms, and TEs of 12ms per echo for 6 echoes, field-of-view of 20x20x25 mm³ and matrix size = 504x504x630 giving an image with 0.040mm isotropic voxels. Total imaging time is currently 14h. To visualize and compare any changes in the mouse brains the images are linearly (6 parameter followed by a 12 parameter) and non-linearly registered together. All scans can then be resampled with the appropriate transform and averaged to create a population atlas representing the average anatomy of the study sample. The result of the registration is to have all scans deformed into alignment with each other in an unbiased fashion, allowing for the analysis of the deformations needed to take each individual mouse's anatomy into the final atlas space to model how deformation fields relate to genotype (8, 9). The jacobian (a measure based on the deformation of each brain indicating expansion or contraction at that voxel) determinants of the deformation fields are then calculated by warping a pre-existing classified MRI atlas onto the population atlas. These measurements were then examined on a voxel-wise basis in order to localize the differences found within regions or across the brain, in a total of 182 different regions. Multiple comparisons in this study were controlled for using the false discovery rate (10).

Magnetic resonance imaging (MRI) of live mice

MRI was performed using a dedicated 7 Tesla small animal MRI (ClinScan, Bruker, Ettlingen, Germany) with a mouse head 4 element phased array receiver surface coil and a linear polarized rat body transmit coil. The animals were anesthetized with a gas mixture of oxygen and about 1% isoflurane. The oxygen was delivered with a flow rate of 0.5L/min and the isoflurane was applied via a vaporizer (Föhr Medical Instruments, Seeheim-Oberbeerbach, Germany). The animal's respiratory rate (about 100/min) was monitored using a small animal monitoring system (SA Instruments, Stony Brook, NY). Animal body temperature was not monitored, but a pad

with circulating water of 37°C covering the back of the animal helped maintaining the body temperature during MRI. MRI for volumetric analysis was done using a 3D constructive interference steady state (CISS) sequence with echo time = 3.87 ms, repetition time = 7.74 ms, flip angle = 50°, readout bandwidth = 200 Hz/pixel, 4 averages, field of view = 16x16x14.4 mm³, matrix = 128x128x120, elliptical k-space sampling and 12:25 min scan time. MRI for microstructural analysis was done using diffusion tensor imaging (DTI) with a diffusion weighting b-value of 1000 s/mm² and 12 spatially non-concordant diffusion weighting directions, echo time = 43 ms, repetition time = 20000 ms, flip angle = 90°, readout bandwidth = 2790 Hz/pixel, 2 averages, field of view = 19x19 mm², matrix = 128x128 with 5/8 partial Fourier acquisition, 30 slices with 0.5 mm slice thickness, no gap between slices and 9:00 min scan time. The DTI parameters apparent diffusion coefficient (ADC) and fractional anisotropy (FA) were calculated using the FMRIB Software Library (FSL) v5.0 software tools *eddy_correct* for correcting eddy currents and movements (11) and *dtifit* for fitting the diffusion tensor model to each voxel (12). Image analysis was done using MRICron (13). Volumes of interest (VOIs) were manually defined for each group (WT, HET, KO) at each time point (PN8, 4-, 16- and 52-weeks of age) on mean CISS or b=0 DTI data sets comprised from realigned data sets of all animals of each group and time point. (Fig. 1c-f and Supplementary Fig. 1k-o). The mean data sets were computed by realigning all individual data sets to one reference data set within the cohort. The FSL tool *flirt* was used for data realignment for improved optimization for the robust and accurate linear registration and motion correction of brain images (14). The reference data set was manually chosen based on the criteria of good image quality, coverage, position and orientation. The realigned data sets were then averaged using the FSL tool *fslmath* to generate the mean data set. Mean data sets were separately calculated for each group at each time point since brains differ too much between groups and time points and allow no sufficient realignment. The VOIs defined on the mean data sets were transferred to the individual image data using the inverse transformation matrix previously calculated by the *flirt* FSL tool. The volumes of the individual VOIs of the CISS data and the median ADC and FA values of these VOIs in the individual DTI data were extracted using an in house written *Matlab* (The Mathworks, Natick, MA, USA) script. Furthermore, the curvature of the cortex VOIs defined in the CISS data was also extracted using an in house written *Matlab* code (Fig. 1d). Briefly, a surface was fitted to the top points of the cortex VOI using the locally weighted scatter plot smoothing algorithm with the

quadratic regression model (*loess*), then the mean and the Gaussian curvature was calculated and averaged over the midline cortex surface area.

In vivo electrophysiology

Animals. All *in vivo* local field potential (LFP) were performed in P7-10 WT controls or P8-10 TAOK2 KO mice. TAOK2 KO mice had significantly smaller body weight at the time of recording than WT pups at similar age (WT: 4.78 ± 0.19 g, n=11, KO: 4.12 ± 0.12 g, n=12, p=0.019). Hence to avoid that any differences between the groups could result from the lower body-weight of the KO mice, an additional control group was matched for body weight (WT: 4.03 ± 0.82 g, n=9, KO: 4.12 ± 0.12 g, n=12, p=0.57) by including younger P7 mice. However, this resulted in a significantly different age distribution between the groups (WT: 7.78 ± 0.32 days, n=9, KO: 9.25 ± 0.22 , n=12, p=0.002).

Surgery. Mouse pups were initially anesthetized with isoflurane (induction 5% in O₂) followed by i.p. administration of urethane (1 g/kg body weight). The head of the pup was fixed into the stereotaxic apparatus using two small metal bars fixed with dental cement on the nasal and occipital bones, respectively. The bone over the regions of interest (prelimbic subdivision (PL) of the PFC, intermediate HC) was carefully removed by drilling holes of less than 0.5 mm in diameter. Removal of the dura mater by drilling was avoided, since leakage of cerebrospinal fluid or blood damps the cortical activity and neuronal firing. The body of the animal was surrounded by cotton and kept at a constant temperature of 37 °C by placing it on a heating blanket. A local anesthetic (0.25% bupivacaine/1% lidocaine) was administered. After 20-30 min recovery period, multi-site electrodes (NeuroNexus Technologies, Inc., Ann Arbor, MI, USA) were inserted perpendicularly to the skull surface into PL until a depth of 1.7-2.5 mm and at a 20° from the vertical plane into HC at a depth of 1.2-1.7 mm. In each experiment, the electrodes were labeled with DiI (1,1'-dioctadecyl-3,3',3'-tetramethyl indocarbocyanine, Invitrogen) to enable post-mortem in histological sections the reconstruction of electrode tracks in PFC and HC (Supplementary Fig. 3a,d). Two silver wires were inserted into cerebellum and served as ground and reference electrodes.

Recording protocols. Simultaneous recordings of LFP and multi-unit activity were performed from the prefrontal subdivision PL (0.5-0.7 mm anterior to bregma and 0.3-0.5 mm from the

midline) and the CA1 area of the intermediate HC (3.5-3.7 mm posterior to bregma, 3.5-3.8 from the midline) using similar protocols as described previously (15). One shank Michigan electrodes with 16 recording sites (0.5-3 M Ω) that were separated by 50 μ m (for HC) or 100 μ m (for PL) were used. The position of recording sites over the PL and CA1 area was confirmed by post-mortem histological evaluation. Both LFP and MUA were recorded for at least 60 min at a sampling rate of 32 kHz using a multi-channel extracellular amplifier (Digital Lynx 4S with no gain, Neuralynx, Bozeman, MT, USA) and the Cheetah acquisition software (Westlake Village, CA, USA). During recording the signal was band-pass filtered between 0.1 Hz and 5 kHz.

Data analysis. Channels for analysis were selected on the basis of post-mortem histological investigation, i.e. which recording sites of fluorescently marked electrodes were confined to PL and hippocampal CA1, and the presence of specific patterns of activity. In the PL, these patterns were nested theta bursts with superimposed beta-low gamma oscillations, whereas in the HC the reversal of LFP over Str. pyramidale was used for the selection of the channel with sharp-waves of minimum amplitude and consequently, lowest contribution to the spectral content of the signal. Data were imported and analyzed off-line using custom-written tools in Matlab software version R2015a (The Mathworks, Natick, MA, USA). For LFP analysis, the signals were low-pass filtered (<1500 Hz) using a third order Butterworth filter before reducing the sampling rate to 3200 Hz. The detection and classification of discontinuous patterns of activity in the neonatal PL and hippocampal CA1 area were performed using a modified version of the previously developed algorithm for unsupervised analysis of neonatal oscillations (16) and confirmed by visual inspection. Signals were filtered between 4-100 Hz using a zero phase filter. Fragmented detection of oscillations was avoided by considering events with inter-event intervals <100 ms for PL and <300 ms for HC as single events. Only oscillatory events lasting >1 s in PL and >1.5 s in HC and containing at least three cycles were considered for further analysis. The detected events in PL and HC were analyzed in their occurrence (defined as the number of bursts per min), duration, mean amplitude and maximum amplitude (defined as the voltage difference between 0 and the maximal positive peak) and power distribution. Power spectral density estimates of every detected oscillatory event were calculated on the raw signal using the MATLAB function 'FFT' and averaged over all events from one animal. Mean power in different frequency bands (4-12 Hz, 12-30 Hz, 30-100 Hz) were quantified. Time-frequency

plots were calculated by transforming the LFP events using the function ‘WT’ from the wavelet coherence Matlab toolbox. As spectral measure of correlation between two signals coherence was calculated from the cross-spectral density and normalized by the individual power spectral density of each. The computation was performed using a multi-taper approach implemented in the Chronux toolbox (chronux.org) by the function ‘coherencysege’ on 4-100 Hz filtered data. For this a continuous signal for either PL or HC was generated by gluing as many multiples of 2s-long segments that could be fitted of all co-occurring oscillatory bursts in either PL or HC into single vectors. Frequency domains with significant coherence were determined by Monte Carlo simulation. For this, LFP segments of 2 s from one region were shuffled and the coherence was calculated between the shuffled LFP from one region and the original LFP from the other region. After 100 iterations, the 95th percentile of the resulting distribution was used as significance threshold. The mean coherence for all frequency components (4-12 Hz, 12-25 Hz and 25-40 Hz) of oscillatory events that were detected as peaks in the power spectra were calculated.

Statistics. Statistical analyses were performed with IBM SPSS Statistics version 21 (SPSS GmbH). Generally all values were tested for normal distribution by the Kolmogorov-Smirnov test. For normally distributed values unpaired t-test was used. For not normally distributed values the Shapiro-Wilk test was used.

In situ electrophysiology

As described previously (17), coronal brain slices (400 μm) were prepared in cold sucrose-based slicing solution containing (in mM): 160 sucrose, 2.5 KCl, 10 MgSO_4 , 1.25 NaH_2PO_2 , 25 glucose, 30 NaHCO_3 , 20 hepes, 5 Na-ascorbate, 3 Na-pyruvate, 2 thiourea, 0.5 CaCl_2 . Slices recovered for 45 minutes at 30°C, followed by 45 min at room temperature. Visually guided whole-cell recordings (BX51WI, Olympus, Tokyo, Japan) were performed using an Axoclamp 700B amplifier (Molecular Devices, Sunnyvale, CA) from patch electrodes (P-97 puller, Sutter Instruments, Novato, CA) containing a cesium based intracellular solution (in mM): 100 CsCl, 100 gluconic acid, 10 hepes, 0.5 EGTA, 10 Na-phosphocreatine, 2 MgATP, 0.5 NaGTP, pH 7.3 with CsOH). Composition of aCSF (in mM): 120 NaCl, 2.5 KCl, 1 MgSO_4 , 26 NaHCO_3 , 10 glucose, 2 CaCl_2 . 1 μM TTX and 100 μM picrotoxin was added to the bathing medium to block Na-dependent action potentials and GABA currents, respectively for mEPSCs. 1 μM TTX and

1mM Kynurenic Acid was added to the bathing medium to block Na-dependent action potentials and mEPSCs, respectively for mIPSCs. Recordings were performed at -70 mV for mEPSCs and +10mV for mIPSCs using Clampex 10.6 (Molecular Devices), corrected for a calculated -10 mV junction potential and analysed using the Template Search function from Clampfit 10.6 (Molecular Devices, Sunnyvale, CA).

Behavior analysis

Animal Husbandry. *Taok2* 6- to 8-weeks old male WT, Het and KO littermates from heterozygous breeding pairs were transferred from the breeding facility into a *vivarium* with an inverted 12h/12h-light/dark cycle (light off at 8:00 am) and maintained under standard housing conditions ($21 \pm 1^\circ\text{C}$, 40-50% humidity, food and water *ad libitum*). Behavioral experiments were performed with 10- to 18-weeks old mice during the dark cycle in a room next to the *vivarium* illuminated with dim red light. Tests started and ended at least two hours after light offset and three hours before light onset, respectively. The experimental material was cleaned with soap, water, and ethanol (70%) before and after each contact with an animal. To avoid a “litter effect”, no more than two animals per genotype were used from the same litter. Tracks representing the position of the mice were created and analyzed with the software EthoVision (Noldus, Wageningen, The Netherlands) as described (18). Manual scoring of behavior was performed by a trained experimenter blind to the genotype of the mice using the *The Observer* software (Noldus). The experimenter trained himself until he repeatedly scored at least 90% of consistency between two analyses performed at different times on the same mouse, as calculated with the Reliability Test provided by the software The Observer (having 1s as maximal time discrepancy between two evaluations).

Open field. The open field test was performed in a box (50 x 50 cm and 40 cm high) illuminated with white light (25 lux). Mice were started from one corner of the box and videorecorded for 30 min. Distance moved, mean minimal distance to walls, time spent in the center (an imaginary 25 x 25 cm square in the middle of the arena) were analyzed with the software EthoVision (Noldus), whereas the parameters rearing on wall (the mouse stands on the hind limbs and touches the wall with at least one fore paw) and self-grooming were analyzed with The Observer (Noldus).

Elevated plus maze. The maze had the shape of a plus with four 30 cm long and 5 cm wide arms, connected by a squared center (5 x 5 cm). Two opposing arms were bordered by 15 cm high

walls (closed arms), whereas the other two arms (open arms) were bordered by a 2 mm rim. The maze was elevated 75 cm from the floor and an infrared camera allowed video recording under total darkness. The mouse was placed into the center facing one open arm and left on the maze for 5 min. The following parameters were analyzed with The Observer: entries into the open and closed arms (calculated when all the four paws were on an arm), total transitions (the sum of entries into the open and closed arms), rearing and self-grooming.

Social preference. Motivation to investigate a social stimulus was tested by giving the experimental mouse the choice to investigate a beaker containing an unfamiliar sex-matched mouse or an empty beaker. The apparatus consisted of a squared box (50 cm x 50 cm and 40 cm high) illuminated with white light (10 lux). Two beakers made of transparent plastic (diameter 10 cm, 15 cm high) with several holes (diameter= 1 cm) drilled at the bottom were placed at two opposite corners of the box. An unfamiliar adult male mouse was confined in one of the two beakers. The test started by placing the experimental mouse next to the beaker containing a male mouse and lasted 5 min. Distance moved and time at the beakers were analyzed with EthoVision, whereas rearing, self-grooming and time spent sniffing the two beakers were analyzed with The Observer.

Spontaneous alternation. The spontaneous alternation paradigm was performed in a Y maze to test for working memory performance (19). The maze consisted of three equally sized arms (34 cm x 5 cm x 30 cm) made of transparent Plexiglas connected such as to make a Y and illuminated with 20 lux. Mice were placed in the center of the maze and allowed to freely explore the maze until they performed 27 transitions or after a maximal given time of 20 min. An entry into any arm with the four paws was considered as transition. An entry into a new arm after having visited the two other arms was considered as alternation. Data were analyzed as percentage of alternations over all transitions. Average time to make a transition was calculated by dividing the duration of the test by the number of total transitions.

Spatial object recognition. This test was designed to test short-term spatial memory. The paradigm is based on the same assumption as the spontaneous alternation test, namely that rodents have an intrinsic tendency to investigate novel stimuli. In other words, the test is a sort of delayed-no-match-to-place paradigm (19). The spatial object recognition test was performed in the same box used for the open field test divided into two equal compartments by a white PVC wall with a hole in the middle to allow a mouse to move between the two compartments. The

arena was illuminated with white light (15 lux). The paradigm consisted of two trials, the exposure (10 min) and the test (5 min) trials with an inter-trial interval of 10 min during which mice were returned to their housing cage. During the exposure trial two identical objects unfamiliar to the mice were placed in one corner of each compartment. During the test trial one of the two objects was displaced to another corner of the compartment. Distance moved, mean velocity and time spent at the objects were analyzed with Ethovision. Sniffing the objects, self-grooming, rearing on wall and rearing on objects were analyzed with The Observer.

Contextual fear conditioning. In the contextual fear conditioning test mice had to learn the association between the unconditioned (electric foot-shock) and conditioned (context) stimuli. Mice were conditioned in the context, a chamber (23.5 cm x 23.5 cm and 19.5 cm high) with Plexiglas walls and ceiling and a stainless grid floor from which an electric shock could be elicited. The chamber was illuminated by white light (20 lux). Mice were placed in the center of the cage and received four electric foot shocks (0.35 mA, 1 s) at 120, 160, and 200 s. At 220 s the recording ended, and the mouse was immediately returned into its home cage. Mice were tested for long-term memory retention one day after conditioning. The conditioned response was analyzed by quantifying the percentage of time spent freezing (defined as absence of body movements for at least 1 s). Freezing behavior was automatically analyzed using a modified version of the system Mouse-E-Motion (Infra-e-motion, Hamburg, Germany).

Water maze. The water maze consisted of a circular tank (145 cm in diameter) circled by dark curtains. The water was made opaque by the addition of non-toxic white paint such that the white platform (14 cm diameter, 9 cm high, 1 cm below water surface) was not visible. Four landmarks (35 X 35 cm) differing in shape and grey gradient were hung on the wall of the maze. Light was provided by four white spotlights placed on the floor around the swimming pool that provided homogeneous illumination of 60 lux on the water surface. Before the experiment started, mice were familiarized for one day to swim and climb onto a platform (diameter of 10 cm) placed in a small rectangular maze (42.5 x 26.5 cm and 15.5 cm high). During familiarization the position of the platform was unpredictable since its location was randomized and training was performed under darkness. After familiarization mice underwent three learning days during which they had to learn the location of a hidden platform. Starting position and position from which mice were taken out of the maze were randomized. At day 1 and 2 mice underwent four learning trials (maximum duration 90 seconds, inter-trial interval of 10 minutes). After staying on the platform

for 15 s, mice were returned to their home cage and warmed up under red light. Day 3 consisted of a 60 seconds long transfer trial (long-term memory transfer trial) during which the platform was removed.

Hippocampal and Cortical Cultures

Primary cortical (17) and hippocampal cultures (20) were prepared as described previously. When mouse cortices from heterozygous *TAOK2* mice were prepared, the cortices were kept in Hibernate E buffer (Life Technologies) until genotyping of the embryos was performed. Cortices from either WT or KO embryos were pooled. Tissue was digested in 0.25% Trypsin or 31.25µg/ml papain and 10µg/ml DNaseI for 10-20min at 37°C. Trypsination or papain digestion was stopped by adding DMEM + 10%FCS or Neurobasal + 10% FBS and washed three times. Dissociated neurons were either plated directly on poly-D-lysine and laminin pre-coated glass coverslips in Neurobasal/B27 medium (Life Technologies) or underwent Amaxa electroporation (Lonza/Amaxa, Cologne, Germany).

Amaxa Electroporation

Amaxa electroporation (Lonza) was carried out according to the manufacturer's protocol and described elsewhere (21). In detail, 5×10^6 cells were resuspended in 100µl Nucleofector solution with a total of 3µg DNA or the indicated amount of shRNA added. Rat hippocampal neurons and mouse cortical neurons were transfected using the appropriate Nucleofector kit using program G-013 or O-005, respectively. Transfected neurons were plated on glass coverslips in 24-well chambers (50,000 cells/well), pre-coated with 0.1mg/mL Poly-D-Lysine (BD Sciences)/3.3ug/mL Laminin (Sigma) for immunocytochemical analysis or on pre-coated 4-well LabTek chamber slides (300,000 cells/chamber, Sarstedt, Nümbrecht, Germany) for time-lapse life imaging.

Immunofluorescence

Dissociated cultures: Neurons were maintained in culture for 14 or 15 days *in vitro* (DIV) to analyze spine formation, for 19DIV to perform spine rescue and spine motility assays or for 21DIV to analyze mature spine morphology and then fixed for immunofluorescence analysis. Neurons were fixed with 4% PFA/4% sucrose at RT for 10 min, followed by 3 washes with PBS.

After blocking in donkey serum (Sigma, St. Louis, MI), neurons were incubated with the primary antibodies. *Cortical sections*: Brains were removed and immersion-fixed overnight in 4% PFA/4% sucrose and thereafter transferred to sterile 30% sucrose/PBS (4 °C, overnight). Brains were embedded in OCT compound (Sakura Finetek, Tokyo, Japan), frozen at -80°C and sectioned in a cryostat (Leica, Wetzlar, Germany). 20- to 30- μ m cryosections were incubated overnight at 4 °C with the primary antibodies.

Quantitative immunofluorescence of cultured cortical neurons

Neurons were cultured for 14DIV after plating and fixed as described above. Total TAOK2, filamentous actin, tubulin and the marker protein GFP_{palm} or Venus (pCAGIG) were visualized by indirect immunofluorescence. The mean intensity gray value of a line drawn along the soma and the neurites was measured using ImageJ. To measure TAOK2 localization, images were put under the same threshold value for each fluorescence channel; green for GFP and red for TAOK2. After threshold was set, total area encompassed by TAOK2 and GFP inside the dendrites or cell body was measured using ImageJ. Total area covered by TAOK2 was then divided by total GFP area of the dendrites or cell body alone.

Electronmicroscopy

Selected coronal vibratome sections (PFC, layer II/III, Bregma 1.94 to 1.34 mm, Interaural 5.74 to 5.14), primary somatosensory regions S1HL, S1Fl, S1BF and S1Tr (SSC, layer V, Bregma -1.46 to -1.94 mm, Interaural 2.34 to 1.86 mm), intermediate hippocampus (iHC, CA1, Bregma -2.54 to -2.80 mm, Interaural 1.26 to 1.00 mm) (22) were collected and prepared for electronmicroscopy as described before (23). Briefly, sections were rinsed three times in 0.1 M sodium cacodylate buffer (pH 7.2–7.4) and after osmication using 1% osmium tetroxide in cacodylate buffer, sections were dehydrated using ascending ethyl alcohol concentration steps, followed by two rinses in propylene oxide. Infiltration of the embedding medium was performed by immersing the tissue in a 1:1 mixture of propylene oxide and Epon and finally in neat Epon and hardened at 60°C. Semithin sections (0.5 μ m) were prepared for light microscopy mounted on glass slides and stained for 1 minute with 1% Toluidine blue. Ultrathin sections (60nm) were examined in an EM902 (Zeiss, Munich, Germany). Pictures were taken with a MegaViewIII digital camera (A. Tröndle, Moorenweis, Germany). Discrimination of dendritic versus spine

innervation was done acc. to (23-26). Briefly, we distinguished spine versus shaft synapses through determining the postsynaptic spine in ‘spine synapses’ as a dendritic spine that contains a clearly recognizable post-synaptic density (PSD), that does not contain mitochondria in close proximity to the PSD and that is filled with electron-dense material. Whereas in a ‘shaft synapse’, the postsynaptic site, defined by the presence of a PSD, is not as morphologically distinct as the spine head and is filled with rather electron-light material (clear appearance). Importantly, dendritic shaft innervation is often, but not necessarily, characterized by the presence of mitochondria in close proximity to the PSD.

Regional brain lysates

Indicated brain regions from postnatal 21- to 24-days old C57BL/6 mice were dissected and homogenized in 20x vol/weight of either sterile-filtered RIPA buffer (50mM Tris-HCl pH 7.4, 150mM NaCl, 1mM EGTA, 1% NP-40, 0.25% sodium deoxycholate, 10mM Na₃VO₄, 10mM NaF, containing proteinase inhibitors (Roche, Basel, Switzerland) or RhoA GTPase buffer (50mM Tris-HCl pH 7.2, 500mM NaCl, 10mM MgCl₂ x 6 H₂O, 1% Triton-100, 0.1% SDS, 10mM Na₃VO₄, 10mM NaF, containing proteinase inhibitors (Roche, Basel, Switzerland). To clear lysates, they were centrifuged at 20.500 x g for 10 min at 4°C. Supernatants were either applied to RhoA GTPase assay immediately or were analyzed by BCA method for concentration determination and afterwards applied to immunoprecipitation or SDS gel separation and western blotting.

Immunoprecipitation

400µg of total lysates for Taok2 immunoprecipitation or 1500µg of total lysates for co-immunoprecipitation of RhoA and Taok2, respectively, were incubated overnight at 4°C with 2µg of pre-conjugated polyclonal TAOK2 antibody (Santa Cruz). Antibodies were conjugated before to 15µl of magnetic Dynabeads™ Protein G (Invitrogen) for 2hrs at RT. The bound immune complexes were collected by a magnet, followed by 5 washes in sterile-filtered RIPA buffer. After the final wash, supernatants were completely removed and beads were eluted in 2x SDS-PAGE sample buffer, boiled for 5 min at 95 °C and separated on 4-20% SDS-polyacrylamide gradient gels (Invitrogen).

RhoA Pull Down

Recombinant GST-Rhotekin-RBD was expressed in freshly transformed DH5 α -E. coli, purified through Glutathione-Sepharose 4B beads (Qiagen) as described elsewhere and stored in appropriate aliquots at -80°C (27). SHSY5Y cell lysates, freshly homogenized cortical or hippocampal brain lysates or LCL lysates were obtained in RhoA pull-down buffer and cleared by centrifugation for 10min at 4°C. The RhoA activation assay was performed using a commercial RhoA Pull-down kit (Cytoskeleton, Biochem Kit, #BK036, Denver, CO) or as described (21). Briefly, the supernatant was applied completely to ~50 μ g of freshly thawed recombinant GST-Rhotekin-RBD beads per pull-down and incubated for 45min at 4°C. A small portion of ~50 μ l of supernatant was saved before for detection of total RhoA in the lysate. After incubation, GST-Rhotekin-RBD beads were washed thoroughly for three times. After the final wash, supernatants were completely removed and beads were eluted in 2x SDS-PAGE sample buffer, boiled for 5 min at 95°C and separated on 4-20% SDS-polyacrylamide gradient gels (Invitrogen).

Transfection and Western blot analysis of transfected cell lines

HEK-293T and SHSY5Y cells (both Invitrogen, Carlsbad, CA) were grown under standard cell culture conditions. HEK-293T cells were transfected with plasmids using Lipofectamine™ LTX with Plus™ reagent according to the manufacturer's protocol (Invitrogen). For transient transfection, HEK293T cells were co-transfected with equal amounts of TAOK2 variants and JNK1a1 with Venus (GFP) or with TAOK2 variants only and Venus. The total amount of transfected DNA was 4 μ g per 1-well in a 6-well plate. The cells were allowed to express the constructs for 48 hours before lysis and analysis. SHSY5Y, a fast growing neuronal cell line, was used for testing the commercial RhoA activator (CN01, Cytoskeleton).

Proteins from cell lysates in RhoA GTPase buffer, containing proteinase inhibitors (Roche) or RIPA buffer, containing proteinase inhibitors (Roche) were either applied to RhoA GTPase assay immediately or were analyzed by BCA method (Thermo, Waltham, MA) for concentration determination and afterwards applied to immunoprecipitation or SDS gel separation and western blotting. Proteins were separated on 4-20% or 10% SDS-polyacrylamide gradient gels (Invitrogen) at 30mA/gel and transferred to Immobilon-P PVDF membranes (Millipore) at 35mV overnight at RT. Membranes were blocked in TBS-T (50mM Tris-HCl pH 7.4, 150mM

NaCl, 0.1% Tween-20) with 3% BSA, Fraction V (Roth, Karlsruhe, Germany) or with 3% skim milk powder for 1h at room temperature and then incubated with primary antibody for 2 h to overnight at 4 °C. Membranes were washed for 30 min in TBS-T and incubated for 2hrs at room temperature with horseradish peroxidase-conjugated secondary antibodies (Dianova or GE healthcare) and then washed for 30 min in TBS-T. Immunoreactivity signals were detected by enhanced chemiluminescence (INTAS, ChemoStar, Ahmedabad, India or Amersham ECL, GE Healthcare, Chicago, IL).

Lymphoblastoid cell lines (LCLs) were created using peripheral blood mononuclear cells isolated from the blood of probands and their families that were immortalized using the Epstein-Barr virus (28). Lymphoblastoid cell lines were cultured under standard suspension condition in RPMI 1640 media (with %15 FBS and 1mM sodium pyruvate). LCLs were cultured at 200.000 cells/ml in 5ml and media was doubled every 3-5 days until reaching 20ml. Cells were counted and 5 million cells were pelleted and lysed. Proteins from cell lysates were either applied to RhoA GTPase assay immediately or were analyzed by BSA method (Thermo) for concentration determination and afterwards applied to immunoprecipitation or gel separation and western blotting. Proteins were separated on 12% SDS-polyacrylamide Tris-glycine gels at 100V and transferred to Immobilon-P PVDF membranes (Millipore) for Western Blot analysis.

Quantitative droplet digital PCR

cDNA was synthesized from RNA extracted from cultured lymphoblastoid cell lines using the Superscript III 1st strand cDNA Synthesis Kit (Thermo) following manufacturer's protocol. Three RNA extractions were performed per sample and cDNA synthesis carried out from each independently. Gene expression estimation of TAOK2 was performed using the QX200 Droplet Digital PCR system (Bio-Rad Laboratories, Hercules, CA) using TBP probe Hs00427621_m1 as an endogenous control in a duplex reaction mode. Two Taqman probes were used to measure TAOK2; Hs01586199, which targeted all isoforms and Hs01026246 which targeted isoform NM_004783.3 only. The 20ml gene expression reaction mix consisted of 10 ml of 2x ddPCR SuperMix for Probes (Cat No. 1863023, Bio-Rad Laboratories), 1 ml of the target assay (labelled with FAM), 1ul of the endogenous control assay (labelled with VIC), 6 ml nuclease free water and 2 ml of cDNA. The gene expressions assays were previously validated by temperature gradient to ensure optimal separation of target and control droplets. Cycling conditions for the

reaction were 95°C for 10 minutes, followed by 45 cycles of 94°C for 30 seconds and 58°C for 1 minute, 98°C for 10 minutes and finally a 10°C hold on a Life Technologies Veriti thermal cycler. Data were analysed using QuantaSoft v1.4 (Bio-Rad Laboratories). No template, no RT and Human Universal RNA controls were tested in parallel with the study samples. The ratio of total TAOK2 or TAOK2 β to TBP mRNA per sample was calculated and compared to the average ratio of the parents and siblings in each family.

Golgi-Cox Staining

Brains were harvested whole from 3-weeks old male offspring on a C57BL/6 background and stained using the FD Rapid GolgiStain kit (FD NeuroTechnologies, Columbia, MD) according to the manufacturer's protocol. Briefly, brains were rinsed with double distilled water and then immersed in a 1:1 mixture of FD Solution A:B. Solution A:B was replaced within the first 24 hours and then kept for 2 weeks at room temperature in the dark. Brains were then transferred to FD Solution C and kept in the dark at 4°C for 48 hours. Solution C was replaced after the first 24 hours. Afterwards, in preparation for vibratome sectioning, individual brains were placed in sterile 30% sucrose at 4°C for 24 hours. Coronal sections of 120 μ m thickness were cut and collected consecutively in a 24-well plate in PBS. This thickness enabled optimal staining and preservation of spines on secondary and tertiary dendritic segments while the complete basal dendritic arbor was acquired as needed for Sholl analysis. Sections were mounted on Superfrost Ultra Plus microscope slides (Thermo Scientific) and allowed to dry for 5-10 min. Sections were developed exactly as described in the FD Rapid GolgiStain instructions. The development solution should be prepared just before use combining FD Solution D:E:H₂O_{dd} in a 1:1:2 mixture. Bottle and staining jar must be always covered to prevent vaporization. Glass slides were immersed for 3-5 min, rinsed in H₂O_{dd} for two times, 4 min each and allowed to dry briefly. Each slide was embedded using 120-150 μ l Mowiol (Sigma).

In utero Electroporation

Time-pregnant mice were given a pre-operative dose of buprenorphine (0.05-0.01 mg/kg body weight) by subcutaneous injections at least 30 min before surgery. Animals were then anesthetized using 2.5% isoflurane/O₂ inhalation. Oxygen was delivered with a flow rate of 0.65L/min and the isoflurane was applied via a vaporizer (Föhr Medical Instruments, Seeheim-

Oberbeerbach, Germany). The uterine horns were exposed and TAOK2 expression plasmids mixed with Fast Green (Sigma) were microinjected into the lateral ventricles of the embryos. The expression plasmid concentration was three times that of the fluorescence marker Venus. Five current pulses (50ms pulse, 950ms interval; 35mV) were delivered across the heads of the embryos. Post-surgery, 2-3 drops of meloxicam (0.1-0.5 mg/kg body weight) were given orally through soft food for 96 hrs. The offspring was delivered and perfused at three weeks of age (PN21-24) for dendritic and spine analysis. Perfusion was carried out as described elsewhere (29) and brains were post-fixed with 4%PFA/4% sucrose overnight at 4°C. Coronal sections of 120 µm thickness were cut using a vibratome and collected consecutively in a 24-well plate in PBS.

Confocal Imaging

15-20 individual pyramidal neurons derived from two to three independent coronal sections per mouse containing the cingulate cortex, area 1 and 2 (cg1 and cg2) and the prelimbic cortex (PrL) of the prefrontal cortex (PFC, layer II/III, Bregma 1.94 to 1.34 mm, Interaural 5.74 to 5.14), the primary somatosensory regions S1HL, S1FL, S1BF and S1Tr (SSC, layer V, Bregma -1.46 to -1.94 mm, Interaural 2.34 to 1.86 mm) or the intermediate hippocampus (iHC, CA1, Bregma -2.54 to -2.80 mm, Interaural 1.26 to 1.00 mm) (22) were imaged. Pyramidal neurons were identified by localization and their distinct morphology within the respective regions. Presence of a completely unclipped basal dendritic arbor was a prerequisite for selection and imaging of the cell. Z-stacks of Golgi-stained dendritic trees (60-90 µm total on Z-axis; optical section thickness per Z-section: 1µm) were taken at 60x magnification on a NiKOn Ti-E research microscope (Nikon, Tokyo, Japan) fitted with a CoolSnap HQ2 camera (Photometrics, Tucson, AZ). Z-stacks of GFP-labeled dendritic trees after *in utero* electroporation (IUE) (60-90 µm total on Z-axis; optical section thickness per Z-section: 1µm) were acquired at 40x on a Zeiss LSM 700 confocal laser-scanning microscope (Zeiss, Oberkochen, Germany).

Sholl Analysis

To perform Sholl-Analysis, images of basal dendritic arbors of Golgi-stained pyramidal cells of the cg cortex (PFC), SSC and intermediate CA1 were opened with Fiji (ImageJ) and first submitted to Phansalkars auto local thresholding (30) with the radius adjusted according to the

background of the images (10-15pxl). The thresholded images were then traced semi-automatically via the Simple Neurite Tracer plugin (31) with Hessian-based analysis enabled, resulting in 2D projections of the basal dendritic trees. The images of the basal dendritic projections were then analyzed via the Sholl Analysis plugin with radii set to 5 μ m intervals (v3.4.5) (32).

For images taken of cells labeled with cytosolic GFP, the images were not thresholded but directly traced via Fiji's (ImageJ) segmented line tool and saved as ROI's. These ROI's were then flattened to create an image of the traced basal dendritic tree and the resulting images were then analyzed via the Sholl Analysis plugin with radii set to 5 μ m intervals (v3.4.5) (32).

Spine analysis Golgi

Spine analysis was performed on images of selected cells, previously taken for Sholl Analysis, according to standardized parameters described (33). In addition, the whole basal dendritic tree was taken into account. Centroids of width measurements of spines were then utilized to create spine maps after matching the centroids of the somata at the origin and aligning the angular position of apical dendrites (34). To emphasize the spine distribution, alpha blending was set to 0.5. Additional read-out from the spine analysis was created by calculating the distance of each spine to the center of the respective cell aiming to visualize the distribution of spines over the distance to the soma. Furthermore, the total number of spines per cell was matched with the respective total basal dendritic length to obtain an estimate of general spine-density. To quantify dendrite spines morphology, we used a semi-automatized protocol (33) using Fiji's (ImageJ) selection tools combined with the ROI manager. Measurements of individual spines were taken according to the cited paper, in an unbiased fashion, and a semi-automated based algorithm categorized the spines. Briefly, spine morphologies were categorized based on spine-head width and length of spines and as a result of width to length ratio. According to these parameters, spines were classified into mushroom, thin, long-thin, filopodia and stubby spines. Branched spines were marked separately during the analysis. Cumulative distributions were employed to visualize parameters of whole populations. Additionally, the distance of each spine to the center of the respective cell was calculated to visualize the distribution of spines over the distance to the soma.

Spine analysis IUE

Images of the basal dendritic trees were re-taken at 40x (to cover the whole basal dendritic field in one image) for spine-analysis with the pinhole set to 1 airy unit (0.9 μ m sections) and imaging at 0.5 μ m steps. To improve the local contrast, the size of images was adjusted through bicubic interpolation (doubling the number of pixel in each direction (2048px \rightarrow 4096px). Spine analysis was then performed as described for Golgi-stained cells. Since different objectives and microscopes were used (60x vs. 40x, light-microscope vs. laser-scanning microscope) and the original thresholds for categorizing spines were applied to Golgi-stained cells (33), the distribution of width measurements from control cells derived from IUE versus Golgi-stained cells in SSC and iHC were compared. Through re-scaling, the width distribution could be matched to the distribution of width measurements from spines of Golgi-stained SSC and CA1 leading to a new threshold for mushroom spines of 0.46154 μ m. Thresholding of length measurements were not adjusted to keep the benefits of a higher resolution through cytosolic GFP-labeling and confocal microscopy.

Heat-maps and assessment of dendritic coverage

Heat-maps were generated to create a visual representation using Fiji (ImageJ) on the basis of dendritic filaments of individual cells. The dendritic width of these filaments was reduced to one pixel, thus taking each dendritic section equally into account. Neurons were then matched in their orientation and aligned regarding their angular position of apical dendrite and the centroid of the soma. Subsequently, filaments were projected onto each other followed by the application of a mean filter (Radius = 5px). To put an emphasis on areas with a higher occurrence of dendritic material, a LUT (warm metal) was applied.

Spine Motility Assay Analysis

Live cell imaging experiments of dendritic filopodia/spines of DIV14 *Taok2* WT, KO and KO with CN01 (1 U/ml) neurons labeled with Lifeact-GFP were carried out in a defined volume of culture medium (Neurobasal/B27 medium (Life Technologies) at 37°C/5% CO₂ and imaged at 60x magnification with a Nikon Ti-E research microscope (Nikon, Tokyo, Japan) fitted with a CoolSnap HQ2 camera (Photometrics, Tucson, AZ). During the acquisition, dendritic segments were recorded with a 2-second frame rate for 5 minutes (capturing 151 frames in total). Cells

treated with the RhoA activator CN01 were imaged once, then treated with the activator and after activation immediately imaged for the second time to allow an optimal pairwise comparison. For analysis, unmodified raw images were opened with Fiji's (ImageJ) for the motility analysis. Images were modified regarding pixel density and sharpness and corrected for bleaching with included plugins. To improve the local contrast, the size of images was adjusted: Scale 2 times in x and y coordinate to 2478 x 2080 pxl through bicubic interpolation; mask unsharp with a radius of 10 pixels and a mask weight of 0.6; bleaching correction with histogram matching. Due to lower fluorescence and higher signal background ratios in LifeAct-transfected cells, we reinforced the modification in the mask unsharp with a radius of 40 pixels and added an additional step of contrast enhancement (enhance contrast with between 0.1% and 1% saturated pixels applied on all 151 frames) in the rescue assay (Fig. 6g, h). To correct for rotation or general movement of the cell, we applied the "Template matching" plugin from Qingzong Tseng (<https://sites.google.com/site/qingzongtseng/template-matching-ij-plugin>) (35). We normalized correlation coefficient; search area: 0; subpixel registration and bicubic interpolation method for subpixel translation. Where applicable, we modified all the cells with similar parameters. We used the "Dendritic Filopodia Motility Analyzer" to perform the motility analysis (36). We selected spines on 3-5 comparably sized dendritic stretches around the soma in a medial distance from 10-70 μ m to the center of the soma and analyzed 10-15 spines per cell. As an upper and lower threshold, we adjusted the default values to have the spines covered in blue with as little black surrounding as background as possible. In order to reduce that influence of the background, we chose an area for the analysis starting at the basement of the spine surrounding the spine covering area as tightly as possible. We used the starting coordinate as a reference for extension/retraction analyses.

Statistical Analysis for Sholl Analysis

Figure 3b (top): WT vs Het, *p for 95 μ m; WT vs KO, *p for 20 μ m, **p for 25 μ m. Figure 3b (bottom): WT vs Het, *p for 65, 70 μ m, **p for 60 μ m, ***p for 10, 15, 55 μ m, ****p for 20-50 μ m; WT vs KO, *p for 75, 80 μ m, **p for 65, 70 μ m, ***p for 60 μ m, ****p for 10-55 μ m. Figure 5b (left): WT+Venus vs Het+Venus, *p for 40, 55, 60, 95 μ m, **p for 65, 85, 90 μ m, ***p for 70-80 μ m; WT+Venus vs Het+TAOK2 α/β , *p for 35 μ m, **p for 10, 15, 30 μ m, ***p for 20, 25 μ m.

Figure 5b (right): Het+TAOK2 α/β vs Het+Venus, *p for 5, 30, 70 μm , **p for 15, 20, 80 μm , ***p for 75 μm , ****p for 10 μm ; Het+TAOK2 α/β vs Het+ TAOK2 α/β A135P, *p for 40, 45, 65 μm , **p for 50, 60, 70, 75 μm , ***p for 55 μm . Het+TAOK2 α/β vs TAOK2 β P1022*, *p for 70, 145, 150 μm , **p for 60, 80, 140 μm , ***p for 65 μm , ****p for 85-135 μm .

Supplementary Figure 8b (top): WT vs Het, *p for 25, 60 μm ; WT vs KO, *p for 80, 120 μm , **p for 20 μm .

Supplementary Figure 8b (bottom): WT vs KO, *p for 35-50 μm .

Supplementary Figure 9b (top): WT vs Het, *p for 35 μm .

Statistical Analysis for Spine Analysis

Figure 3e: WT vs Het, *p for 45-50 μm , **p for 40 μm , ***p for 35 μm . WT vs KO, *p for 90 μm , ***p for 80-85 μm , ****p for 35-75 μm .

Figure 5d: TAOK2 α/β vs Venus, *p for 25, 55 μm , **p for 30 μm ; TAOK2 α/β vs TAOK2 β P1022*, *p for 45-50, 125 μm , **p for 25-30, 40 μm , ***p for 35 μm ; TAOK2 α/β A135P vs Venus, *p for 45,50, 60 μm , **p for 40 μm , ***p for 25, 35, 55 μm , ****p for 30 μm ; TAOK2 β P1022* vs Venus, *p for 20, 60, 70 μm , ****p for 25-55 μm .

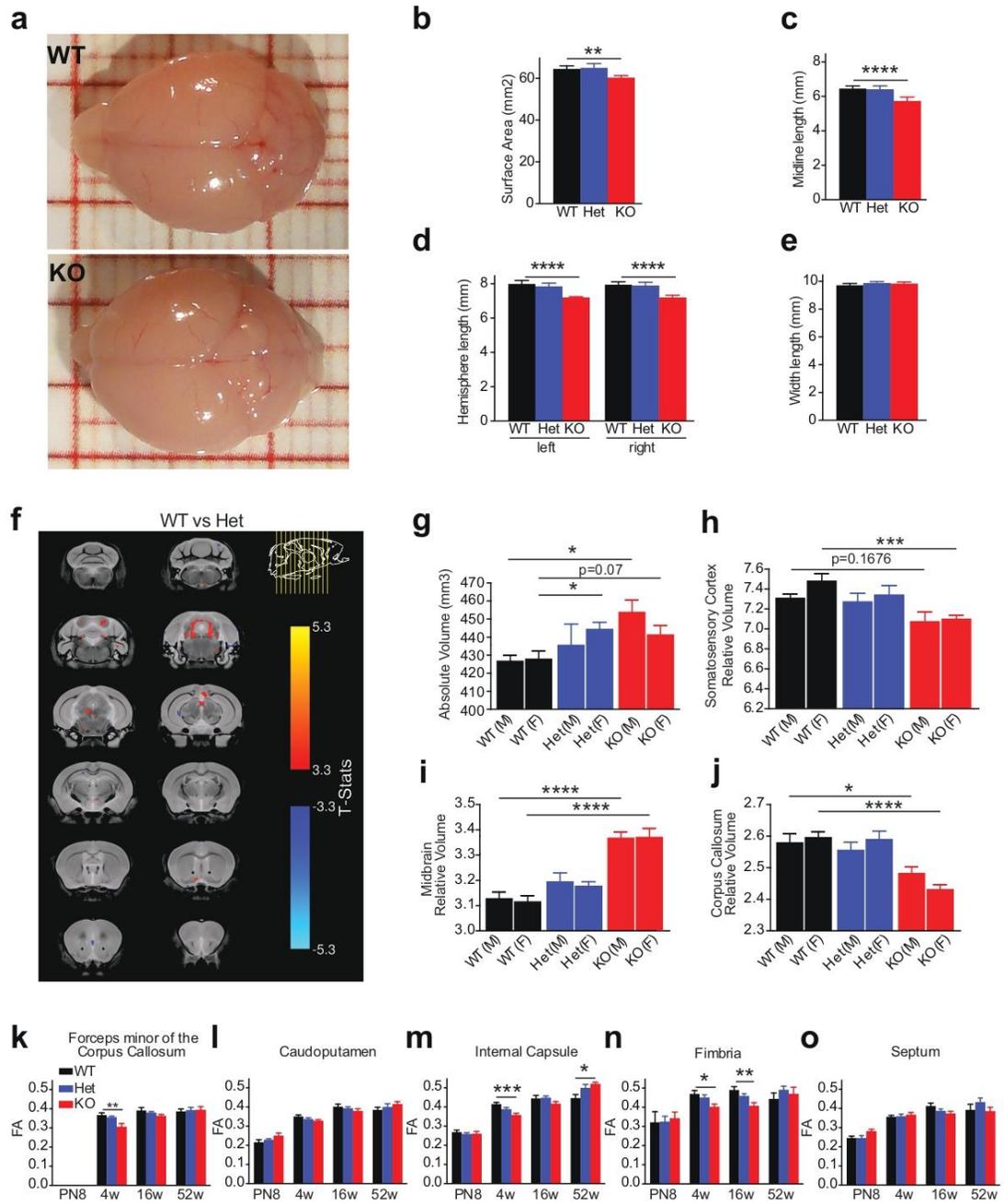
Supplementary Figure 9e (top): WT vs KO, *p for 65 μm .

1. Kapfhamer D, Taylor S, Zou ME, Lim JP, Kharazia V, Heberlein U. Taok2 controls behavioral response to ethanol in mice. *Genes, brain, and behavior*. 2013 Feb;12(1):87-97. PubMed PMID: 22883308.
2. Wiegert JS, Oertner TG. Long-term depression triggers the selective elimination of weakly integrated synapses. *Proc Natl Acad Sci U S A*. 2013 Nov 19;110(47):E4510-9. PubMed PMID: 24191047. Pubmed Central PMCID: 3839749.
3. de Anda FC, Rosario AL, Durak O, Tran T, Graff J, Meletis K, et al. Autism spectrum disorder susceptibility gene TAOK2 affects basal dendrite formation in the neocortex. *Nat Neurosci*. 2012 Jun 10;15(7):1022-31. PubMed PMID: 22683681. Pubmed Central PMCID: 4017029.
4. Riedl J, Crevenna AH, Kessenbrock K, Yu JH, Neukirchen D, Bista M, et al. Lifeact: a versatile marker to visualize F-actin. *Nature methods*. 2008 Jul;5(7):605-7. PubMed PMID: 18536722. Pubmed Central PMCID: 2814344.
5. Riedl J, Flynn KC, Raducanu A, Gartner F, Beck G, Bosl M, et al. Lifeact mice for studying F-actin dynamics. *Nature methods*. 2010 Mar;7(3):168-9. PubMed PMID: 20195247.
6. Noren NK, Pasquale EB. Eph receptor-ephrin bidirectional signals that target Ras and Rho proteins. *Cellular signalling*. 2004 Jun;16(6):655-66. PubMed PMID: 15093606.

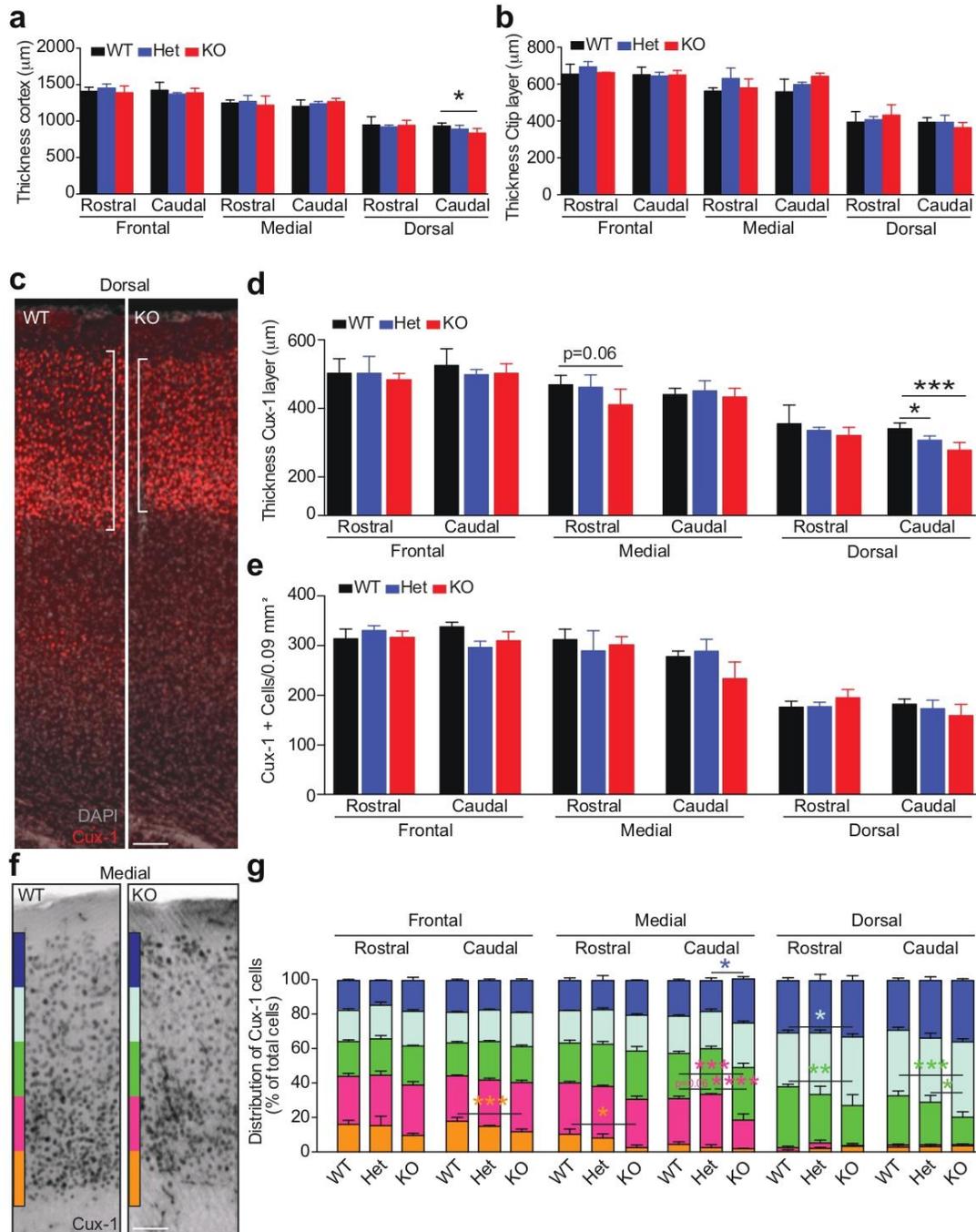
7. Bock M, Umatham R, Zuehlsdorff S, Volz S, Fink C, Hallscheidt P, et al. Interventional magnetic resonance imaging: an alternative to image guidance with ionising radiation. *Radiation protection dosimetry*. 2005;117(1-3):74-8. PubMed PMID: 16464829.
8. Nieman BJ, Flenniken AM, Adamson SL, Henkelman RM, Sled JG. Anatomical phenotyping in the brain and skull of a mutant mouse by magnetic resonance imaging and computed tomography. *Physiological genomics*. 2006 Jan 12;24(2):154-62. PubMed PMID: 16410543.
9. Lerch JP, Pruessner J, Zijdenbos AP, Collins DL, Teipel SJ, Hampel H, et al. Automated cortical thickness measurements from MRI can accurately separate Alzheimer's patients from normal elderly controls. *Neurobiology of aging*. 2008 Jan;29(1):23-30. PubMed PMID: 17097767.
10. Genovese CR, Lazar NA, Nichols T. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*. 2002 Apr;15(4):870-8. PubMed PMID: 11906227.
11. Andersson JL, Sotiropoulos SN. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *NeuroImage*. 2016 Jan 15;125:1063-78. PubMed PMID: 26481672. Pubmed Central PMCID: 4692656.
12. Behrens TE, Woolrich MW, Jenkinson M, Johansen-Berg H, Nunes RG, Clare S, et al. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magnetic resonance in medicine*. 2003 Nov;50(5):1077-88. PubMed PMID: 14587019.
13. Rorden C, Karnath HO, Bonilha L. Improving lesion-symptom mapping. *Journal of cognitive neuroscience*. 2007 Jul;19(7):1081-8. PubMed PMID: 17583985.
14. Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*. 2002 Oct;17(2):825-41. PubMed PMID: 12377157.
15. Hartung H, Cichon N, De Feo V, Riemann S, Schildt S, Lindemann C, et al. From Shortage to Surge: A Developmental Switch in Hippocampal-Prefrontal Coupling in a Gene-Environment Model of Neuropsychiatric Disorders. *Cereb Cortex*. 2016 Sep 09. PubMed PMID: 27613435. Pubmed Central PMCID: 5066837.
16. Cichon NB, Denker M, Grun S, Hanganu-Opatz IL. Unsupervised classification of neocortical activity patterns in neonatal and pre-juvenile rodents. *Frontiers in neural circuits*. 2014;8:50. PubMed PMID: 24904296. Pubmed Central PMCID: 4034041.
17. Kwan V, Meka DP, White SH, Hung CL, Holzapfel NT, Walker S, et al. DIXDC1 Phosphorylation and Control of Dendritic Morphology Are Impaired by Rare Genetic Variants. *Cell reports*. 2016 Nov 08;17(7):1892-904. PubMed PMID: 27829159.
18. Freitag S, Schachner M, Morellini F. Behavioral alterations in mice deficient for the extracellular matrix glycoprotein tenascin-R. *Behavioural brain research*. 2003 Oct 17;145(1-2):189-207. PubMed PMID: 14529817.
19. Morellini F. Spatial memory tasks in rodents: what do they model? *Cell and tissue research*. 2013 Oct;354(1):273-86. PubMed PMID: 23793547.
20. Dail M, Richter M, Godement P, Pasquale EB. Eph receptors inactivate R-Ras through different mechanisms to achieve cell repulsion. *Journal of cell science*. 2006 Apr 01;119(Pt 7):1244-54. PubMed PMID: 16522685.

21. Richter M, Murai KK, Bourgin C, Pak DT, Pasquale EB. The EphA4 receptor regulates neuronal morphology through SPAR-mediated inactivation of Rap GTPases. *J Neurosci*. 2007 Dec 19;27(51):14205-15. PubMed PMID: 18094260.
22. Franklin KBJ, Paxinos, G. *The mouse brain in stereotaxic coordinates*. New York: Academic Press. 2001.
23. Koster JD, Leggewie B, Blechner C, Brandt N, Fester L, Rune G, et al. Inositol-1,4,5-trisphosphate-3-kinase-A controls morphology of hippocampal dendritic spines. *Cellular signalling*. 2016 Jan;28(1):83-90. PubMed PMID: 26519023.
24. Gray EG. Axo-somatic and axo-dendritic synapses of the cerebral cortex: an electron microscope study. *Journal of anatomy*. 1959 Oct;93:420-33. PubMed PMID: 13829103. Pubmed Central PMCID: 1244535.
25. Gray EG. Electron microscopy of synaptic contacts on dendrite spines of the cerebral cortex. *Nature*. 1959 Jun 06;183(4675):1592-3. PubMed PMID: 13666826.
26. DeFelipe J. The dendritic spine story: an intriguing process of discovery. *Front Neuroanat*. 2015;9:14. PubMed PMID: 25798090. Pubmed Central PMCID: 4350409.
27. Pellegrin S, Mellor H. Rho GTPase activation assays. *Current protocols in cell biology*. 2008 Mar;Chapter 14:Unit 14 8. PubMed PMID: 18360815.
28. Hui-Yuen J, McAllister S, Koganti S, Hill E, Bhaduri-McIntosh S. Establishment of Epstein-Barr virus growth-transformed lymphoblastoid cell lines. *Journal of visualized experiments : JoVE*. 2011 Nov 08(57). PubMed PMID: 22090023. Pubmed Central PMCID: 3308597.
29. Gage GJ, Kipke DR, Shain W. Whole animal perfusion fixation for rodents. *Journal of visualized experiments : JoVE*. 2012 Jul 30(65). PubMed PMID: 22871843. Pubmed Central PMCID: 3476408.
30. Phansalskar N, More, S., Sabale, A. . Adaptive local thresholding for detection of nuclei in diversity stained cytology images. *International Conference on Communications and Signal Processing (ICCS)*. 2011:218-20.
31. Longair MH, Baker DA, Armstrong JD. Simple Neurite Tracer: open source software for reconstruction, visualization and analysis of neuronal processes. *Bioinformatics*. 2011 Sep 01;27(17):2453-4. PubMed PMID: 21727141.
32. Ferreira TA, Blackman AV, Oyrer J, Jayabal S, Chung AJ, Watt AJ, et al. Neuronal morphometry directly from bitmap images. *Nature methods*. 2014 Oct;11(10):982-4. PubMed PMID: 25264773. Pubmed Central PMCID: 5271921.
33. Risher WC, Ustunkaya T, Singh Alvarado J, Eroglu C. Rapid Golgi analysis method for efficient and unbiased classification of dendritic spines. *PloS one*. 2014;9(9):e107591. PubMed PMID: 25208214. Pubmed Central PMCID: 4160288.
34. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. 2009.
35. Tseng Q, Duchemin-Pelletier E, Deshiere A, Balland M, Guillou H, Filhol O, et al. Spatial organization of the extracellular matrix regulates cell-cell junction positioning. *Proc Natl Acad Sci U S A*. 2012 Jan 31;109(5):1506-11. PubMed PMID: 22307605. Pubmed Central PMCID: 3277177.
36. Tarnok K, Gulyas M, Bencsik N, Ferenc K, Pfizenmaier K, Hausser A, et al. A new tool for the quantitative analysis of dendritic filopodial motility. *Cytometry Part A : the journal of the International Society for Analytical Cytology*. 2015 Jan;87(1):89-96. PubMed PMID: 25257846.

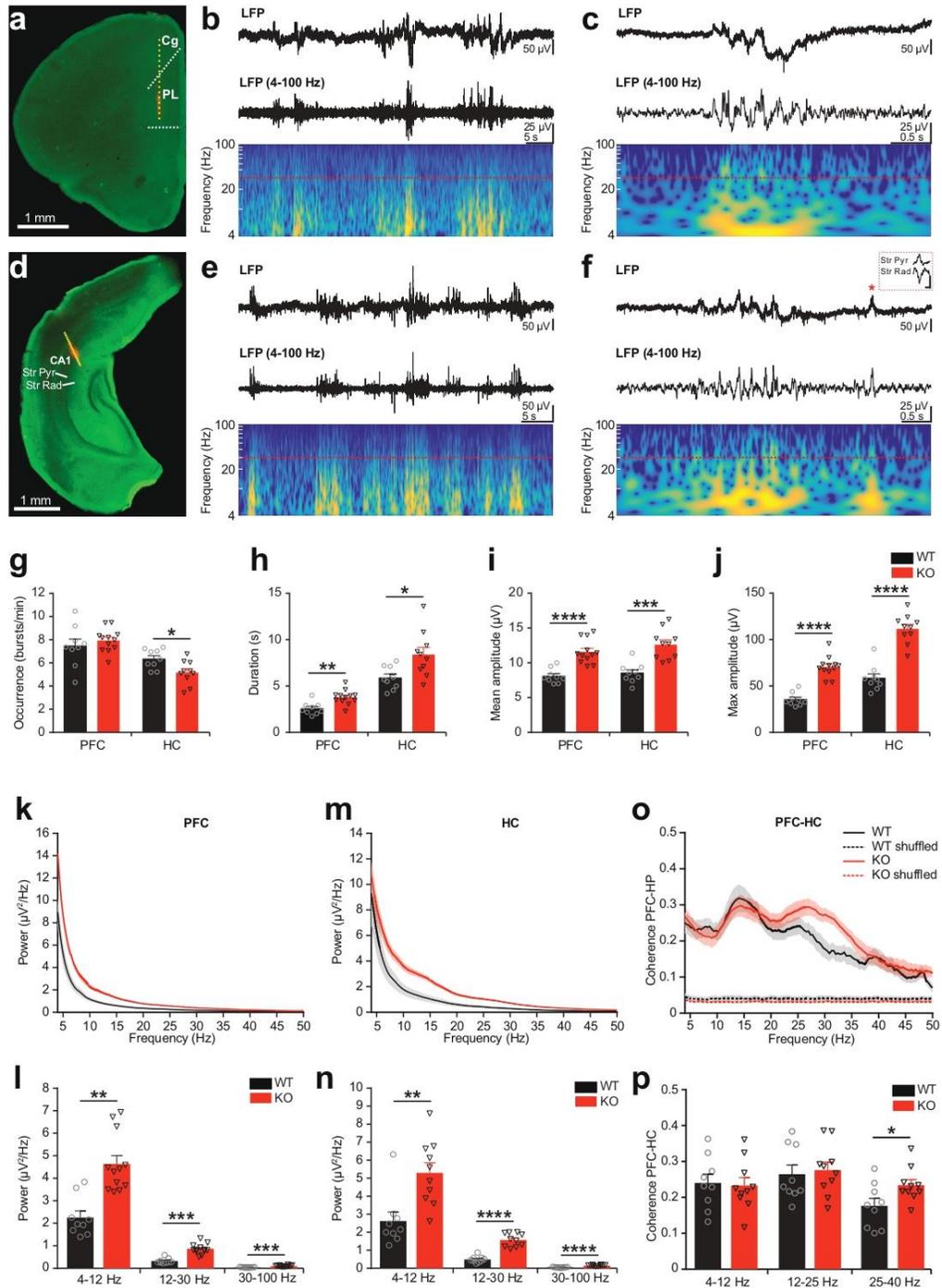
Supplementary Figure 1



Supplementary Figure 2



Supplementary Figure 3

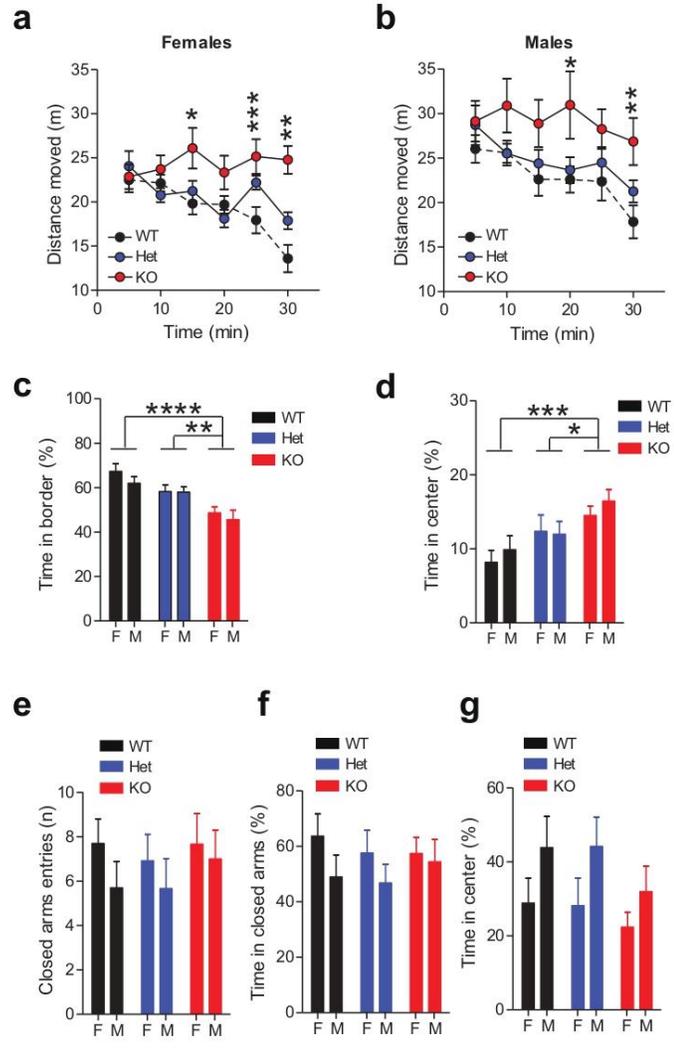


Supplementary Table 2

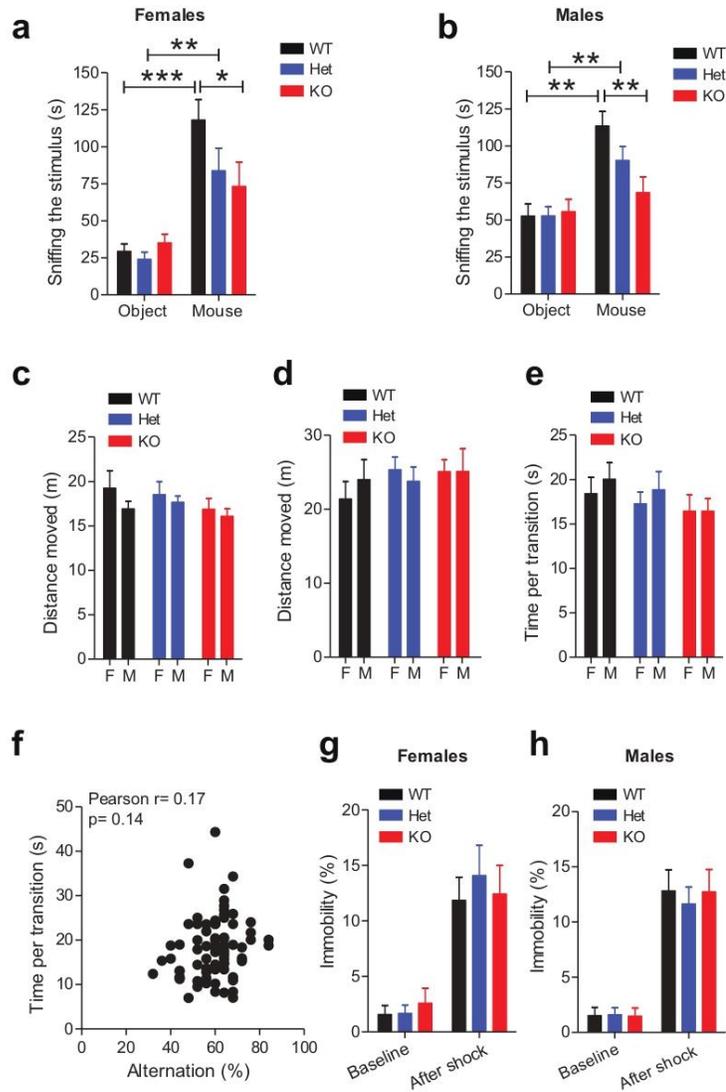
Mean prefrontal-hippocampal coherence in different frequency bands as well as properties of discontinuous patterns of oscillatory activity in PFC and HC of neonatal control and TAOK2 KO mice. Data are shown as mean \pm s.e.m. Significance levels for comparisons between controls and TAOK2 KO mice are listed.

		Controls (n=9)	TAOK2 KO (n=10-12)	p value
PFC-HC	Coherence (4-12 Hz)	0.240 \pm 0.025	0.233 \pm 0.022	0.838
	Coherence (12-25 Hz)	0.265 \pm 0.025	0.276 \pm 0.022	0.746
	Coherence (25-40 Hz)	0.176 \pm 0.021	0.234 \pm 0.015	0.035
PFC	Occurrence (bursts/min)	7.45 \pm 0.59	7.93 \pm 0.29	0.442
	Duration (s)	2.57 \pm 0.22	3.78 \pm 0.22	0.001
	Mean amplitude (μ V)	8.10 \pm 0.34	11.58 \pm 0.49	0.00003
	Maximum amplitude (μ V)	35.54 \pm 2.36	70.46 \pm 3.20	1.07 * 10 ⁻⁷
	Power 4-12 Hz (μ V ² /Hz)	2.26 \pm 0.29	4.63 \pm 0.37	0.001
	Power 12-30 Hz (μ V ² /Hz)	0.33 \pm 0.04	0.86 \pm 0.07	0.0002
	Power 30-100 Hz (μ V ² /Hz)	0.04 \pm 0.003	0.10 \pm 0.01	0.0001
HC	Occurrence (bursts/min)	6.34 \pm 0.28	5.15 \pm 0.33	0.014
	Duration (s)	5.88 \pm 0.41	8.39 \pm 0.78	0.013
	Mean amplitude (μ V)	8.52 \pm 0.47	12.61 \pm 0.68	0.0001
	Maximum amplitude (μ V)	58.39 \pm 4.61	111.52 \pm 4.93	5.02 * 10 ⁻⁷
	Power 4-12 Hz (μ V ² /Hz)	2.62 \pm 0.50	5.30 \pm 0.57	0.003
	Power 12-30 Hz (μ V ² /Hz)	0.49 \pm 0.06	1.57 \pm 0.12	3.99 * 10 ⁻⁷
	Power 30-100 Hz (μ V ² /Hz)	0.05 \pm 0.01	0.15 \pm 0.01	3.99 * 10 ⁻⁷

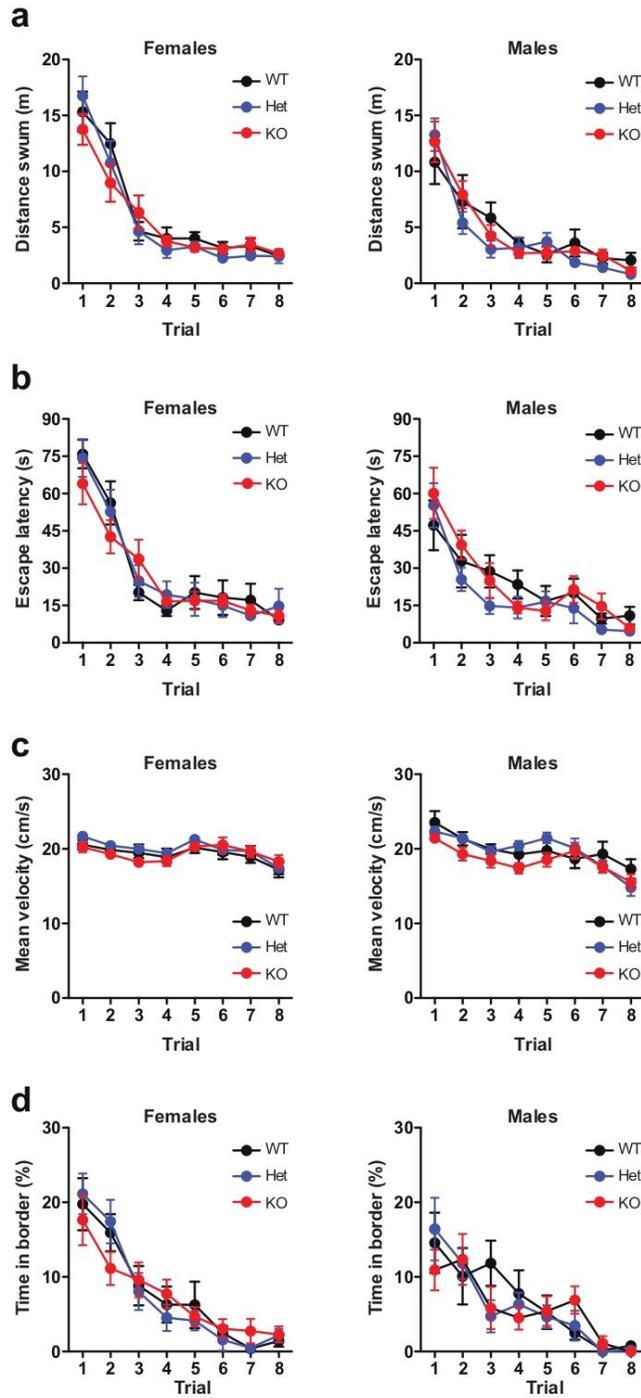
Supplementary Figure 4



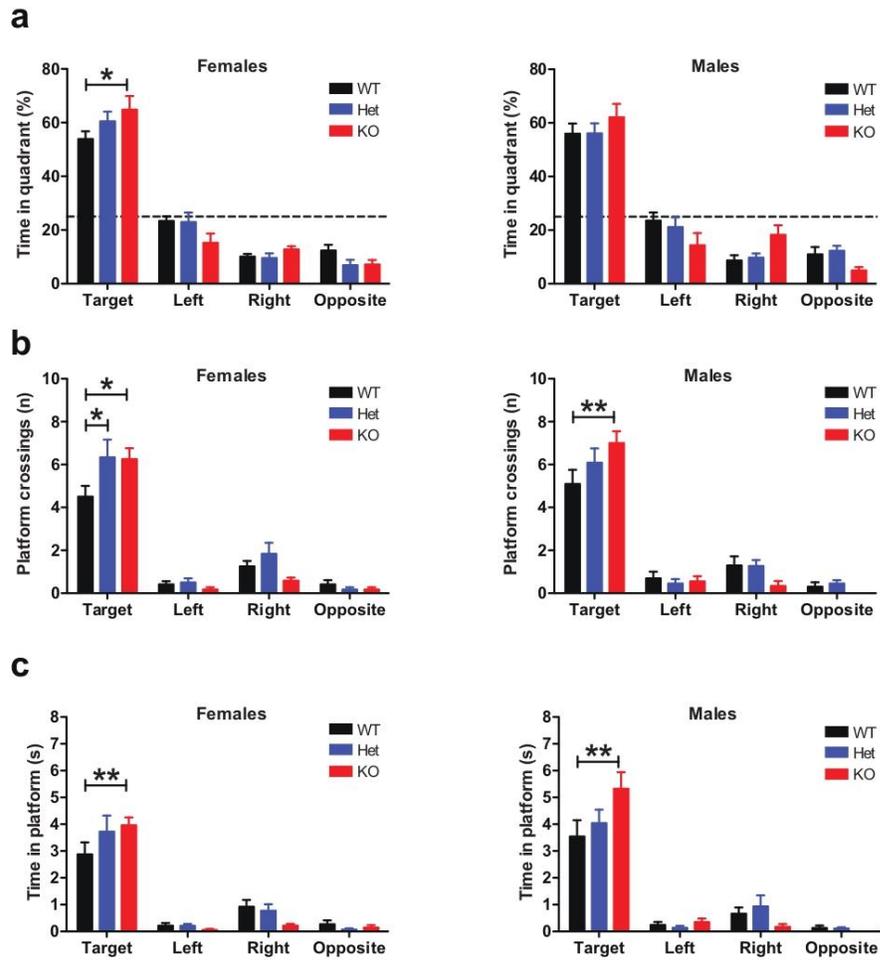
Supplementary Figure 5



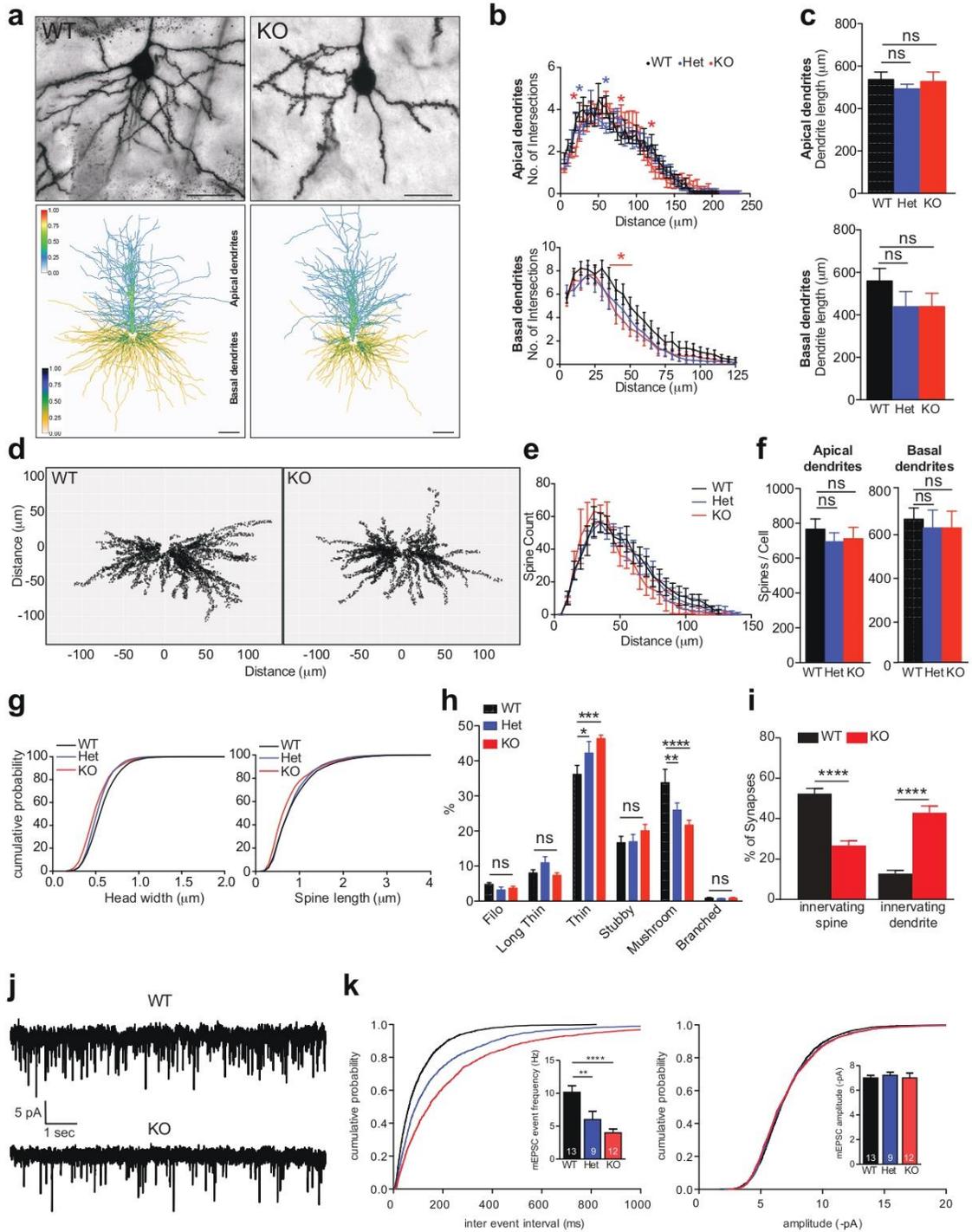
Supplementary Figure 6



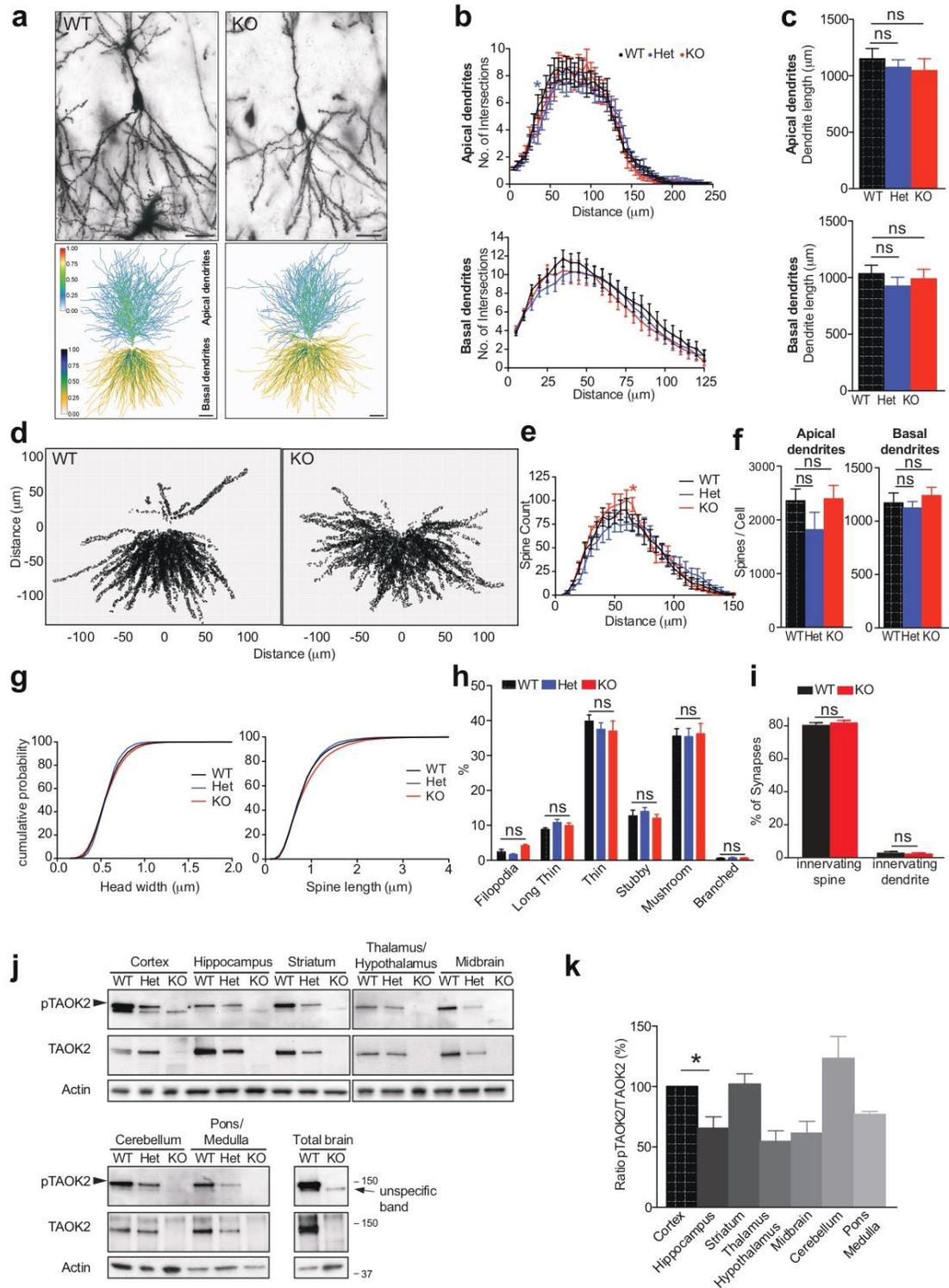
Supplementary Figure 7



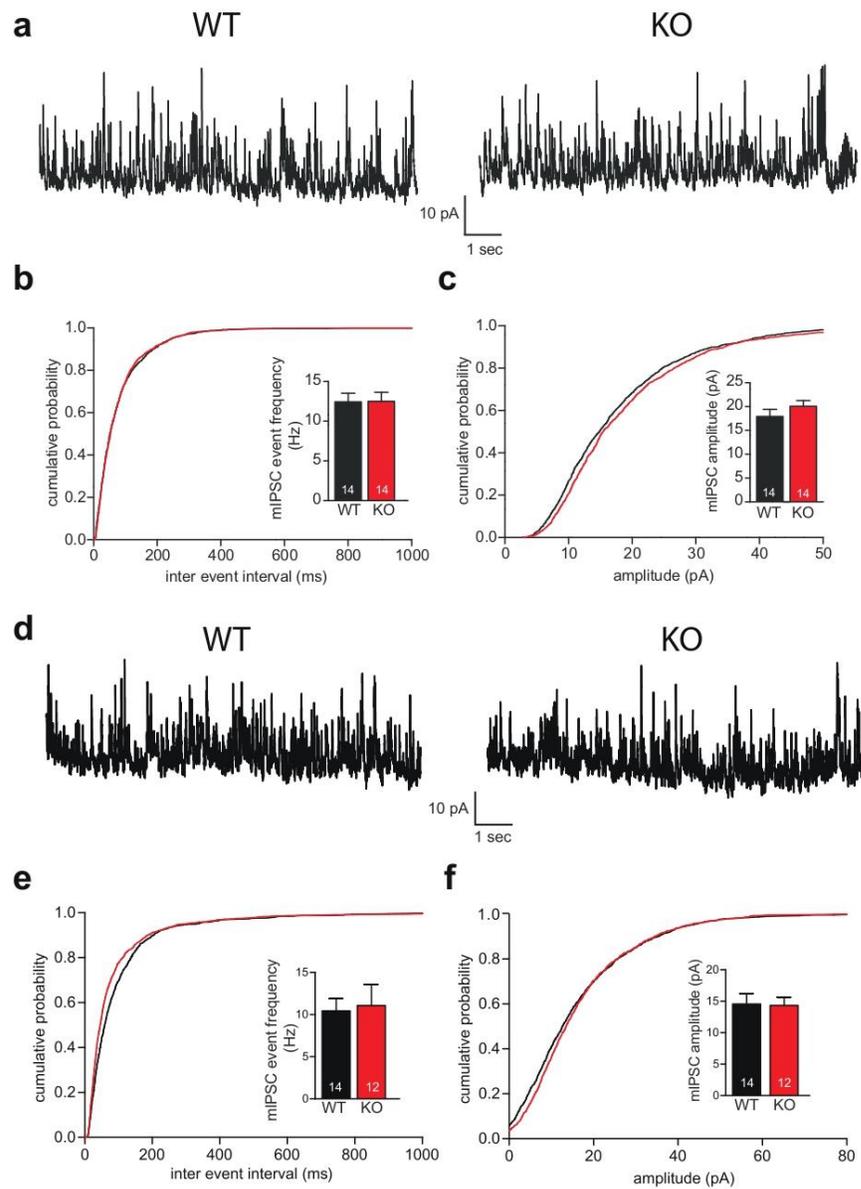
Supplementary Figure 8



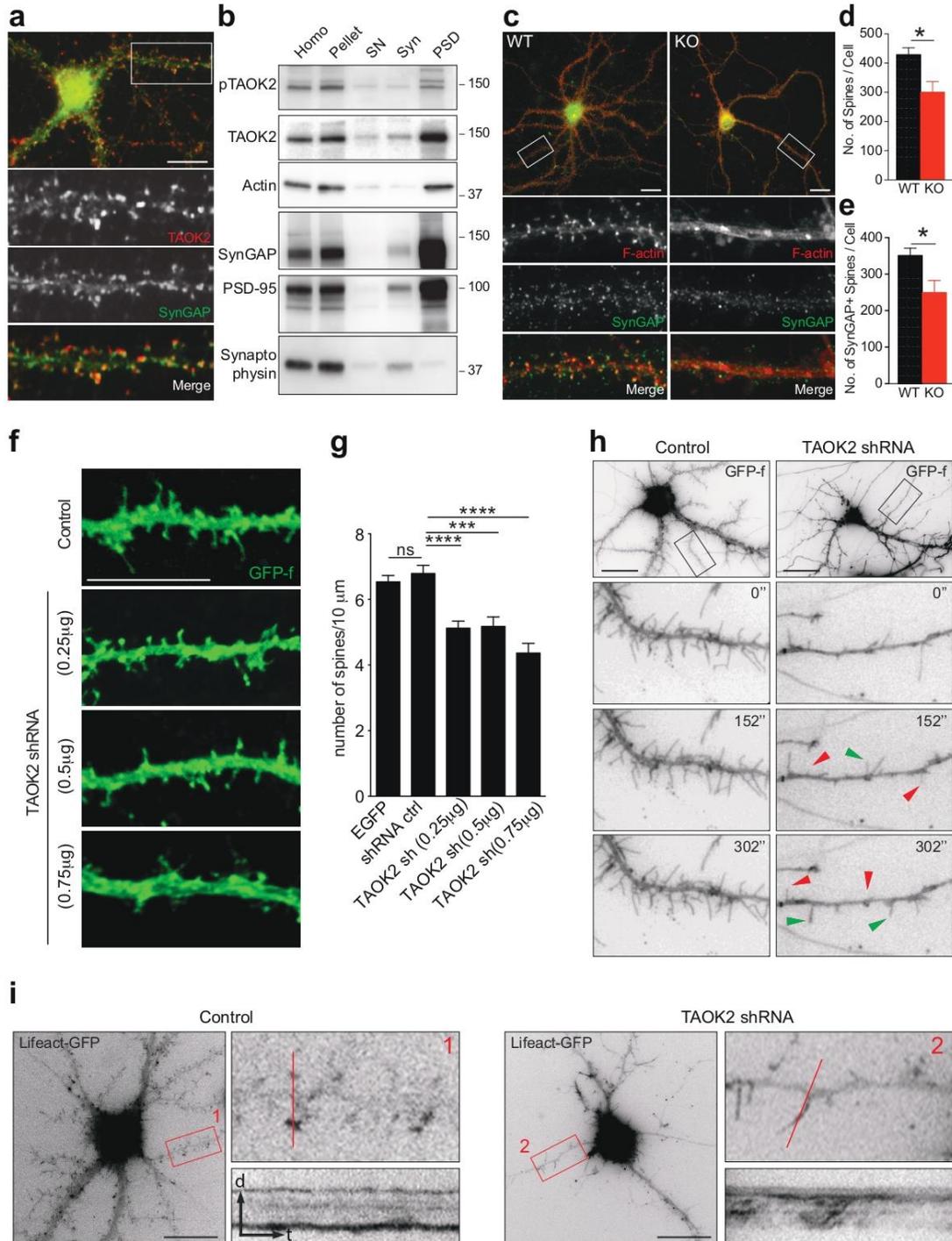
Supplementary Figure 9



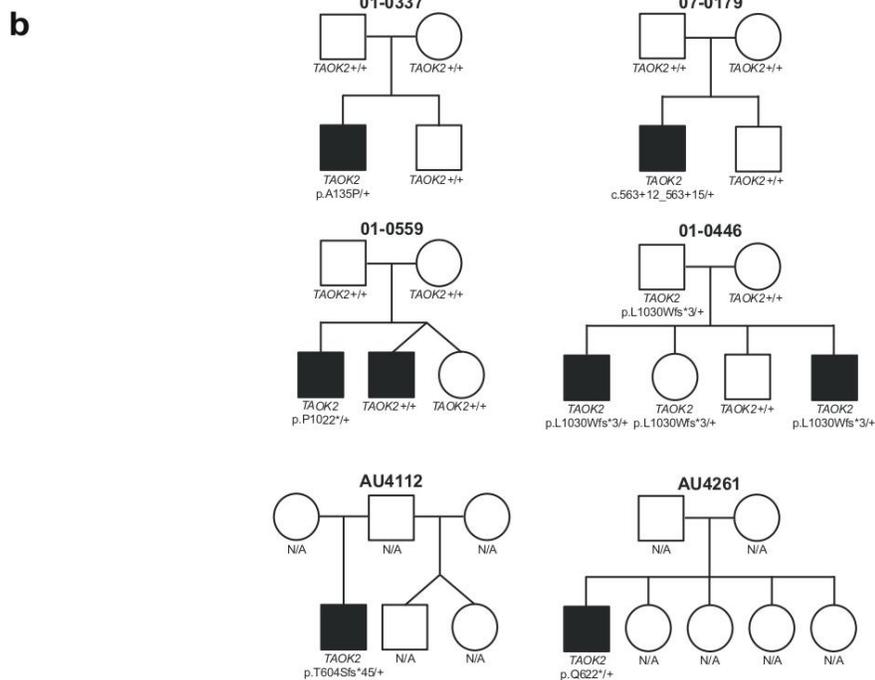
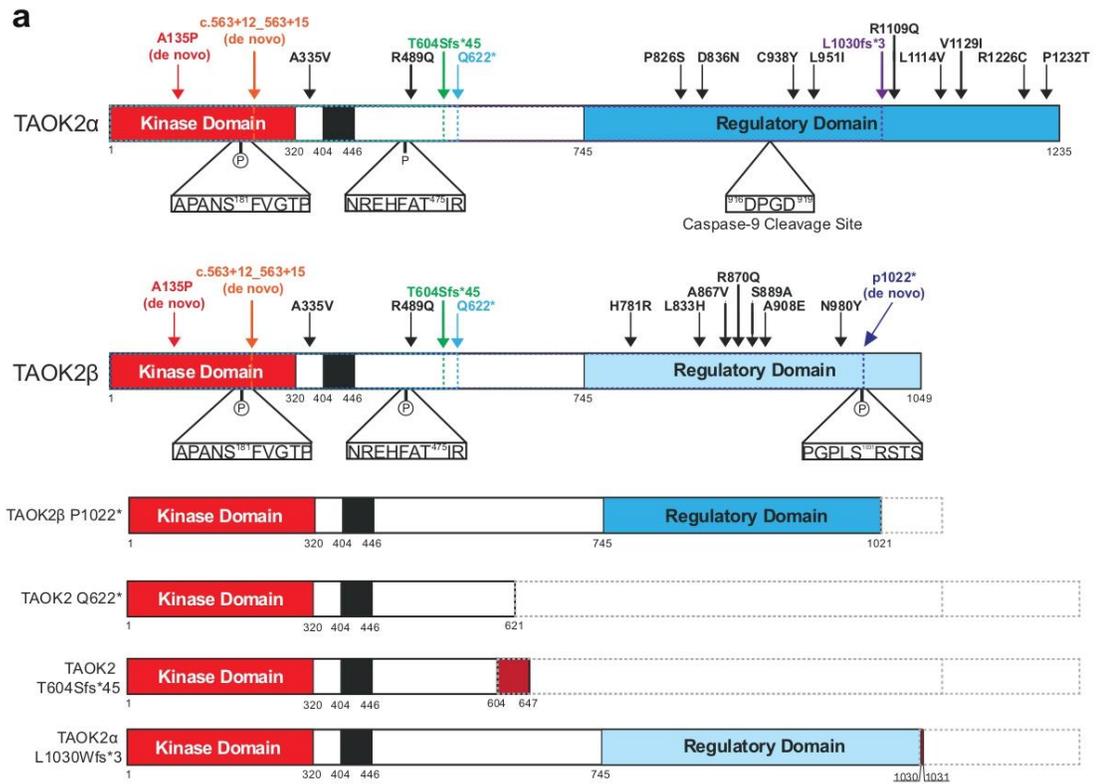
Supplementary Figure 10



Supplementary Figure 11



Supplementary Figure 12

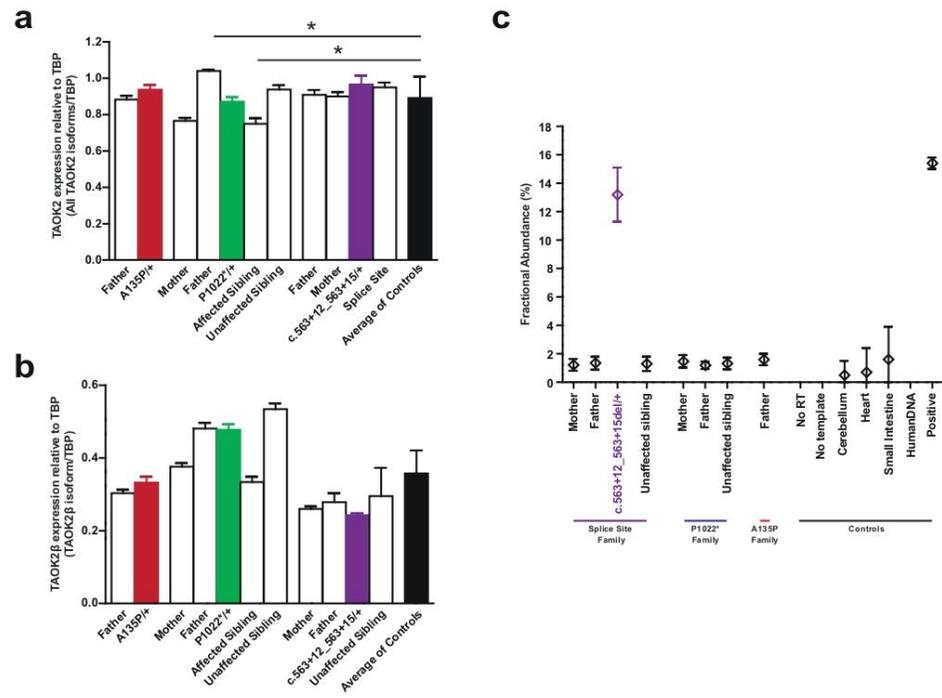


Supplementary Table 3

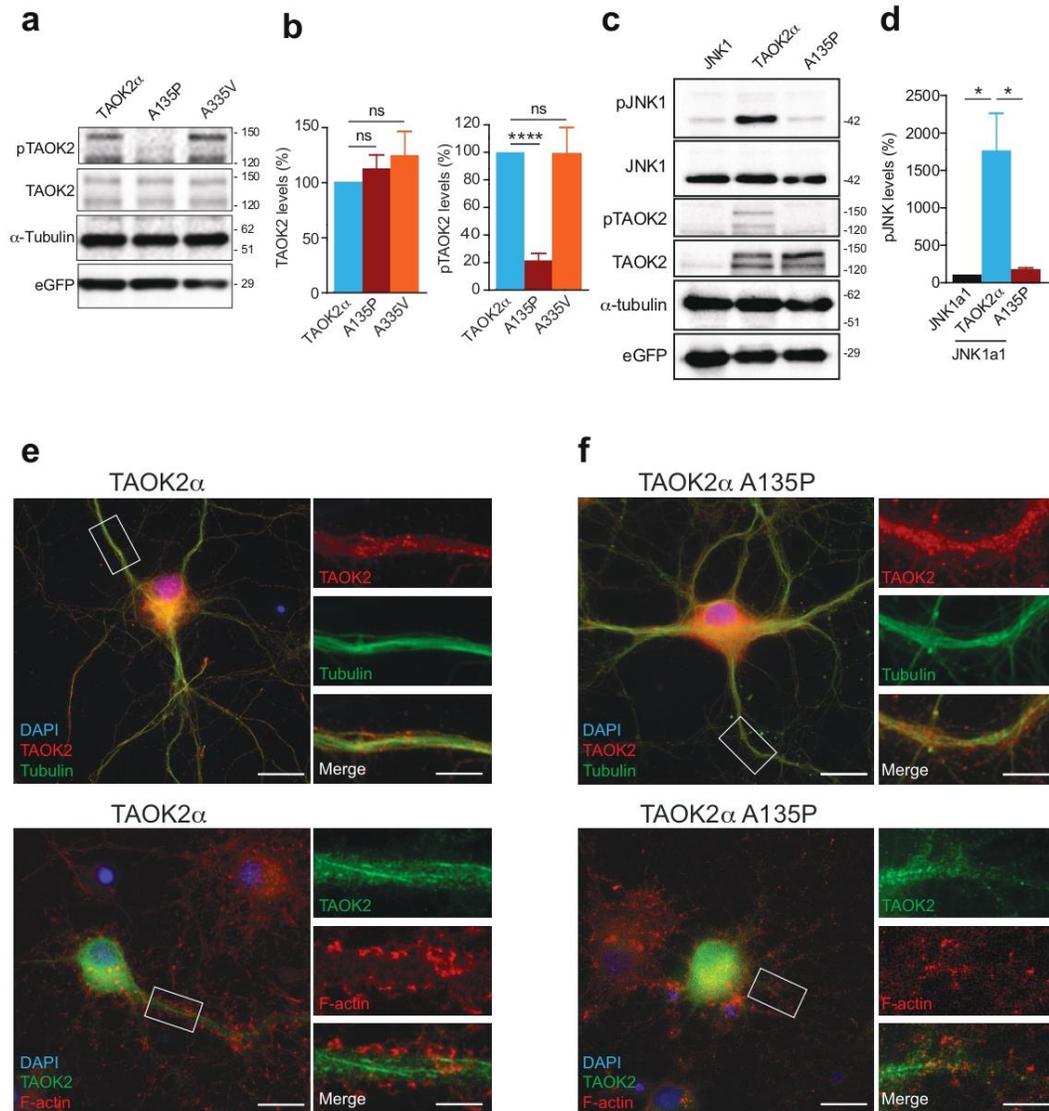
Position	Reference	Type	Genotype	RefSeq Effect	RefSeq Codon Change	Amino Acid Change	RefSeq Transcript	Affected Isoforms	Inheritance	Minor Allele Frequency (%)	Polyphen Score
29990345	G	HET	G/C	MISSENSE	Gcg/Ccg	A135P	NM_016151	Alpha Beta Gamma	de novo	-	0.97
30002803	GGGCCCCCTGC TGCCGGGTGCCC	HET	GGGCCCCCTGC TGCCGGGTGCCC/--	FRAMESHIFT DELETION	-	P1022*	NM_004783	Beta	de novo	-	N/A
29990629	GTGA	HET	GTGA/--	SPICE SITE DELETION	N/A	c.563+12_563+15	NC_000016.9	Alpha Beta Gamma	de novo	-	N/A
29996995	CCAGCACTCCAAG CGGAGAAGGCCG AATGCTGTCTCGG CAGAAGGACACT	HET	CCAGCACTCCAAG CGGAGAAGGCCG AATGCTGTCTCGG CAGAAGGACACT/--	FRAMESHIFT DELETION	-	T6045fs*45	NM_016151	Alpha Beta Gamma	Unknown	-	N/A
29997054	C	HET	C/T	NONSENSE	Cag/Tag	Q622*	NM_016151	Alpha Beta Gamma	Unknown	-	N/A
29998680	C	HET	C/--	FRAMESHIFT DELETION	-	L1030Wfs*3	NM_016151	Alpha Gamma	Paternal Present in unaffected and affected sibling	-	N/A
29994397	C	HET	C/T	MISSENSE	gCc/gTc	A335V	NM_016151	Alpha Beta Gamma	Paternal	-	.008
29996577	G	HET	G/A	MISSENSE	cGg/cAg	R489Q	NM_016151	Alpha Beta Gamma	Paternal from paternal grandmother Not present in affected sibling	-	.999
29998069	C	HET	C/T	MISSENSE	Ccc/Tcc	P826S	NM_016151	Alpha	Paternal	-	0.002
29998099	G	HET	G/A	MISSENSE	Gat/Aat	D836N	NM_016151	Alpha	Paternal	-	0.001
29998406	G	HET	G/A	MISSENSE	tGa/tAc	C938Y	NM_016151	Alpha Gamma	Paternal	-	0
29998444	C	HET	C/A	MISSENSE	Ctc/Atc	L951I	NM_016151	Alpha Gamma	Paternal Not present in unaffected sibling	-	.996
29998919	G	HET	G/A	MISSENSE	cGg/cAg	R1109Q	NM_016151	Alpha Gamma	Maternal	-	0.034
29998933	C	HET	C/G	MISSENSE	Ctg/Gtg	L1114V	NM_016151	Alpha Gamma	Maternal Present in affected sibling	-	0.546
29998978	G	HET	G/A	MISSENSE	Gtc/Atc	V1129I	NM_016151	Alpha Gamma	Maternal Present in affected sibling	-	0.002
29999269	C	HET	C/T	MISSENSE	Cgc/Tgc	R1226C	NM_016151	Alpha Gamma	Maternal Present in affected sibling	-	.988
29999287	C	HET	C/A	MISSENSE	Ccc/Acc	P1232T	NM_016151	Alpha Gamma	Maternal Present in affected sibling	-	.996
30001023	A	HET	A/G	MISSENSE	cAc/cGc	H781R	NM_004783	Beta	Maternal	0.0539%	.998
30002157	T	HET	T/A	MISSENSE	cTt/cAt	L833H	NM_004783	Beta	Unknown	-	1
30002259	C	HET	C/T	MISSENSE	gCg/gTg	A867V	NM_004783	Beta	Paternal	-	.07
30002288	G	HET	G/A	MISSENSE	cGg/cAg	R870Q	NM_004783	Beta	Maternal	0.04%	.998
30002404	T	HET	T/G	MISSENSE	Tcc/Gcc	S889A	NM_004783	Beta	Maternal	-	.311
30002462	C	HET	C/A	MISSENSE	gCg/gAg	A908E	NM_004783	Beta	Not Maternal No paternal DNA	-	.271
30002677	A	HET	A/T	MISSENSE	Aac/Tac	N980Y	NM_004783	Beta	Maternal	-	.056

red: de novo mutations; dark red: variants resulting in truncations; black: rare-inherited variants found in all isoforms; green: rare-inherited variants affecting alpha isoform only; blue: rare-inherited variants affecting alpha and gamma isoforms; purple: rare-inherited variants affecting beta isoform only

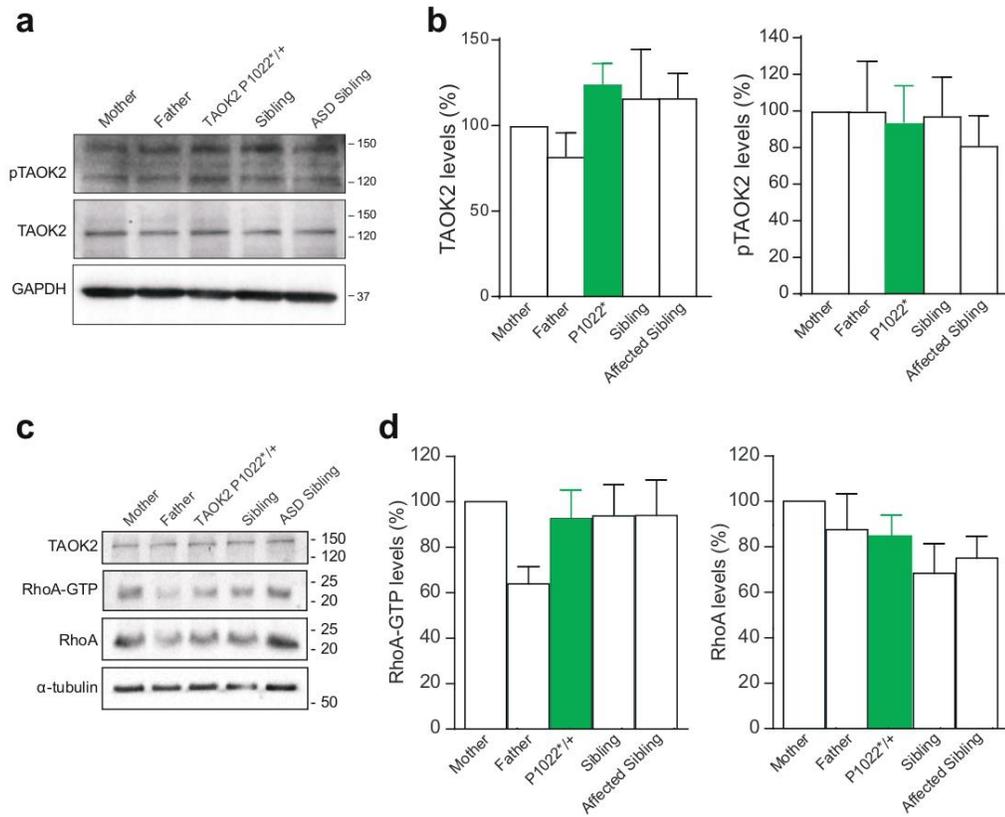
Supplementary Figure 13



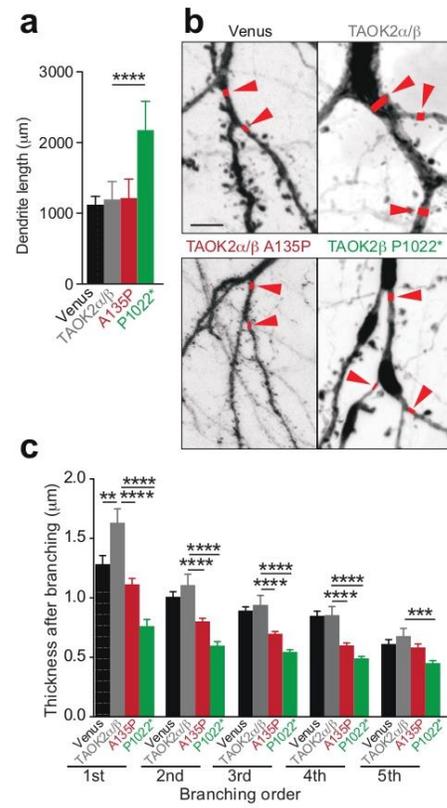
Supplementary Figure 14



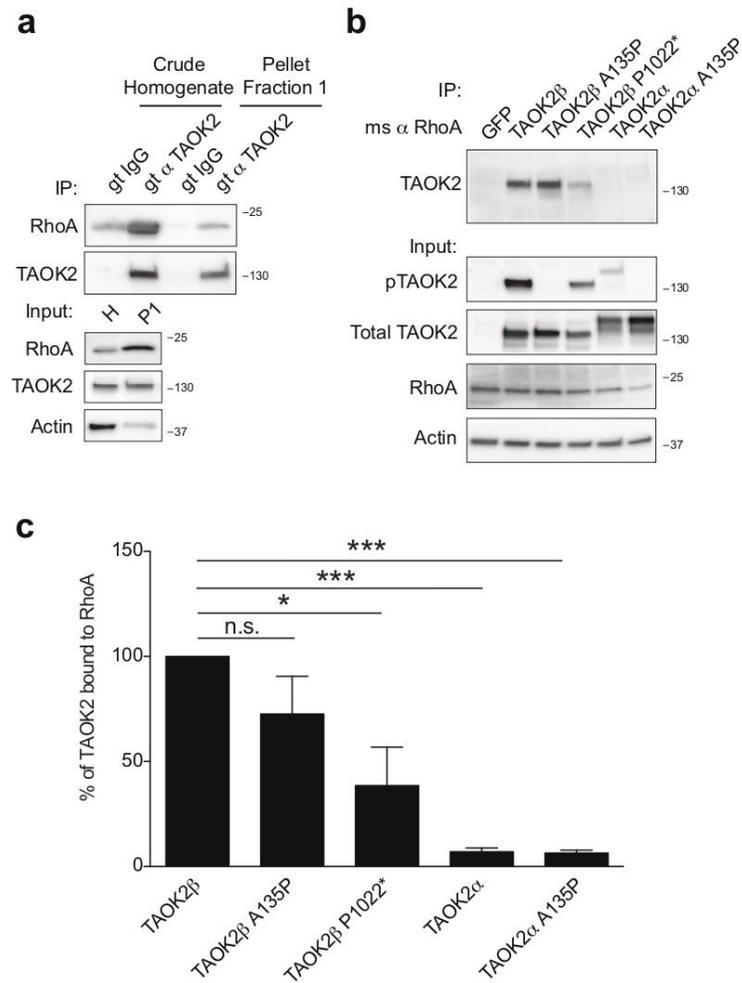
Supplementary Figure 15



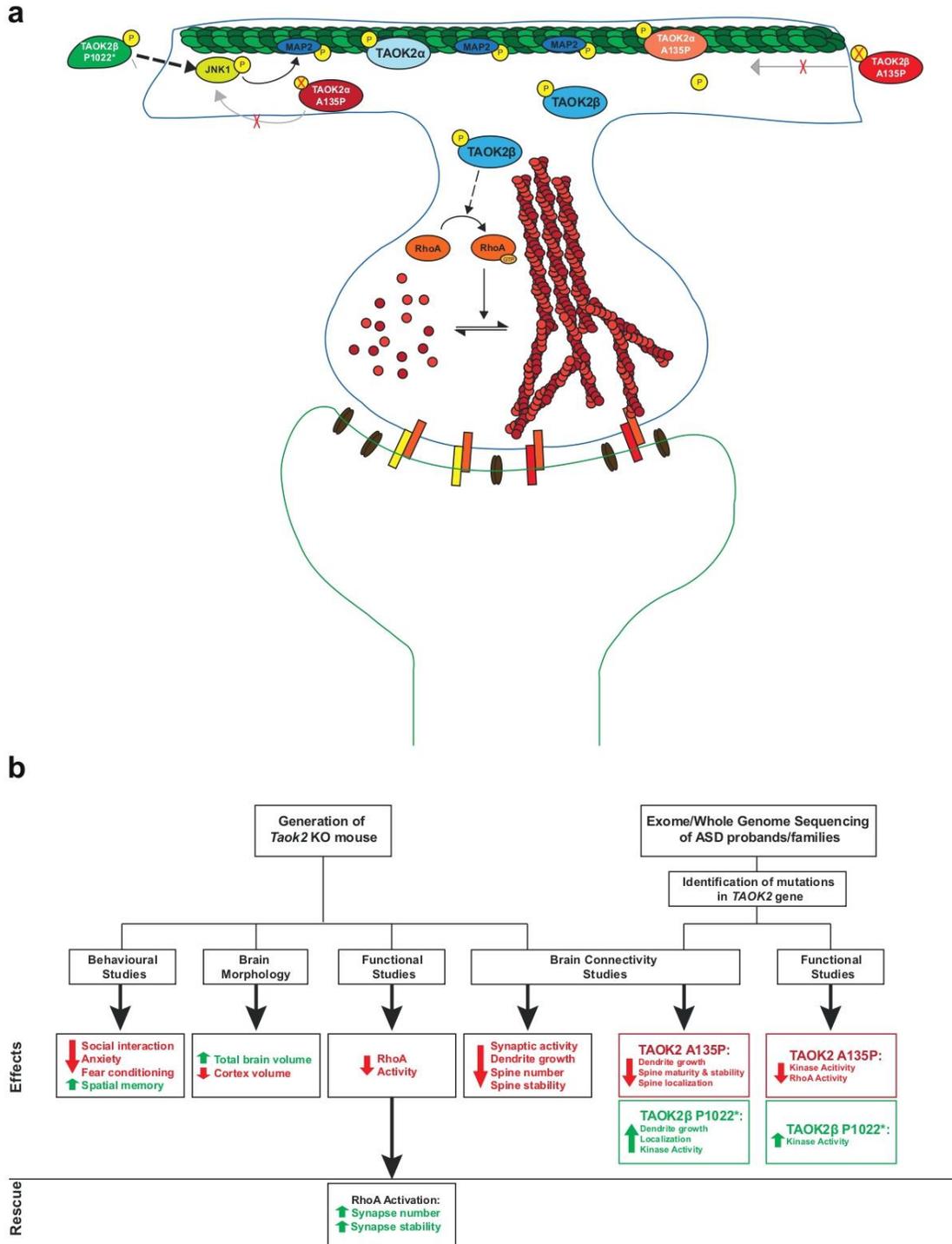
Supplementary Figure 16



Supplementary Figure 17



Supplementary Figure 18



Supplementary Figure 1. *Taok2* KO mice have altered brain morphology. (a) Top view of WT and *Taok2* KO mouse brains. (b-e) The cortex of *Taok2* KO mice have decreased surface area (mm²), midline length (mm) and bilateral hemisphere lengths (mm), with no change in width (WT=6, Het=7 and KO=5 brains from 3 different litters; Surface Area: one-way ANOVA, *post hoc* Dunnett's test; F(2,15)=14.49, p=0.0003; WT vs KO p=0.0010; Midline length: one-way ANOVA, *post hoc* Dunnett's test; F(2,15)=23.54, p<0.0001; WT vs KO p<0.0001; Left Hemisphere length: F(2,15)=32.01, p<0.0001; WT vs KO p<0.0001; Right Hemisphere length: F(2,15)=32.01, p<0.0001; WT vs KO p<0.0001; Width length: one-way ANOVA, *post hoc* Dunnett's test F(2,15)=2.942, p=0.0836). (f) Voxel-wise analysis highlighting significant relative volume differences (images show the lowest threshold of 20% false discovery rate for *Taok2* Het mice, FDR) throughout the brain between the WT and *Taok2* Het mice. T-stats of 3.3 to 5.3 indicates decreasing false discovery rate, where 3.3 = 20% FDR and positive or negative T-stat indicates positive or negative change compared to WT brain. (g) *Taok2* KO mice and female *Taok2* Het have increased absolute brain volume compared to WT mice (WT(M)=4, WT(F)=10, Het(M)=5, Het(F)=8, KO(M)=11, KO(F)=12 mice from 3 different cohorts, statistics by linear model corrected for multiple comparisons using FDR; WT(M) vs KO (M) p=0.0127, WT(F) vs KO (F) p=0.0711, WT(M) vs Het(M) p=0.4497, WT(F) vs Het(F) p=0.0167). (h) *Taok2* KO female mice have decreased relative volume of the somatosensory cortex compared to WT mice, with male mice also trending towards a decrease (WT(M)=4, WT(F)=10, Het(M)=5, Het(F)=8, KO(M)=11, KO(F)=12 mice from 3 different cohorts, statistics by linear model corrected for multiple comparisons using FDR; WT(M) vs KO (M) p=0.1678, WT(F) vs KO (F) p=0.0009, WT(M) vs Het(M) p=0.4348, WT(F) vs Het(F) p=0.4293). (i) *Taok2* KO male and female mice show an increase in relative volume of the midbrain compared to WT mice (WT(M)=4, WT(F)=10, Het(M)=5, Het(F)=8, KO(M)=11, KO(F)=12 mice from 3 different cohorts, statistics by linear model corrected for multiple comparisons using FDR; WT(M) vs KO (M) p<0.0001, WT(F) vs KO (F) p<0.0001, WT(M) vs Het(M) p=0.8045, WT(F) vs Het(F) p=0.5753). (j) The relative volume of the corpus callosum is reduced in male and female *Taok2* KO mice (WT(M)=4, WT(F)=10, Het(M)=5, Het(F)=8, KO(M)=11, KO(F)=12 mice from 3 different cohorts, statistics by linear model corrected for multiple comparisons using FDR; WT(M) vs KO (M) p=0.0278, WT(F) vs KO (F) p<0.0001, WT(M) vs Het(M) p=0.9171, WT(F) vs Het(F) p=0.9375). (k-o) Diffusion tensor imaging reveal delayed development of the forceps minor of

the corpus callosum, internal capsule and fimbria in *Taok2* KO mice (WT=5-7, Het=6-15 and KO=6-11 brains from 3 different cohorts; forceps minor of the corpus callosum: one-way ANOVA, *post hoc* Dunnett's test; 4wks: $F(2,29)=7.094$, $p=0.0031$; WT vs KO $p=0.0041$; internal capsule: one-way ANOVA, *post hoc* Dunnett's test; 4wks: $F(2,29)=8.225$, $p=0.0015$; WT vs KO $p=0.0008$; 52wks: $F(2,14)=5.377$, $p=0.0185$; WT vs KO $p=0.0117$; fimbria: one-way ANOVA, *post hoc* Dunnett's test; 4wks: $F(2,29)=4.832$, $p=0.0154$; WT vs KO $p=0.0131$; 16wks: $F(2,31)=6.113$, $p=0.0058$; WT vs KO $p=0.0044$). * $p<0.05$, ** $p<0.01$, *** $p<0.001$ and **** $p<0.0001$. Values are mean +/- s.e.m.

Supplementary Figure 2. Abnormal neuronal and cortical positioning after loss of *Taok2*

(a) Cortex thickness is decreased in *Taok2* KO compared with WT littermates (WT=5-9 brains, Het=3-6 brains, KO=2-6 brains; dorsal-caudal: one-way ANOVA, *post hoc* Dunnett's test $F(2,13)=4.312$, $p=0.0366$; WT vs KO $p=0.0219$). (b) Thickness of Ctip+ lower cortex is not altered in *Taok2* Het or KO cortices compared with WT littermates (WT=5-8 brains, Het=3-6 brains, KO=2-4 brains; one-way ANOVA, not significant in all regions). (c) Cux-1+ upper cortex is thinner in *Taok2* KO cortices. Scale bars represent 200 μ m. (d) Quantification of thickness of Cux-1+ upper cortex (WT=5-9 brains, Het=3-6 brains, KO=2-6 brains; medial-rostral: one-way ANOVA, *post hoc* Dunnett's test, $F(2, 12)=3.03$, $p=0.0861$; WT vs KO $p=0.0633$; $p=0.045$ by t-test; dorsal-caudal: one-way ANOVA, *post hoc* Dunnett's test, $F(2,13)=17.02$, $p=0.0002$; WT vs KO $p=0.0001$, WT vs Het $p=0.0115$). (e) Quantification of cell density of Cux-1+ neurons in the upper cortex (WT=5-9 brains, Het=3-5 brains, KO=2-4 brains; one-way ANOVA, not significant in all regions). (f) Cux-1+ cells differentially distribute in *Taok2* KO cortices with more Cux-1+ cells clustered in the medial to upper portion of the upper cortex in the medial-caudal and dorsal cortical region. Scale bars represent 20 μ m. (g) Quantification of Cux-1+ cells distribution in the binned upper cortex of frontal, medial and dorsal cortical regions (WT=5-8 brains, Het=3-6 brains, KO=2-5 brains; frontal-rostral: two-way ANOVA, *post hoc* Bonferroni's test, $F(2,35)=1.197e-14$, $p>0.9999$ between genotypes; frontal-caudal: two-way ANOVA, *post hoc* Bonferroni's test, $F(2,80)=4.657e-13$, $p>0.9999$ between genotypes; WT vs KO: orange $p=0.0004$; medial-rostral: two-way ANOVA, *post hoc* Bonferroni's test, $F(2,45)=4.252e-13$, $p>0.9999$ between genotypes; WT vs KO: orange $p=0.0156$; medial-caudal: two-way ANOVA, *post hoc* Bonferroni's test, $F(2,50)=0.01929$,

$p=0.9809$ between genotypes; WT vs KO: magenta $p=0.0003$, WT vs Het: magenta $p=0.0673$, Het vs KO: blue $p=0.0118$, magenta $p<0.0001$; dorsal-rostral: two-way ANOVA, *post hoc* Bonferroni's test, $F(2,40)=4.764e-13$, $p>0.9999$ between genotypes; WT vs KO: cyan $p=0.0287$, green $p=0.0012$; dorsal-caudal: two-way ANOVA, *post hoc* Bonferroni's test, $F(2,45)=8.032e-13$, $p>0.9999$ between genotypes; WT vs KO: green $p=0.0002$, Het vs KO: green $p=0.0177$. * $p<0.05$, ** $p<0.01$, *** $p<0.001$, and **** $p<0.0001$. Values are mean \pm s.e.m).

Supplementary Figure 3. Patterns of oscillatory activity in the prelimbic subdivision of PFC and hippocampal CA1 area of neonatal *Taok2* KO mice. (a) Digital photomontage reconstructing the location of the DiI-labelled recording electrode (orange) in PFC. The superimposed yellow dots mark the 16 recording sites covering the prefrontal subdivisions, cingulate cortex (Cg) and PL. (b) Extracellular LFP recording of discontinuous oscillatory activity in PL from a P7 WT mouse displayed unfiltered (top) and after band-pass (4-100 Hz) filtering (middle). Bottom: traces are accompanied by the color-coded wavelet spectra of the LFP at identical time scale. The red dotted line indicates the lower border of the gamma frequency range (30 Hz). (c) Characteristic theta burst unfiltered (top) and after band-pass (4-100 Hz) filtering (middle). Bottom: Color-coded frequency plots show the wavelet spectra at identical time scale. Note the presence of gamma rhythms superimposed on the underlying theta rhythm. (d) Same as (a) for recording electrode in hippocampal CA1 of a P9 WT mouse. The superimposed yellow dots mark the 16 recording sites covering stratum pyramidale (Str Pyr) and radiatum (Str Rad). (e) Same as (b) for hippocampal LFP from a P8 WT mouse. (f) Characteristic theta burst with sharp wave (red asterisk) displayed unfiltered and after band-pass (4-100 Hz) filtering accompanied by the color-coded frequency plot at identical time scale. Inset, sharp wave reversing between Str Pyr and Str Rad and displayed at a larger time scale. Scale bars for inset correspond to 50 μ V and 0.05s. Bar diagrams illustrating the occurrence (g), duration (h), mean amplitude (i) and maximum amplitude (j) of oscillatory events in PFC and HC of WT and *Taok2* KO mice (WT=9 and KO=10-12 mice per group, unpaired t-test, PFC: occurrence $t(19)=-0.785$, $P=0.442$; duration $t(19)=-3.736$, $P=0.001$; mean amplitude $t(19)=-5.393$, $P=0.00003$; maximum amplitude $t(19)=-8.245$, $P=1.07*10^{-7}$, HC: occurrence $t(17)=2.725$, $P=0.014$; duration $t(17)=-2.759$, $P=0.013$; mean amplitude $t(17)=-4.829$, $P=0.0001$; maximum amplitude $t(17)=-7.813$, $P=5.02*10^{-7}$). (k) Power spectra of discontinuous oscillations in

neonatal PFC when averaged for all WT controls. (l) *Taok2* KO mice have increased power in discontinuous oscillations in theta (4-12Hz), beta (12-30Hz) and gamma (30-100Hz) frequency ranges in PFC (WT=9 and KO=12 mice per group; Shapiro-Wilk test; 4-12 Hz $Z=-3.269$, $P=0.001$; 12-30 Hz $Z=-3.695$, $P=0.0002$; 30-100 Hz $Z=-3.838$, $P=0.0001$). (m) Power spectra of discontinuous oscillations in neonatal HC when averaged for all WT controls. (n) *Taok2* KO mice have increased power in discontinuous oscillations in the PFC in theta (4-12Hz), beta (12-30Hz) and gamma (30-100Hz) frequency ranges (WT=9 and KO=10 mice per group, Shapiro-Wilk test 4-12 Hz $Z=-2.939$, $P=0.003$; unpaired t-test 12-30 Hz $t(17)=-7.948$, $P=3.99 \times 10^{-7}$, 30-100 Hz $t(17)=-7.948$, $P=3.99 \times 10^{-7}$). * $p<0.05$, ** $p<0.01$, *** $p<0.001$, and **** $p<0.0001$. Values are mean \pm s.e.m. (o) Coherence spectra for co-occurring discontinuous network oscillations in neonatal PL and HC when averaged for all WT controls. (p) *Taok2* KO mice have increased coherence between the PFC and HC in the beta frequency (25-40Hz) (WT=9 and KO=10 mice per group; unpaired t-test: $t(17)=-2.297$, $P=0.035$). * $p<0.05$, ** $p<0.01$, and *** $p<0.001$. Values are mean \pm s.e.m.

Supplementary Figure 4. Hyperactivity and locomotor activity control tests in open-field and anxiety-related behavior for *Taok2* Het and KO mice. (a, b) Distance moved during the open field test displayed in time bins of 5 minutes for female (a) and male (b) mice, respectively, as compared to WT mice within the same time bin. (WT(F)=16, WT(M)=12, Het(F)=17, Het(M)=12, KO(F)=17, KO(M)=9 mice from three different cohorts; three-way mixed ANOVA, *post hoc* Bonferroni; effect of the interaction “genotype x time bin” $F(10,385)=2.91$, $p=0.002$; Females: WT vs KO: 15min $p<0.05$, 25min $p<0.001$, 30min $p<0.01$; Males: WT vs KO: 20min $p<0.05$, 30min $p<0.01$). (c) Time spent in a 5 cm border of the open field. (WT(F)=16, WT(M)=12, Het(F)=17, Het(M)=12, KO(F)=16, KO(M)=10 mice from three different cohorts; two-way ANOVA, *post hoc* Bonferroni’s test, effect of genotype $F(2,77)=14.43$, $p<0.0001$; WT vs KO $p<0.0001$, Het vs KO $p=0.0027$). (d) Time spent in the center (25 cm x 25 cm) of the open field. (WT(F)=16, WT(M)=12, Het(F)=17, Het(M)=11, KO(F)=17, KO(M)=11 mice from three different cohorts; two-way ANOVA, *post hoc* Bonferroni’s test, effect of genotype $F(2,77)=6.343$, $p=0.0028$; WT vs KO $p<0.0001$, Het vs KO $p=0.0027$). (e-g) Number of closed arm entries (e), time spent in the closed arms (f) and time spent in the center (g) in the elevated plus maze test was not affected by genotype. (WT(F)=10, WT(M)=10, Het(F)=12, Het(M)=12,

KO(F)=9, KO(M)=9 mice from three different cohorts; two-way ANOVA, *post hoc* Bonferroni's test, effect of genotype (e) $F(2,56)=0.3404$, $p=0.7130$ (f) $F(2,56)=0.2884$, $p=0.7505$ (g) $F(2,56)=0.9611$, $p=0.3887$ showing no differences in the number of closed arm entries or time spent in the closed arm or time spent in the center). ns >0.05 , * $p<0.05$, ** $p<0.01$, *** $p<0.001$ and **** $p<0.0001$. Values are mean \pm s.e.m.

Supplementary Figure 5. Gender control in sociability and locomotor activity controls in social, working and spatial memory and fear learning paradigms for *Taok2* Het and KO mice. (a, b) Time spent sniffing the unfamiliar mouse and novel object the social preference test by female (a) and male (b) mice. (WT(F)=11, WT(M)=11, Het(F)=8, Het(M)=15, KO(F)=12, KO(M)=12 mice from three different cohorts; three-way mixed ANOVA, *post hoc* Bonferroni; effect of the interaction “genotype x stimulus” $F(2,63)=4.96$, $p=0.001$; Females: WT vs KO: Mouse $p<0.05$, WT: Object vs Mouse $p<0.001$, Het: Object vs Mouse $p<0.01$; Males: WT vs KO: Mouse $p<0.01$, WT: Object vs Mouse $p<0.01$, Het: Object vs Mouse $p<0.01$). (c) Distance moved during the social preference test was not affected by genotype. (WT(F)=11, WT(M)=11, Het(F)=11, Het(M)=15, KO(F)=15, KO(M)=10 mice from three different litters; two-way ANOVA, *post hoc* Bonferroni's test, effect of genotype $F(2,67)=0.4582$, $p=0.6344$). (d) Distance moved during the recall trial of the spatial object recognition test was not affected by genotype. (WT(F)=11, WT(M)=11, Het(F)=11, Het(M)=15, KO(F)=15, KO(M)=10 mice from three different litters; two-way ANOVA, *post hoc* Bonferroni's test, effect of genotype $F(2,67)=0.4396$, $p=0.6462$). (e) No effects of genotype was detected for the time required to do a transition during the spontaneous alternation test in the Y-maze, although average values indicate that KO tended to be faster than WT. (WT(F)=11, WT(M)=11, Het(F)=15, Het(M)=13, KO(F)=16, KO(M)=8 mice from three different cohorts; two-way ANOVA, *post hoc* Bonferroni's test, effect of genotype $F(2,68)=0.1177$, $p=0.8891$). (f) No correlation was detected between time required to do a transition and percentage of alternation in the spontaneous alternation test in the Y-maze. (g, h) No difference between genotype was detected in their time spent immobile during baseline and conditioning trial of the contextual fear conditioning test as tested for female (g) and male mice (h). (WT(F)=9, WT(M)=8, Het(F)=9, Het(M)=9, KO(F)=9, KO(M)=8 mice from three different cohorts; three-way mixed ANOVA, *post hoc* Bonferroni,

with no difference in the time spent immobile during the baseline and conditioning trial. ns >0.05, *p<0.05, **p<0.01 and ***p<0.001. Values are mean ± s.e.m.

Supplementary Figure 6. Learning curves during training in the water maze test of WT and *Taok2* Het and KO mice. Mice were trained for eight trials over two days (4 trials per day, inter trial interval of 10 min) to find a hidden platform in a circular water maze. No effect of genotype, nor of the interaction “genotype x sex”, “genotype x trial” or “genotype x sex x trial” was detected as measured with genotype and sex as between groups factor and trial as within group factor. (a) Distance swum to reach the platform by female (left panel) and male (right panel) mice. (b) Escape latency of female (left panel) and male (right panel) mice. (c) Mean swimming velocity of female (left panel) and male (right panel) mice. (d) Time spent in the border (a 10 cm annulus adjacent to the pool wall) of female (left panel) and male (right panel) mice. (WT(F)=12, WT(M)=10, Het(F)=12, Het(M)=11, KO(F)=12, KO(M)=12 mice from three different cohorts; three-way mixed ANOVA, *post hoc* Bonferroni; effect of the interaction “genotype x trial” showed no difference in the learning curves with no difference in total distance swum, mean velocity, and time spent at the border. ns p>0.05. Values are mean ± s.e.m.

Supplementary Figure 7. *Taok2* deficient mice show enhanced preference for platform position in Morris water maze. Tests were conducted during the transfer trial performed 24 hours after the last training session. (a) Percentage of time spent in the four quadrants for female (left panel) and male (right panel) mice. Dotted lines indicate chance level (25%). (b) Number of crossings of imaginary platforms equally located in the center of each quadrant for female (left panel) and male (right panel) mice. (c). Time spent in imaginary platforms equally located in the center of each quadrant for female (left panel) and male (right panel) mice. (WT(M)=10, Het(F)=12, Het(M)=11, KO(F)=12, KO(M)=9 mice from three different cohorts; three-way mixed ANOVA, *post hoc* Bonferroni; effect of the interaction “genotype x position” (a) F(6,180)=3.3, p=0.0042,; Females/Target: WT vs KO p<0.05 (b) F(6,180)=4.9105, p=0.0001; Females/Target: WT vs KO p<0.05, WT vs Het p<0.05, Males/Target: WT vs KO p<0.01 (c) F(6,180)=4.495, p=0.0003; Females/Target: WT vs KO p<0.01, Males/Target: WT vs KO p<0.01 showing significant effects of the interaction “genotype x position” with genotype and

sex as between groups factors and quadrant or platform position as within groups factor. ns>0.05, *p<0.05 and **p<0.01. Values are mean ± s.e.m.

Supplementary Figure 8. *Taok2* Het and KO mice have milder reduction in dendrite growth and synaptic connectivity in the somatosensory cortex. (a) Top: Golgi-stained SSC neurons from P21 WT and *Taok2* Het and KO mice. Scale bars represent 20µm. Bottom: Dendritic heat maps of superimposed neuron tracings for each condition. Blue to red (apical) and yellow to blue (basal) indicates increased probability of dendrite presence. . Scale bars represent 20µm. (b) Top: No major difference in apical dendritic complexity in in layer 2 SSC neurons in *Taok2* Het and KO mice (WT=16, Het=17 and KO=15 neurons from 3 different brains; two-way ANOVA, *post hoc* Dunnett's test; $F(2,2115)=0.8267$, $p=0.4376$ between genotypes; *represents ranges of significance; WT vs Het (blue), WT vs KO (red); see supplemental statistics). Bottom: Less significant reduction in basal dendritic complexity in layer 2/3 SSC neurons in *Taok2* Het and KO mice compared to WT mice (WT=19, Het=18, and KO=18 neurons from 3 different brains; two-way ANOVA, *post hoc* Dunnett's test; $F(2,1204)=17.92$, $p<0.0001$ between genotypes; *represents ranges of significance; WT vs KO (red); see supplemental statistics). (c) Top: No significant difference in apical and basal dendrite length between WT and *Taok2* Het and KO SSC neurons (apical: WT=15, Het=17 and KO=15 neurons from 3 different mice brains; one-way ANOVA, *post hoc* Dunnett's; $F(2,44)=0.4811$, $p=0.6213$; basal: WT=19, Het=12 and KO=16 neurons from 3 different mice brains; one-way ANOVA, *post hoc* Dunnett's; $F(2,44)=1.32$, $p=0.2775$). (d) Dendritic spine distribution maps of P21 Golgi-stained WT and *Taok2* KO SSC neurons. (e) No significant change in spine distribution between WT and *Taok2* Het and KO SSC neurons (WT=6, Het=6, and KO=6 neurons from 3 different brains; two-way ANOVA, *post hoc* Dunnett's test; $F(2,615)=1.905$, $p=0.1497$ between genotypes). (f) No significant difference in apical and basal dendritic spines per cell between WT and *Taok2* Het and KO SSC neurons (WT=6, Het=6, and KO=6 neurons from 3 different mice brains; Left: one-way ANOVA, *post hoc* Dunnett's test; $F(2,15)=0.3927$, $p=0.6820$; Right: one-way ANOVA, *post hoc* Dunnett's test; $F(2,15)=0.1042$, $p=0.9017$;). (g) Cumulative probability histograms show shift towards reduced dendritic spine lengths (right) *Taok2* KO SSC neurons and reduced head widths (left) in *Taok2* Het and KO SSC neurons. (h) *Taok2* Het and KO SSC neurons have a significant increase in thin spines and reduction in mushroom-like spines when compared to WT

SSC neurons (WT = 4111 spines, Het = 3688, KO = 3729 spines from 6 cells per condition from three different brains; two-way ANOVA, *post hoc* Sidak's test; $F(2,90)=1.964e-13$, $p>0.9999$ between genotypes; WT vs KO F: $p=0.8822$, LT: $p=0.9557$, T: $p=0.0002$, S: $p=0.2832$, M: $p<0.0001$, and B: $p>0.9999$; WT vs Het F: $p=0.7578$, LT: $p=0.3811$, T: $p=0.0291$, S: $p=0.9864$, M: $p=0.0044$, and B: $p=0.9987$). (i) *Taok2* KO SSC neurons imaged by electron microscopy show decreased percentage of synapses on postsynaptic spines (left) and increased synapses on dendrites (right) (WT=95, KO=65 neurons; left: unpaired t-test; $t(162)=7.261$, $p<0.0001$; right: $t(162)=8.839$, $p<0.0001$). (j) Representative traces of mEPSC spikes from WT and *Taok2* KO SSC neurons. Scale: 5pA by 1sec. (k) Left: Longer inter-event intervals in *Taok2* Het and KO SSC neurons shown on a cumulative probability histogram. Inside: Reduced mEPSC event frequency in *Taok2* Het and KO SSC neurons (WT=13, Het=9 and KO=11 neurons from 3 different mice brains; one-way ANOVA, *post hoc* Dunnett's test; $F(2,31)=12.86$, $p<0.0001$; WT vs Het $p=0.0088$, WT vs KO $p<0.0001$). Right: No change in mEPSC amplitude in *Taok2* WT, Het and KO SSC neurons shown on a cumulative probability histogram. Inside: Average mEPSC amplitude of *Taok2* WT, Het and KO neurons (WT=12, Het=9 and KO=13 neurons from 3 different mice brains; kruskal-walis ANOVA, *post hoc* Dunn's test; H-value=0.01588, $p=0.9921$; WT vs Het $p>0.9999$, WT vs KO $p>0.9999$). * $p<0.05$, *** $p<0.001$ and **** $p<0.0001$. Values are mean \pm s.e.m.

Supplementary Figure 9. TAOK2 KO mice have no alterations in dendrite growth and synaptic connectivity in the hippocampus. (a) Top: Golgi-stained CA1 HC neurons from P21 WT and *Taok2* Het and KO mice. Scale bars represent 20 μ m. Bottom: Dendritic heat maps of superimposed neuron tracings for each condition. Blue to red (apical) and yellow to blue (basal) indicates increased probability of dendrite presence. . Scale bars represent 30 μ m. (b) Top: No major difference in apical and basal dendritic complexity in in layer 2 HC neurons in *Taok2* Het and KO mice (apical: WT=15, Het=14 and KO=18 neurons from 3 different brains; two-way ANOVA, *post hoc* Dunnett's test; $F(2,2156)=2.156$, $p=0.1160$ between genotypes, WT vs Het $p=0.0202$ at 35 μ m). Bottom: No significant change in basal dendritic complexity in CA1 HC neurons in *Taok2* Het and KO mice compared to WT mice (WT=18, Het=15, and KO=16 neurons from 3 different brains; two-way ANOVA, *post hoc* Dunnett's test; $F(2,1472)=7.277$, $p=0.0007$ between genotypes). (c) No significant difference in apical and basal dendrite length

between WT and *Taok2* Het and KO HC neurons (Apical: WT=15, Het=14, and KO=18 neurons from 3 different mice brains; one-way ANOVA, *post hoc* Dunnett's test; $F(2,44)=0.3342$, $p=0.7177$. Basal: WT=23, Het=16, and KO=22 neurons from 3 different mice brains; one-way ANOVA, *post hoc* Dunnett's test; $F(2,58)=0.4335$, $p=0.6503$). (d) Dendritic spine distribution maps of P21 Golgi-stained WT and *Taok2* KO HC neurons. (e) No major change in spine distribution between WT and *Taok2* Het and KO HC neurons (WT=6, Het=6, and KO=6 neurons from 3 different mice brains; Two-way ANOVA, *post hoc* Dunnett's test; $F(2, 600)=2.636$ between genotypes, $p=0.0725$; WT vs KO $65\mu\text{m}$ $p=0.0363$). (f) No significant difference in apical and basal dendritic spines per cell between WT and *Taok2* Het and KO HC neurons (WT=6, Het=6, and KO=6 neurons from 3 different mice brains; Left: one-way ANOVA, *post hoc* Dunnett's test; $F(2,15)=1.445$, $p=0.1668$); Right: one-way ANOVA, *post hoc* Dunnett's test; $F(2, 15)=0.5609$, $p=0.5822$). (g) Cumulative probability histograms show no shift in dendritic spine lengths (right) and head widths (left) in *Taok2* Het and KO HC neurons. (h) No significant difference in spine type between WT and *Taok2* KO HC neurons (WT = 7032, Het = 6747, and KO = 7489 spines from 6 cells per condition from three different brains; two-way ANOVA, *post hoc* Dunnett's test; $F(2,90)=1.317\text{e-}13$, $p>0.9999$ between genotypes). (i) *Taok2* KO neurons imaged by electron microscopy shows no change in percentage of synapses on postsynaptic spines (left) or dendrites (right) (WT=81, KO=60 neurons; *left*: unpaired t-test; $t(139)=0.7799$, $p=0.4368$; *right*: $t(139)=0.7245$, $p=0.4700$). (j) Western blot of pTAOK2 and TAOK2 levels in different regions of P21 *Taok2* WT, Het and KO mice brains (Cortex, Hippocampus, Striatum, Thalamus/Hypothalamus, Midbrain, Cerebellum, Pons/Medulla, Total Brain). Arrow head indicates pTAOK2 bands. Arrow indicates unspecific band. (k) The pTaok2/Taok2 ratio in the HC is significantly reduced compared to the cortex, (CTX=5, HC=5, STR=4, T/HT=5, MB=5, CB=4 and P/M=3 brains from 3 different litters; one-way ANOVA, *post hoc* Bonferroni's test; $F(6,24)=6.767$, $p=0.0003$; CTX vs HC $*p=0.0145$). $*p<0.05$. Values are mean \pm s.e.m.

Supplementary Figure 10. No change in mIPSC frequency and amplitude in PFC and SSC neurons in *Taok2* KO mice. (a) Representative traces of mIPSC spikes from WT and *Taok2* KO PFC neurons. Scale: 10pA by 1sec. (b) No change in inter-event intervals between WT and *Taok2* KO PFC neurons shown on a cumulative probability histogram. Inside: No change in

mean mIPSC event frequency between WT and *Taok2* KO PFC neurons (WT=14 and KO=14 neurons from 3 different mice brains; unpaired t-test; $t(26)=0.03925$, $p=0.9690$). (c) No change in mean mIPSC amplitude between WT and *Taok2* KO PFC neurons shown on a cumulative probability histogram. Inside: No change in average mIPSC amplitude between WT and *Taok2* KO PFC neurons (WT=14 and KO=14 neurons from 3 different mice brains; unpaired t-test; $t(26)=1.117$, $p=0.2742$). (d) Representative traces of mIPSC spikes from WT and *Taok2* KO SSC neurons. Scale: 10pA by 1sec. (e) No change in inter-event intervals between WT and *Taok2* KO SSC neurons shown on a cumulative probability histogram. Inside: No change in mean mIPSC event frequency between WT and *Taok2* KO SSC neurons (WT=14 and KO=12 neurons from 3 different mice brains; unpaired t-test; $t(24)=0.2274$, $p=0.8221$). (f) No change in mIPSC amplitude between WT and *Taok2* KO SSC neurons shown on a cumulative probability histogram. Inside: No change in mean mIPSC event frequency between WT and *Taok2* KO SSC neurons (WT=14 and KO=12 neurons from 3 different mice brains; unpaired t-test; $t(24)=0.09472$, $p=0.9253$). Values are mean \pm s.e.m.

Supplementary Figure 11. TAOK2 is enriched in the post synaptic density and loss of Taok2 decreases synaptic density and stability.

(a) Co-staining of Taok2 and SynGAP on DIV18-19 cortical neurons. Scale bars represent 10 μ m. (b) Western blot of synaptosome fractions (Homogenized fraction (Homo), Cell pellet (Pellet), Synaptonerosome fraction (SN), Synaptic fraction (Syn) and post-synaptic density fraction (PSD)) from 1-month-old WT mice shows that Taok2 is enriched in the PSD fraction. (c) Co-staining of WT and *Taok2* KO DIV18-19 cortical neurons with phalloidin and SynGAP. Scale bars represent 10 μ m. (d) Cultured KO cortical neurons have decreased number of spines per cell compared to WT neurons (WT=15 and KO=29 neurons from 3 different cultures; unpaired t-test; $t(42)=2.427$ $p=0.0196$). (e) Cultured *Taok2* KO cortical neurons have decreased number of SynGAP and phalloidin positive spines per cell compared to WT neurons (WT=15 and KO=29 neurons from 3 different cultures; unpaired t test; $t(42)=2.099$ $p=0.0418$). (f) Cortical neurons transfected with GFP and control shRNA or TAOK2 shRNA at various dosages (0.25, 0.50, and 0.75 μ g) Scale bars represent 10 μ m. (g) Cortical neurons transfected with TAOK2 shRNA show reduced dendritic spine density (spines/10 μ m dendrites) compared to EGFP and control shRNA transfected neurons (EGFP=25, ctrl=22, shTAOK2(0.25 μ g)=21, shTAOK2(0.50 μ g)=14, shTAOK2(0.75 μ g)=16 cells from three

different cultures; one-way ANOVA; *post hoc* Dunnett's test; $F(4,93)=18.22$, $p<0.0001$; EGFP vs shTAOK2(0.25 μ g) $p<0.0001$, EGFP vs shTAOK2(0.50 μ g) $p=0.0001$, EGFP vs shTAOK2(0.75 μ g) $p<0.0001$). (h) Images of DIV24 hippocampal neuron cultures transfected with only GFP (control) or TAOK2 shRNA. Boxes are magnified and shown at 0 min, 152min and 302min. Scale bars represent 10 μ m. (i) Images of DIV24 hippocampal neuron cultures transfected with Lifeact-GFP or TAOK2 shRNA and Lifeact-GFP. Boxes are magnified and showed to the right. Kymographs show increased movement of dendritic spines on neurons with acute loss of Taok2 (Bottom right). Scale bars represent 10 μ m. ns $p>0.05$, * $p<0.05$, *** $p<0.001$ and **** $p<0.0001$. Values are mean \pm s.e.m.

Supplementary Figure 12. Location of all *de novo* and rare-inherited mutations on *TAOK2* and pedigrees of *de novo* and truncation mutation probands. (a) Top: Detailed diagram of TAOK2 α and TAOK2 β isoforms and location of *de novo*, truncating, and rare-inherited mutations. Different protein domains are represented by colored boxes (kinase domain: red, MEK binding domain: black, regulatory domains: blue and light blue. TAOK2 α has two phosphorylation sites (ser181 and thr475) and caspase-9 cleavage site (⁹¹⁶DPGD⁹¹⁹). TAOK2 β has 3 known phosphorylation sites (ser181, thr475, and ser1031). Bottom: Diagrams of truncated TAOK2 due to the P1022*, Q622*, T604Sfs*45 and L1030Wfs*3 mutations. Dark red box represent frame shifts before introduction of premature stop. P1022* only effects the β isoform and L1030Wfs*3 only effects the α isoform. (b) Family pedigrees of 3 *de novo* and 3 truncating mutations. Families include 01-0337 (A135P/+), 07-0179 (c.563+12_563+15/+), 01-0559 (P1022*), 01-0446 (L1030Wfs*3), AU4112 (T604Sfs*3) and AU4261 (Q622*). Black filled squares represent individuals diagnosed with autism spectrum disorder. Genotype details of the AU4122 and AU4261 family members are not available.

Supplementary Figure 13. *TAOK2* mRNA expression analysis of 01-0337, 07-0179, and 01-0559 families. (a) No changes in *TAOK2* expression in LCLs from the A135P, P1022* and c.563+12_563+15 probands compared to average of controls (3 passages and 2 technical replicates; one-way ANOVA, *post hoc* Dunnett's test; $F(11,29)=3.928$, $p=0.0015$; Control vs P1022* father $p=0.0327$, Control vs P1022* affected sibling $p=0.0438$). (b) No changes in *TAOK2* β isoform-specific expression in LCLs from the A135P, P1022* and c.563+12_563+15

proband compared to average of controls (3 passages and 2 technical replicates; one-way ANOVA, *post hoc* Dunnett's test; $F(11,29)=2.841$, $p=0.0120$). RNA expression was measured using ddPCR and normalized to TBP mRNA levels. Parents and siblings of the 3 probands were used as controls for *TAOK2* expression. (c) c.563+12_563+15 proband have 13% fold increase in intron 7 retention. * $p<0.05$. Values are mean +/- s.e.m.

Supplementary Figure 14. The effect of the A135P mutation on TAOK2 α function and localization (a) Western blot of HEK 293 cell lysates 48 hours post transfection with TAOK2 α , TAOK2 α A135P, and TAOK2 α A335V. (b) TAOK2 α A135P has reduced ser181 phosphorylation (right) compared to TAOK2 α , while protein levels (left) remain unchanged. TAOK2 α A335V shows no difference compared to TAOK2 α . (n=5 western blots; one-sample t-test; TAOK2 levels: A135P: $t(4)=1.0171$, $p=0.3444$, A335V: $t(4)=1.151$, $p=0.3138$; pTAOK2 levels: A135P: $t(4)=15.80$, $p<0.0001$, A335V: $t(4)=0.05114$, $p=0.9617$). (c) Representative western blot of HEK293 cell lysates 48 hours post transfection with JNK1 α only or with TAOK2 α , TAOK2 α A135P. (d) TAOK2 α phosphorylates JNK to a higher degree ($1757\% \pm 505$) than TAOK2 β ($384\% \pm 89$). TAOK2 α A135P significantly decreases phosphorylation of JNK1 α in HEK 293 cells compared to TAOK2 α (n=5 western blots; TAOK2 α vs TAOK2 α A135P: unpaired t-test; $t(8)=3.138$, $p=0.0138$; Jnk1 α vs TAOK2 α : one-sample t-test, $t(4)=3.282$, $p=0.0304$). (e) Top: Staining of TAOK2 (red) on DIV14 cortical neurons transfected with palmitoylated GFP shows that TAOK2 α localizes with tubulin (green). Bottom: TAOK2 α (green) does not localize with F-actin rich protrusions (stained with rhodamine-labelled phalloidin; red). Scale bars represent 10 μ m. White boxes are magnified and shown to the right, with scale bars representing 3 μ m. (f) Top: Staining of TAOK2 (red) on DIV14 cortical neurons shows that TAOK2 α A135P does not affect localization with tubulin (green). Bottom: TAOK2 α A135P (green) also does not localize with F-actin rich protrusions (stained with rhodamine-labelled phalloidin; red). Scale bars represent 10 μ m. White boxes are magnified and shown to the right, with scale bars representing 3 μ m. All neurons are co-stained with DAPI. * $p<0.05$ and **** $p<0.0001$. Values are mean +/- s.e.m.

Supplementary Figure 15. P1022* *de novo* mutation does not affect TAOK2 auto-phosphorylation and stability and RhoA levels in LCLs compared to family members (a)

Western blot of LCL lysates from the P1022* proband and their family members. (b) No significant difference in TAOK2 (left) and pTAOK2 (right) levels in the P1022* proband compared to the rest of the family, where the mother is set at 100% (n=6 western blots; TAOK2 levels: one-sample t-test; 100% vs P1022*: $t(5)=2.023$, $p=0.0989$; pTAOK2 levels: 100% vs P1022*: $t(5)=0.3183$, $p=0.7631$) (c) Western blot of LCL lysates from the P1022* proband and their family members after RhoA-GTP pull-down. (d) No significant difference in RhoA-GTP (left) and RhoA (right) levels in the P1022* proband compared to the rest of the family, where the mother is set at 100% (n=6 western blots; RhoA-GTP levels: one-sample t-test; RhoA-GTP levels: 100% vs P1022*: $t(5)=0.5872$, $p=0.5826$; RhoA levels (normalized to tubulin): one-sample t-test; 100% vs P1022*: $t(5)=1.683$, $p=0.1532$). Values are mean +/- s.e.m.

Supplementary Figure 16. P1022* elongates basal dendrites and TAOK2 A135P and P1022* reduce dendrite integrity. (a) TAOK2 β P1022* increases dendrite length in WT *Taok2* cortical neurons (6 neurons from 3 different brains per condition; one-way ANOVA, *post hoc* Dunnett's test; $F(3,20)=19.02$, $p<0.0001$; TAOK2 α/β vs β P1022* $p<0.0001$). (b) Images of dendrite branching in cortical neurons from *in utero* electroporated WT *Taok2* P21 mice. Red arrow heads indicate the measurements on dendrites after branching and red bars highlight the width of dendrites in that region. Scale bars represent 3 μ m. (c) TAOK2 α/β A135P and TAOK2 β P1022* reduce dendrite thickness after branching (6 neurons from 3 different brains per condition; one-way ANOVA, *post hoc* Dunnett's test; 1st: $F(3,168)=18.72$, $p<0.0001$; Venus vs TAOK2 α/β $p=0.0062$, TAOK2 α/β vs α/β A135P $p<0.0001$, TAOK2 α/β vs β P1022* $p<0.0001$; 2nd: $F(3,265)=18.6$, $p<0.0001$; TAOK2 α/β vs α/β A135P $p<0.0001$, TAOK2 α/β vs β P1022* $p<0.0001$; 3rd: $F(3,274)=23.75$, $p<0.0001$; TAOK2 α/β vs α/β A135P $p<0.0001$, TAOK2 α/β vs β P1022* $p<0.0001$; 4th: $F(3,164)=24.61$, $p<0.0001$; TAOK2 α/β vs α/β A135P $p<0.0001$, TAOK2 α/β vs β P1022* $p<0.0001$; 5th: $F(3,74)=7.031$, $p=0.0003$; TAOK2 α/β vs β P1022* $p=0.0002$).

Supplementary Figure 17. RhoA interacts with Taok2 in the mouse cortex and preferentially with TAOK2 β isoform in HEK239 cells. (a) Western blot of Taok2 immunoprecipitation (IPs) from a crude homogenate (H) and a crude membrane fraction (Pellet fraction 1, P1) from WT mouse cortices show direct or indirect interaction of Taok2 and RhoA.

Membranes were immunoblotted against Taok2, RhoA and β -actin (loading control). Unspecific goat IgGs were used as IP controls. (b) Western blot of RhoA immunoprecipitations (IPs) from lysates from HEK293 cells transfected with GFP or TAOK2 α and β variants. Lysates were blotted against pTAOK2, TAOK2, RhoA, and β -actin as loading control. (c) Quantification of Taok2 co-immunoprecipitation with RhoA from HEK293 cells (shown in b) reveals that TAOK2 β preferentially binds RhoA. % of TAOK2 β bound to RhoA was set to 100%. TAOK2 β P1022*, TAOK2 α , and TAOK2 α A135P has significantly reduced binding to RhoA compared to TAOK2 β . (N=3 western blots, one-way Anova, post hoc Dunnett's test; F(4,10)=12.75, p=0.0006; TAOK2 β vs TAOK2 β A135P p=0.3245, TAOK2 β vs TAOK2 β P1022* p=0.0118, TAOK2 β vs TAOK2 α p=0.0007, TAOK2 β vs TAOK2 α A135P p=0.0006).

Supplementary Figure 18. The role of human-derived TAOK2 mutations identified in ASD probands at the synapse. (a) TAOK2 α associates with microtubules, while TAOK2 β localizes to dendritic spines. It is suggested that TAOK2 β increases levels of active RhoA-GTP, which controls stability of the actin cytoskeleton and dendritic spine. The A135P mutation in TAOK2 α does not alter localization to microtubules, while TAOK2 β A135P cannot localize to dendritic spines resulting in a LOF phenotype. It is suggested that the TAOK2 β P1022* has increased kinase activity and thus may increase phosphorylation of JNK leading to elongation of microtubules and dendrites. TAOK2 β P1022* also does not localize to dendritic spines. Phosphorylation of JNK1 by TAOK2 α A135P is also disrupted. (b) Flow chart describing the characterization of Taok2 KO mice by studying changes in behavior, brain morphology, and functional cellular changes in cortical neurons, cell lines, and human cell samples. This is combined with the identification of *de novo* and rare-inherited mutation in TAOK2 by whole exome/genome sequencing of ASD families. These human-derived mutations were further studied using functional *in vitro* assays using murine and human cell systems, cortical neuron cultures derived from Taok2 KO mice as well as *in vivo* assays in *Taok2* Het mice.

Video 1. Time-lapse imaging of 14DIV neurons expressing control shRNA and TAO2 shRNA together with palmitoylated GFP plasmid. Epi-fluorescence imaging was performed on an inverted Nikon microscope (Eclipse, Ti) with a 60x objective (NA 1.4). Duration of time-lapse imaging: 5 min acquiring images every 2 sec.

Video 2. Time-lapse imaging of 14DIV neurons expressing TAO2 β , TAO β A135P, TAO β P1022 together with palmitoylated GFP plasmid. Epi-fluorescence imaging was performed on an inverted Nikon microscope (Eclipse, Ti) with a 60x objective (NA 1.4). Duration of time-lapse imaging: 5 min acquiring images every 2 sec.

Video 3. Time-lapse imaging of 14DIV TAO2 KO neuron expressing Lifeact-GFP plasmid. Cell was imaged before CN01 treatment for 5 min (acquiring images every 2 sec). After CN01 treatment the same cell was imaged for 5 min (acquiring images every 2 sec). Epi-fluorescence imaging was performed on an inverted Nikon microscope (Eclipse, Ti) with a 60x objective (NA 1.4).

**CHAPTER 3: NEURON-SPECIFIC PROTEIN NETWORK MAPPING OF
AUTISM RISK GENES IDENTIFIES MITOCHONDRIAL RESPIRATION AND
METABOLISM AS A CONVERGENT BIOLOGICAL PATHWAY**

AUTHORS

Nadeem Murtaza^{1,2}, Annie A. Cheng², Durga Praveen Meka³, Shuai Hong³, Brianna K. Unda², Sansi Xing¹, Eric Deneault⁴, Froylan Calderon de Anda³, Bradley W. Doble⁵, James Ellis^{6,7}, Gary D. Bader⁷, Stephen W. Scherer^{6,7}, Yu Lu¹, Karun K. Singh^{2,8}*

¹Department of Biochemistry and Biomedical Sciences, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada

²Krembil Research Institute, University Health Network, Toronto, ON, Canada

³Center for Molecular Neurobiology Hamburg (ZMNH), Research Group Neuronal Development, University Medical Center Hamburg-Eppendorf (UKE), Hamburg, Germany

⁴Centre for Biologics Evaluation, Biologic and Radiopharmaceutical Drugs Directorate, Health Products and Food Branch, Health Canada, Ottawa, Ontario, Canada ⁵Department

of Paediatrics, Department of Biochemistry and Medical Genetics, University of Manitoba, Winnipeg, MB, Canada ⁴Department of Laboratory Medicine and

Pathobiology, Faculty of Medicine, University of Toronto, Toronto, ON, Canada

⁶The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, ON, Canada

⁷Department of Molecular Genetics, Faculty of Medicine, University of Toronto,
Toronto, ON, Canada

⁸Department of Laboratory Medicine and Pathobiology, Faculty of Medicine, University
of Toronto, Toronto, ON, Canada

***Correspondence:** karun.singh@uhnresearch.ca

ACKNOWLEDGEMENTS

We thank C.I. and P.dG. for proof reading the manuscript. We also thank K.J.B. and M.B.F. for providing the NRXN1 cDNA. Flowcharts were created with BioRender.com (SD235B8ORF, KW235KT7TM, RZ235KTA0S). Work in the Singh Lab was supported by the Canadian Institute of Health Research (CIHR), the Ontario Brain Institute (OBI), the Network for European Funding for Neuroscience Research (NEURON ERA-NET), and the Donald K. Johnson Eye Institute at University Health Network.

PREFACE

The combined manuscript from this chapter and Chapter 4 is currently in the first round of revisions at Cell Reports as of July 2022. It has been modified and kept separate to match the formatting of this thesis. References for Chapter 3 and 4 are combined with the thesis references. The aim of this chapter was to identify the convergent biological pathways and mechanisms shared between ASD-risk genes. The findings in Chapter 2, which showed that disruptions in TAOK2 cause multiple deficits in neurodevelopment associated with ASD, led us towards identifying the proteins that are directly associated

with the TAOK2 protein. However, the use of BioID to study one ASD-risk gene led to the combined idea, by Dr. Karun Singh and myself, to study as many genes as possible to investigate the convergence of ASD-risk gene functions.

Dr. Karun Singh and I designed the study and all experiments and I generated all of subsequent DNA constructs and performed all experiments and data analysis unless otherwise specified. Annie Cheng and Dr. Brianna Unda created the initial BioID2 lentiviral construct backbone. Annie Cheng assisted with all lentiviral generations. Sansi Xing and Yu Lue ran samples through the mass spectrometer and assisted with data acquisition. Dr. Durga Meka, Shuai Hong, Birgit Schwanke and Froy Calderon de Anda performed all mitochondrial activity and content experiments in mouse cortical neurons. Eric Deneault, Dr. James Ellis, and Dr. Steve Scherer assisted in CRISPR/Cas9 editing the human *TAOK2* KO and A135P iPSC lines. Dr. Gary Bader advised on pathway analyses used in the project. Dr. Karun Singh and I wrote the manuscript with input from Annie Cheng and Dr. Bradley Doble.

3.1 ABSTRACT

Autism spectrum disorder (ASD) is a genetically heterogeneous disorder. Sequencing studies have identified hundreds of risk genes for autism spectrum disorder (ASD), but the signaling networks of actively expressed genes at the protein level remain largely unexplored. To address this gap, we used neuron-specific proximity-labeling proteomics (BioID) to identify protein-protein interaction (PPI) networks of 41 ASD-risk genes. Neuron-glia co-cultures provided neuron-specific PPI networks that were not

largely identified in frequently used HEK293 cells. These neuron-specific networks included synaptic transmission and dendritic arborization proteins, for which we showed that patient *de novo* missense variants disrupt the PPI network of synaptic ASD-risk genes and cause deficits in synaptic transmission. Network analysis revealed that the combined 41 risk gene PPI network map had a 50-fold increase in shared connectivity between distantly related ASD-risk genes than those in existing public databases. Through the PPI network map, we identified convergent pathways between established and uncharacterized risk genes, including synaptic transmission, mitochondrial/metabolic processes, Wnt signaling pathways, ion channel activity and MAPK signaling. Using CRISPR gene knockouts, we further investigated metabolic dysfunction, which has been peripherally associated with ASD, and revealed a functional association between mitochondrial activity and multiple ASD-risk genes not previously connected with this pathway. Together, our data reveal that using cell type-specific PPI networks to map ASD risk genes can identify previously unknown individual and convergent neuronal signaling networks and reveal biological insight into disease mechanisms and sub-cohorts in ASD.

3.2 INTRODUCTION

Autism spectrum disorders (ASD) are a heterogeneous group of neurodevelopmental conditions that manifest early in life, occurring in 1 in 44 children under the age of 8 (Of et al., 2018). The risk of developing ASD has a strong genetic basis, including common and rare genetic risk variants (Gaugler et al., 2014; Iossifov et

al., 2014; Robinson et al., 2016; Weiner et al., 2017). As such, numerous large scale WES and WGS studies have identified hundreds of genes associated with ASD risk (Glessner et al., 2009; Iossifov et al., 2012; Neale et al., 2012; O’Roak et al., 2012b, 2012a; Sanders et al., 2012, 2015; Yuen et al., 2017; Feliciano et al., 2019; Ruzzo et al., 2019; Satterstrom et al., 2020; Wilfert et al., 2021). While the mechanisms by which different risk genes lead to disease are poorly understood, one hypothesis is that they converge functionally within brain signaling networks. Understanding signaling convergence can help reveal the risk genes that work through common pathways and have functional relationships. In turn, this could help classify autism risk genes based on biological pathways, prioritize the discovery of new risk genes, and identify convergent pathways that could be harnessed for targeted therapy development.

The majority of convergent ASD-associated pathways discovered to date are based on exome and genome sequencing, transcriptomics, and gene co-expression analyses, including CRISPR/Cas9 knockout screens combined with single cell RNA sequencing (Parikshak et al., 2013; Willsey et al., 2013, 2021; Sanders et al., 2015; Yuen et al., 2017; Cederquist et al., 2020; Jin et al., 2020; Ramaswami et al., 2020; Satterstrom et al., 2020). These studies have implicated pathways such as synaptic transmission, translation, transcription, chromatin remodeling and splicing (Voineagu et al., 2011; O’Roak et al., 2012a; Chang et al., 2015; Velmeshev et al., 2019; Ramaswami et al., 2020). However, the majority of autism risk genes encode proteins, and protein-protein interactions (PPIs) are an essential mechanism of signaling (Neale et al., 2012; O’Roak et al., 2012b). Therefore, non-protein interaction-based networks, while important, lack

information regarding which ASD-risk genes interact with each other or converge into common signaling networks at the protein level. Given that a large proportion of ASD genes have functions that do not take place in the nucleus and or that are not involved in gene expression (De Rubeis et al., 2014; Satterstrom et al., 2020), identification of PPI networks provides an unbiased approach to gain insights into unknown convergent ASD disease processes at the protein level outside the nucleus (Kuzmanov and Emili, 2013; Murtaza et al., 2020). Previous ASD sequencing studies have shown that that risk genes are part of core PPI networks (Neale et al., 2012; O’Roak et al., 2012a; Chang et al., 2015; Chen et al., 2020), and large yeast-two-hybrid (Y2H) studies have identified PPI networks shared between ASD-risk genes (Sakai et al., 2011; Corominas et al., 2014). However, these data are extracted from databases that are largely derived from non-neuronal cell lines and tissues, and do not represent brain-specific networks(Lage, 2014). The lack of ASD risk-gene PPI networks in disease-relevant cell types represents a missing link towards understanding the biological mechanisms of ASD.

Multiple techniques can be used to identify PPIs, including affinity purification or proximity-labeling proteomics combined with mass spectrometry (reviewed in Richards *et al.*, 2021(Richards et al., 2021). Both are powerful approaches to identify PPI networks in cells but have caveats that can be mitigated by using appropriate controls and validations. Further, many brain-expressed genes are large in size, including ASD-risk genes(Casanova et al., 2019), which limits the systems that can be used for expression in cells and allow identification of their PPI networks. We took an approach that balances gene size limitations, while at the same time captures strong and transient interactions

and proteins in close proximity to build comprehensive BioID PPI networks for ASD risk genes. We developed a lentiviral *in vitro* proximity-labeling proteomics (BioID2) system that uses mouse primary neurons. Proximity-labeling proteomics has been used successfully to capture physiologically relevant interactomes in neural cell-types both *in vitro* and *in vivo* (Loh et al., 2016; Uezu et al., 2016; Chung et al., 2017; Spence et al., 2019; Hamdan et al., 2020) or to map cellular compartments (Hung et al., 2017; Markmiller et al., 2018; Go et al., 2021). Given the implication of cortical neurons in ASD pathology (Nowakowski et al., 2017; Velmeshev et al., 2019), we captured PPI networks from cortical neurons, while allowing them to grow with their glial counterparts to promote proper maturation (Barres, 2008; Stogsdill and Eroglu, 2017; Wilton et al., 2019).

In the current study, we address the lack of brain cell type-specific PPI networks for ASD-risk genes. We designed a screen to identify the interactome of 41 ASD-risk proteins in neurons by using proximity-dependent biotinylation paired with mass spectrometry. We targeted non-nuclear proteins (e.g., cytosolic proteins, receptors, kinases, scaffold proteins, and intracellular signaling proteins) because nuclear proteins have a high level of endogenous biotinylation and categorically different functional pathways. Our screen found 1770 protein-level connections and neighbourhood proteins between the 41 genes in neurons, which was approximately 20-times that reported in the STRING database (at lowest confidence) (Snel et al., 2000). Convergent protein networks included synaptic transmission, mitochondrial/metabolic processes and Wnt signaling. Further investigation of genes not previously linked to mitochondrial/metabolic

processes, through gene-knockout approaches, revealed that multiple genes regulate mitochondrial cellular respiration in mouse and human neurons.

Taken together, we demonstrate that neuron-specific PPI networks provide a powerful approach to reveal novel individual and convergent disease mechanisms in ASD. Given the scalability of our method and its underutilization in ASD research, we believe our PPI network resource and screening system can be applied more broadly to additional autism risk genes to identify previously unknown or overlooked disease mechanisms that are not captured with current approaches.

3.3 METHODS AND MATERIALS

Antibodies

The following antibodies were used for immunostaining and immunoblotting experiments: rabbit anti-FLAG (IB 1:2,000, MilliporeSigma, F7425), mouse anti-FLAG (IF 1:1,000, IB: 1:2000, MilliporeSigma, F3165), rabbit anti-turboGFP (IF 1:1,000, IB 1:1,000, Fisher, PA5-22688), chicken anti-MAP2 (IF 1:1,000, Cedarlane, CLN182), rabbit anti- β -actin (IB 1:1,000, Cell Signaling, 8457S), mouse anti- β -actin (IB 1:5,000, MilliporeSigma, A5316), goat anti-TAOK2 α/β (IB 1:1,000, Santa Cruz Biotechnology, sc-47447), rabbit anti-TAOK2 β (IB 1:1,000, Synaptic Systems, 395 003), mouse anti-Synapsin1 (IF: 1:1000, Synaptic Systems, 106 001), mouse anti-TOMM20 (IF 1:100, US Biological, 134604), DAPI (IF 300mM, ThermoFisher, D21490), Hoechst (IF 1:10,000, Invitrogen, 1050083), Phalloidin-488 (IF 1:120, Cytoskeleton Inc., PHDG1), Anti-mouse-Cy3 (IF 1:500, Jackson ImmunoResearch, 715-165-151), Alexa 488 anti-rabbit (IF

1:500, Jackson ImmunoResearch, 711-545-152), Alexa 488 anti-chicken (IF 1:500, Jackson ImmunoResearch, 703-545-155), 405 conjugated-streptavidin (IF 1:500, Jackson ImmunoResearch, 016-470-084), 405 anti-chicken (IF 1:500, Jackson ImmunoResearch, 703-475-155), Alexa 647 anti-mouse (IF 1:500, Jackson ImmunoResearch, 715-605-150).

Generation of constructs

All cloning was accomplished using the In-Fusion HD cloning kit (Takara). To create the BioID2 fusion constructs, we obtained an expression construct containing a 198bp (13x “GGGGS” repeat) linker sequence upstream of a C-terminal 3xFLAG-tagged BioID2 sequence with BioID2 (Genscript). For lentiviral expression, 13xlinker-BioID2-3xFLAG was amplified and cloned into the lentiviral backbone pLV-hSYN-RFP (Addgene #22909)(Nathanson et al., 2009). For ease of visualization and to create a bicistronic construct, the RFP in the pLV-hSYN-RFP backbone was replaced with the TurboGFP(tGFP)-P2A from pCW57-GFP-2A-MCS (Addgene #71783)(Barger et al., 2019). NheI digest sites were added after the P2A sequence and before the 13xLinker to allow easy insertion of ASD-risk bait genes. The final construct being pLV-hSyn-tGFP-P2A-Bait-13xLinker-BioID2-3xFLAG (referred to as the BioID2 fusion construct). For the control luciferase construct a second P2A was cloned in between the luciferase ORF and the 13xLinker, creating the pLV-hSyn-tGFP-P2A-Luciferase-P2A-13xLinker-BioID2-3xFLAG construct (referred to as the Luciferase control construct). ASD-risk genes open reading frames (ORFs) were purchased from Addgene and Genscript or amplified from human adult and fetal brain RNA (Takara) (see Table S10)(Seeling et al.,

1999; Furlong et al., 2000; Ohno et al., 2002; Cummins and Vogelstein, 2004; Howarth et al., 2005; Urano et al., 2007; Sowa et al., 2009; Solowska et al., 2010; Johannessen et al., 2010; Braun et al., 2011; Alford et al., 2012; Butko et al., 2012; Lu et al., 2014; Malecki et al., 2015; Wang et al., 2016a; Kim et al., 2016; Braganza et al., 2017; Hiday et al., 2017). For mouse electrophysiology experiments, the GRIA1, GRIA1 R208H and GRIA1 A636T ORFs were inserted between the GFP-P2A and 3xFLAG. The pLV-CMV-Cas9-T2A-EGFP plasmid was made by replacing the UBC promoter-rTetR in the FUW-M2rtTA plasmid (Addgene #20342)(Hockemeyer et al., 2008) with CMV-Cas9-T2A-EGFP from PX458 (Addgene #48138)(Ran et al., 2013). All generated constructs are available upon request.

Animal housing

Taok2 Het (*Taok2* +/-) and KO (*Taok2* -/-) mice were created by Kapfhamer *et al.*(Kapfhamer et al., 2013). The E15-16 or E18 mouse embryo brains were used for cortical neuronal cultures. P21-P23 mice were used for mass spectrometry or RNA sequencing experiments. Animals housed at the Central Animal Facilities at McMaster University were approved for experiments and procedures by the Animal Research Ethics Board (AREB) at McMaster University. Animals housed at the University Medical Center Hamburg-Eppendorf, Hamburg were approved for experiments and procedures by local authorities of the city-state Hamburg (Behörde für Gesundheit und Verbraucherschutz, Fachbereich Veterinärwesen) and the animal care committee of the University Medical Center Hamburg-Eppendorf. All procedures were performed

according to the German and European Animal Welfare Act. Animals housed at the Animal Resource Center at University Health Network were approved for experiments and procedures by the University Health Network animal care committees.

Mouse Cortical Neuron Cultures

E15-E16 CD1 mice (Charles River) embryo cortices were harvested using a dissecting microscope and kept in HBSS. Cortices were then digested in 300 µg/ml of papain (Worthington) and 2 U/ml of DNase (Thermo) for 20 minutes at 37 °C. Cortices were then washed three times with mouse plating media (Neurobasal media supplemented with 2 mM GlutaMAX (Thermo), Pen-Strep (Thermo), and 10% FBS(Gibco)). Digested cortices were triturated and put through 40 µm strainer. Cells were counted, suspended in plating media, and plated at 600,000 cells per well of a 12-well plate. Plates were coated with 100 µg/ml poly-D-lysine (mol wt > 300,000, Sigma) and 3 µg/ml Laminin (Sigma). For immunostaining, 12 mm coverslips (Fisher) were placed in the well prior to coating. The cells were incubated at 37 °C (with 5 % CO₂) for one hour, after which plating media was removed and replaced with mouse culturing media (Neurobasal media supplemented with 2 mM GlutaMAX, Pen-Strep, and B27). Cells were grown at 37 °C (with 5 % CO₂) and half media changes were done on day 7 and every 3-4 days onwards.

CRISPR/Cas9 editing of human induced pluripotent stem cells (iPSCs)

All work with the human iPSCs was performed with the approval of the Hamilton Integrated Research Ethics Board. Human iPSCs were maintained on Matrigel (Corning) coated plates using mTeSR1 media (Stem Cell Technologies) and passaged every 3-4 days using ReLeSR (Stem Cell Technologies). Human iPSCs were edited for homozygous knockout of *TAOK2* or heterozygous knock-in of the A135P mutation as described in Deneault *et al.* (Deneault et al., 2018). MGB probes were ordered from ThermoFisher scientific and single-stranded oligodeoxynucleotide (ssODN) were designed on Benchling.com (Biology Software) and ordered from Integrated DNA Technologies. For the A135P mutation a mutant and wildtype ssODNs containing the A135P (G to C) mutation and a PAM site mutation or just the PAM site mutation, respectively, were used to create a heterozygous knock-in.

Human iPSC to neuron differentiation via NGN2 induction

Human iPSCs were cultured on Matrigel (Corning) coated plates using mTeSR1 media (Stem Cell Technologies) and passaged every 3-4 days using ReLeSR (Stem Cell Technologies) until neural induction. A modified NGN2 induction protocol (Zhang *et al.* 2013) was used to differentiate human iPSCs into excitatory NGN2 neurons (Zhang et al., 2013). Human iPSCs were dual infected with pTet-O-NGN2-P2A-EGFP and FUW-M2rtTA lentiviruses for dox-inducible expression and were titered for > 90% infection efficiency. On Day -1 iPSCs were singularized using Accutase (Stem Cell Technologies) and plated with mTeSR1 media (supplemented with 10 μ M Y-27632) on Matrigel at 400,000 cells per well in a 6-well plate. On Day 0, media exchanged and supplemented

with Doxycycline (1 µg/ml). On Day 1 and 2, media was replaced with iNPC media (DMEM/F12 media (Gibco) supplemented with N2 (Gibco), MEM NEAA (Thermo), 2mM GlutaMAX, and Pen-Strep) with Doxycycline and Puromycin (2µg/ml). On Day 3, media was then replaced with iNi media (Neurobasal media with SM1 (Stem Cell Technologies), 2mM GlutaMAX, Pen-Strep, 20 ng/ml BDNF, 20 ng/ml GDNF, and 1 µg/ml Laminin) with Doxycycline. On day 4, differentiated neurons were singularized using Accutase and re-plated at 100,000 cells per well in a 24-well plate in only iNi media. Plates were pre-coated with 20 µg/ml Laminin and 67 µg/ml Poly-ornithine (Sigma). Mouse glial cells were plated on top of the differentiated neurons after 24 hours at a density of 50,000 cells per well. Half-media changes were carried out every other day, and iNi media was supplemented with 2.5 % FBS on Day 9 and onwards. Neurons were grown until day 28 post NGN2-induction.

Generation of high-titer lentivirus

All viruses were made using the 2nd generation lentiviral packaging systems in Lenti-X HEK293 FTT cells (Takara). Lenti-X cells were passaged maximum 3 times before being used for virus production in HEK media (High glucose DMEM with 4 mM GlutaMAX, 1 mM Sodium Pyruvate, and 10 % FBS). Lenti-X cells were passaged once with 500µg/ml Gentamycin (Thermo) to increase T antigen expressing cells. Cells were plated into T150 flasks and each flask was transfected with the BioID2 lentiviral plasmid and the packaging plasmids, pMD2.G and pPAX2 (Addgene #12259 and #12260), using Lipofectamine 2000 in a 3:5 Opti-MEM: HEK media mix. Media was exchanged for

fresh media after 5.5-6 hours. Media was harvested twice, first at 48 hours and then at 72 hours post-transection and spun at 100,000xG for 2 hours (maximum acceleration and deceleration). The virus was resuspended in PBS and kept at -80°C until they were used. Larger and less stable viruses were spun at 20,000xG for 4 hours in a table top centrifuge using a 20 % sucrose cushion(Yacoub et al., 2012). See nature exchange protocol for detailed procedure.

Infection of mouse cortical neurons for BioID2 screen

One plate of 7.2 million mouse cortical neurons was considered as one biological replicate. Each cortical neuron culture produced at least 5 plates for four separate BioID2 bait gene samples and one luciferase control sample. Three separate cultures were done in a 3 days span in one week to get 3 biological replicates per protein-of-interest (POI). On days *in vitro* (DIV) 14, the conditioned media from the mouse neuron cultures were removed, leaving only 0.5 ml of media per well. Extra wells with and without coverslips were infected at the same MOI for flow cytometry measurements of GFP positive neurons and immunostaining, respectively. On DIV14, lentivirus with BioID2 fusion constructs were added to each well at an MOI of 0.7 and on DIV17 each well was supplemented with 50 µM of Biotin. After 18-20 hours, cells for mass spectrometry were lysed with RIPA buffer (1 % NP40, 50 mM Tris-HCl, 150 mM NaCl, 0.1 % SDS, 0.5 % deoxycholic acid, and protease inhibitor cocktail (PIC)) and flash frozen in liquid nitrogen. Cells for flow cytometry were dissociated with 0.25 % Trypsin-EDTA (Fisher) and resuspended in PEF media (PBS with 2 mM EDTA and 5 % FBS) (See flow

cytometry section). Cells for immunostaining were fixed with 4 % PFA for 20 minutes, washed with PBS, and kept at 4 °C for staining.

Transfection of HEK293 FT cells for BioID2 screen

10 million HEK293 FT cells were plated in a 10 cm culture dish and transfected 24 hours later with the BioID2 fusion construct plasmids using Lipofectamine 2000. Media was changed 6 hours after transfection and 50µM biotin was added 48 hours post-transfection. Cells were lysed 72 hours post-transfection in RIPA buffer and flash frozen in liquid nitrogen. Each individual plate was considered as biological replicate and three plates were used for each gene and the luciferase control. An extra plate was used for flow cytometry measurements of GFP positive cells.

Processing of mouse cortical neuron and HEK293 FT cell BioID2 samples

Lysed cells were thawed and DNA was digested using benzonase (Sigma). Lysates were then sonicated at high speed for 5 seconds and centrifuged at 20,000xG for 30 minutes. The lysate supernatants were incubated with streptavidin Sepharose beads (GE Healthcare) at 4 °C for 3 hours. Following the incubation, the supernatant was spun down at 100xG for 2 minutes and the supernatant was removed. The beads were then washed once with RIPA buffer, and then six times with 100 mM triethylammonium bicarbonate (TEAB) with centrifugation between each wash. After the final wash, the beads are then resuspended in 100 mM TEAB and sequencing-grade trypsin (Promega) was added to digest the biotinylated proteins on the beads into peptides. The beads were incubated at 37 °C for 16 hours while rotating, and additional trypsin was added and

incubated for a further 2 hours. The beads were then pelleted and the supernatant was transferred to a new tube. The beads were washed twice with 100 mM TEAB and each wash was added to the supernatant. The supernatant was then transferred to a 1.5 mL screw cap tube and speed vacuum dried. The dried peptides were stored at 4 °C for TMT-labeling.

Multiplex TMT-labeling of BioID2 samples

Dried peptides were resuspended in 100 mM TEAB. Each sample was TMT-labeled using the TMT 10plex Isobaric Mass Tagging Kit (Thermo). The four genes (proteins-of-interest, POI) were divided into two separate batches and the luciferase control samples were divided between the batches. Each batch had three biological replicates of the two genes and the luciferase control. One luciferase sample chosen at random was divided and labeled with two different labels to determine variance due to labeling efficiencies. In brief, TMT-label resuspended in acetonitrile was added to each sample and incubated at room temperature for one hour. To stop the reaction, 5 % hydroxylamine was then added to the samples and incubated for 15 minutes at room temperature. All ten samples were combined into one tube and divided into two samples. Both samples were then speed vacuum dried. One sample was kept at -80 °C for storage and the second sample was kept at 4 °C to be run in the mass spectrometer.

Identification of biotinylated proteins from BioID2 screen samples using LC-MS/MS

Peptide samples were resuspended in 0.1% Trifluoroacetic acid (TFA) and loaded for liquid chromatography, which was conducted using a home-made trap-column (5 cm x 200 μ m inner diameter; POROS 10 μ m 10R2 C18 resin) and a home-made analytical column (50 cm x 50 μ m inner diameter; Monitor 5 μ m 100A C18 resin), running a 120min (label free) or 180min (TMT) reversed-phase gradient at 70nl/min on a Thermo Fisher Ultimate 3000 RSLCNano UPLC system coupled to a Thermo QExactive HF quadrupole-Orbitrap mass spectrometer. A parent ion scan was performed using a resolving power of 120,000 and then up to the 20 most intense peaks were selected for MS/MS (minimum ion count of 1000 for activation), using higher energy collision induced dissociation (HCD) fragmentation. Dynamic exclusion was activated such that MS/MS of the same m/z (within a range of 10 ppm; exclusion list size = 500) detected twice within 5 seconds were excluded from analysis for 30 seconds. Data were analyzed using Proteome Discoverer 2.2 (Thermo). For protein identification, search was against the Swiss-Prot mouse proteome database (55,366 protein isoform entries)(Bateman et al., 2021), while the search parameters specified a parent ion mass tolerance of 10 ppm, and an MS/MS fragment ion tolerance of 0.02 Da, with up to two missed cleavages allowed for trypsin. Dynamic modification of +16@M was allowed.

Analysis for the identification of ASD-risk and cellular compartment protein PPI networks

Only proteins identified with two unique peptides were used for analysis. Flow cytometry was used to calculate the total GFP in infected neuron samples. If the POI

sample had less GFP than the luciferase control sample, the factor needed to equalize the amount of GFP was applied to the protein abundances of the POI samples. Protein abundances were also normalized to the highest total protein count sample for each set of biological replicates. Unpaired one-tailed student's test was used to determine significantly enriched biotinylated proteins in the POI sample using the Log₂ abundances of the three biological replicates of the POI samples compared to the luciferase control samples ($p < 0.05$) (Uezu et al., 2016). Significance B outlier test was used to identify significantly biotinylated proteins in the POI sample compared to the luciferase control sample using the average abundance and protein abundance ratio between POI and luciferase samples (SigB $p < 0.05$). Only proteins that were found to be significant from both analyses were included in the PPI network. The protein abundance ratio between the luciferase control replicate samples, which were labeled with different TMT labels, was considered to be the minimal ratio required for significance. Any protein that did not surpass this ratio was considered to be a false positive, even if statistically significant, and not included in the PPI network.

Pathway enrichment analyses

All pathway enrichment analysis was done using the g:Profiler GOst functional profiling tool (<https://biit.cs.ut.ee/gprofiler/gost>) (Raudvere et al., 2019). We used internal sources without electronic GO annotations for GO biological processes and GO cellular component (compartment), and curated Reactome pathway gene sets from the Bader lab (http://download.baderlab.org/EM_Genesets/) (Merico et al., 2010). All three sources

were used for the shared ASD-risk gene network proteins. Only GO cellular component enrichment was used for the HEK293 FT cell BioID2 PPI networks, neuron cellular compartment BioID2 PPI networks, and *de novo* missense mutation network BioID2 PPI network comparisons. We compared the protein lists against a custom statistical domain of proteins identified through fractionated mass spectrometry of the mouse brain (Sharma et al., 2015) and combined with any additional proteins identified in the BioID2 screen. The final mouse brain proteome background had a total of 11992 proteins after removing multiple isoforms of the same protein. HEK293 BioID2 PPI networks were compared to all annotated gene lists. The g:Profiler Benjamini-Hochberg FDR multiple correction was used and only pathways with an adj. p-value < 0.05 were considered significantly enriched. For cellular component enrichment for *de novo* missense variant BioID2 PPI networks, the ggplot package in R was used to create the dot plots. For *de novo* missense variant BioID2 bait genes, all proteins identified in the wildtype samples were used for analysis, while for the shared PPI network map only proteins found in all wildtype samples were used for pathway enrichment analysis.

Virus titering and GFP normalization for BioID2 screen

Mouse cortical neurons were cultured as described above and infected on DIV3 at three dilutions of virus (1:100, 1:333, 1:1000). On DIV 5, infected mouse cortical neurons were singularized using 0.25 % Trypsin-EDTA (Fisher) and resuspended in PEF media (PBS with 2 mM EDTA and 5 % FBS). For GFP normalization, DIV18 mouse neurons infected with the BioID2 lentiviruses were dissociated with Trypsin and

resuspended in PEF media. CytoFLEX-LX or CANTO II flow cytometers were used to measure the percentage of GFP-positive cells with the 488 laser and 525/40 or 525/50 filters, respectively, using CytExpert software (Beckman Coulter). Functional titers were calculated based on the linear relationship between virus amount and percent of GFP positive cells. Mouse cortical neurons were infected at an MOI of 0.7, where 70 percent of cells were expected to be infected with the BioID2 lentiviral constructs. For normalization, the total GFP per 20,000 GFP-positive cells were quantified by taking the area under the GFP intensity histogram. GFP percentage and total amount was calculated using FlowJo software.

Western blots

HEK293 FT cells were transfected with the BioID2 constructs using Lipofectamine 2000 (Invitrogen) in Opti-MEM: HEK media. Cells were harvested 48 hours post-transfection and lysed with RIPA buffer (with fresh PIC). Lysates were either snap-frozen in liquid nitrogen or taken directly for western blot sample preparation. Thawed or fresh lysates were sonicated at high frequency for 5 seconds and centrifuged at 20,000xG for 5 minutes at 4 °C. Lysates were then quantified using the Bio-Rad Bradford protein assay (Bio-Rad) by measuring absorbance with the SPECTROstar Nano machine and MARS Data analysis software (BMG LABTECH) and diluted to equal concentrations with RIPA buffer. 30-50 µg of protein were run on 8 % or 10 % SDS-PAGE Tris-Glycine gels (depending on the size of the proteins) at 100V for initial stacking and then 140V for 1-1.25 hours in a Tris-Glycine running buffer. Proteins were

then transferred onto PVDF membrane using a Tris-Glycine buffered wet transfer system for 2 hours at constant 200 mA. Blots were then blocked with 5% milk in TBS-T (Tris buffered saline pH 7.4 with 0.1 % Tween). Blots were incubated with primary antibodies overnight in 5 % milk/TBS-T. The next day, membranes were washed three times with TBS-T for 5 minutes each and then incubated with secondary antibodies in 5 % milk/TBS-T for 1 hour. Blots were imaged by incubating them with the Amersham ECL western blotting detection reagent (VWR) for 1 minute and then imaging every 10 seconds for 5 minutes on the ChemiDoc XRS+ machine (Bio-Rad). ImageLab (Bio-Rad) was used for band intensity quantification.

Staining and imaging of mouse cortical neurons and human iPSC-derived neurons

Mouse cortical neurons and human iPSC-derived neurons on coverslips were fixed with 4% paraformaldehyde (PFA) for 10 minutes at room temperature, washed once with PBS, and stored in PBS at 4 °C protected from light. Fixed coverslips were then blocked and permeabilized in BP solution (PBS with 10% donkey serum and 0.3% Triton-X) for 45 minutes at room temperature. Coverslips were then incubated with primary antibodies at 4 °C overnight. The following day, coverslips were washed three times with PBS for eight minutes each. Coverslips were then incubated with secondary antibodies for one hour at room temperature, followed by three washes with PBS. For human iPSC-derived neuron Synapsin1 staining, coverslips were incubated with 300 mM of DAPI for 15 minutes, before the third wash with PBS. Excess liquid was then removed

from the coverslips and they were mounted onto VistaVision glass microscope slides (VWR) with 10 μ L of Prolong Gold Anti-Fade mounting medium (Life Technologies). For TOMM20 staining, mouse neurons were fixed with 4% PFA at 37°C for 10 min and then permeabilized with 0.5 % Triton X-100 for 10 minutes. Non-specific binding was blocked by incubation with 5 % donkey serum in PBS for 50 minutes at room temperature, followed by primary antibody incubation. The secondary antibody was added for 50 minutes at room temperature. Primary and secondary antibodies were diluted in PBS with 0.5 % BSA, 2.5 % Donkey-serum, and 0.15 % Triton X-100. After primary and secondary antibody incubation, three washing steps with PBS were performed. Then, coverslips were incubated with Phalloidin-488, for F-actin labeling, and Hoechst dye for 45 minutes at room temperature followed by three PBS washes. Coverslips were mounted onto slides using Fluoromount-G® (Southern Biotech) and were stored protected from light. Synapsin1 and BioID2 stained images were taken on the Zeiss LSM 700 confocal microscope with 63x or 40x oil objective, respectively. Mito-dsRed images were taken on the Echo Revolve microscope with a 20x objective.

Proteomic profiling of Taok2 KO mice cortical post-synaptic density fraction through LC-MS/MS

The right cortical lobes of three P21-23 *Taok2* KO mice and five P21-23 wildtype littermates were harvested and differential centrifugation was used to obtain the crude post-synaptic density fraction (Kwan et al., 2016). PSD fractionations were validated by western blot for PSD-95 and synaptophysin (data not shown). Final post-synaptic density

pellets were resuspended using 8 M urea and 100 mM ammonium bicarbonate. Protein samples were then reduced with 10 mM Tris(2-carboxyethyl) phosphine for 45 min at 37 °C, alkylated with 20 mM iodoacetamide for 45 min at room temperature, and digested by trypsin (Promega) (1:50 enzyme-to-protein ratio) overnight at 37 °C. The peptides were desalted with the 10 mg SOLA C18 Plates (Thermo Scientific), dried, and labeled with Multiplex 10-plex TMT labels (Thermo) in 100 mM triethylammonium bicarbonate, and quenched with 5% hydroxylamine before combined. 40 µg of the pooled sample was separated into 60 fractions by high-pH reverse-phase liquid chromatography (RPLC) using a homemade C18 column (200 µm × 30 cm bed volume, Waters BEH 130 5 µm resin) running a 70 min gradient from 11 to 32% acetonitrile– 20 mM ammonium formate (pH 10) at a flow rate of 5 µL/min. Each fraction was then loaded onto a homemade trap column (200 µm × 5 cm bed volume) packed with POROS 10R2 10 µm resin (Applied Biosystems), followed by a homemade analytical column (50 µm × 50 cm bed volume) packed with Reprosil-Pur 120 C18-AQ 5 µm particles (Dr. Maisch) with an integrated Picofrit nanospray emitter (New Objective). LC-MS experiments were performed on a Thermo Fisher Ultimate 3000 RSLCNano UPLC system that ran a 3 h gradient (11– 38% acetonitrile–0.1% formic acid) at 70 nL/min coupled to a Thermo QExactive HF quadrupole-Orbitrap mass spectrometer. A parent ion scan was performed using a resolving power of 120 000; then, up to 30 of the most intense peaks were selected for MS/MS (minimum ion counts of 1000 for activation) using higher energy collision-induced dissociation (HCD) fragmentation. Dynamic exclusion was activated such that MS/MS of the same m/z (within a range of 10 ppm; exclusion list size = 500)

detected twice within 5 seconds was excluded from the analysis for 30 seconds. Data were analyzed using Proteome Discoverer 2.2 (Thermo). For protein identification, search was against the SwissProt mouse proteome database (55,366 protein isoform entries), while the search parameters specified a parent ion mass tolerance of 10ppm, and an MS/MS fragment ion tolerance of 0.02Da, with up to two missed cleavages allowed for trypsin. Dynamic modification of +16@M was allowed. Only proteins with two unique peptides were used for further analysis. Differentially expressed proteins (DEPs) were calculated through Significance B outlier test using the Perseus software (Tyanova et al., 2016), and only proteins that had adj. p-value < 0.05 were considered as DEPs.

Transcriptome profiling of Taok2 KO mice cortices through RNA sequencing

Cortices from *Taok2* KO and wildtype littermates were also harvested for RNA at post-natal day 21-23, with 3 males and 3 females from each genotype. The RNA was extracted using Trizol and was sent for total RNA sequencing at the Center for Applied Genomics (TCAG). mRNA was purified using poly(A) selection to avoid contamination of ribosomal RNAs and miRNAs. All samples were run on one lane resulting in ~31-34 million of read pairs per sample. All analysis was carried out using the open-source platform Galaxy (usegalaxy.org) (Afgan et al., 2018). RNA reads were checked for good quality using the FastQC tool. The trimmomatic tool was used to identify and trim off known adaptors and remove any bases that have a Phred score of less than 20. FastQC was used again to ensure that adaptor sequences were removed and that the quality of the reads was not affected. We next used the HISAT2 alignment program for alignment of

the RNA sequences to the mouse genome GRCm38 (NCBI). On average 85% of reads from mouse samples were aligned once and 5% were aligned more than once to distinct genome locations. Moving on, the featureCounts tool was used to count the number of reads per gene using the same reference genome as the HISAT2 tool. The DESeq2 tool was used to determine the significant differentially expressed genes (DEGs) between *Taok2* WT and KO mouse cortices. Genes were considered as DEGs if they had an adjusted p-value lower than 0.05.

Gene set enrichment analysis (GSEA) of Taok2 KO mouse proteome and transcriptome

DEGs and DEPs were ranked based on the equation $-\log_{10}(\text{adj. p-value}) * \text{Ln}(\text{fold change})$. GSEA 4.1.0 (Broad Institute)(Daly et al., 2003; Subramanian et al., 2005) was used to run the GSEA preranked test. Tests were run with 1000 permutations, weighted enrichment statistics, and excluding gene sets smaller than 15 and larger than 500 genes. All other settings were kept as default. All mouse GO term gene sets without electronic GO annotations (http://download.baderlab.org/EM_Genesets/) were used for the analysis(Merico et al., 2010). Visualization of the enriched gene sets was done on Cytoscape 3.8.2 using the EnrichmentMap app and the AutoAnnotate app was used for clustering similar gene sets(Shannon et al., 2003; Bader et al., 2014; Kucera et al., 2016). All visualized gene sets had an FDR < 0.1.

Seahorse assay of in vitro mouse and human iPSC-derived neurons

Mouse cortical neurons were cultured as described above at a density of 30,000 cells/well in the Seahorse XF96 cell culture microplate. CRISPR/Cas9 KO mouse neurons were infected at DIV 7 and assayed at DIV 14. Human iPSC-derived NGN2 neurons were plated on day 4 of dox induction at a density of 50,000 cells per well in the Seahorse XF96 cell culture microplates (Agilent), pre-coated with 20 $\mu\text{g/ml}$ Laminin and 67 $\mu\text{g/ml}$ Poly-ornithine (Sigma). Mouse glia was plated on top of the neurons at a density of 25,000 cells per well, 24 hours later. Cells were used for the Seahorse assay on day 7. The day prior to the seahorse assay the Seahorse XFe96 sensor cartridge was filled with Calibrant XF solution and incubated at 37 °C (without CO₂) overnight. On the day of the assay the Seahorse XF96 microplates were washed twice with 200 μl per well of pre-warmed MST media (Seahorse XF DMEM pH 7.4 media supplemented with 1 mM sodium pyruvate, 2.5 mM GlutaMAX, and 17.5 mM Glucose). The plate was then filled with 180 μl per well MST media and incubated at 37 °C (without CO₂) for 1 hour. During the incubation, the mitochondrial stress test drugs were added to the XFe96 sensor cartridge (1 μM Oligomycin for mouse neurons and 3 μM Oligomycin for human neurons, 1 μM FCCP, and 1 μM Rotenone/Antimycin A resuspended in MST media). The cartridge plate with the drug compounds were then put in the Seahorse XFe96 analyzer for calibration. After calibration, the microplate was placed into the Seahorse XFe96 analyzer for the pre-set mitochondrial stress test protocol. Oxygen consumption rates (OCR) were recorded every seven minutes and the drug compounds were added in 21-minute intervals. Oligomycin was used to inhibit ATP-synthase to measure ATP-synthase dependant respiration, FCCP was added to decouple the inner membrane to

measure maximal respiration, and Rotenone and Antimycin A were added together to measure non-mitochondrial respiration. After the assay, microplates were frozen at -80°C overnight and cell content was measured using the Cyquant cell proliferation assay (Thermo) by measuring fluorescence with the CLARIOStar machine and MARS data analysis software (BMG LABTECH). Cellular respiration analysis was performed using the Wave software (Agilent) and OCR values were normalized to the number of cells per well.

CRISPR/Cas9 knockout in mouse cortical neurons

Mouse cortical neurons were infected at DIV7 with the pLV CMV-Cas9-T2A-EGFP (MOI 1) and pLV U6-sgRNA/EF1a-mCherry (MOI 3) lentiviruses. Cultures were allowed to recover until DIV14 and were then taken for the Seahorse assay. The GeneArt genomic cleavage detection kit (Thermo) was used to detect insertions or deletions in the targeted sites.

Measuring mitochondrial activity in shRNA knockdown mouse neurons

Embryonic age E15 C57BL6/J mouse pups were *in utero* electroporated with Taok2 shRNA and control shRNA. Electroporated mouse embryo cortices were then harvested and cultured at E18. Mouse neuron cultures were imaged at DIV5 after incubation with 2nM TMRM (Thermo). Images were analyzed on ImageJ. Soma regions were delineated and integrated density in the soma (soma area x mean intensity) was

measured. For background correction, mean background intensity was obtained from the neighbouring region.

Measuring mitochondrial activity and content in mouse cortical neurons

DIV 6 mouse cortical neurons cultured from *Taok2* WT, Het, and KO mouse embryos were incubated with 2 nM TMRM (Invitrogen, #T668) and/or 100 nM MitoTracker Green (Cell Signaling Technology, #9074P) were directly added to the conditioned medium, and incubated for 15 minutes. Cells were then imaged within 30 minutes after the incubation time. Images were loaded onto ImageJ, background mean intensity was measured from the region without TMRM and MitoTracker signals inside the cell, then the cell was delineated and the background was removed. After background correction, using the JACoP plugin the TMRM-MitoTracker signal colocalization was analyzed using Manders' correlation coefficients. For Manders' correlation coefficients, threshold values for TMRM (red channel) and MitoTracker (green channel) were set to 335 ± 55 and 640 ± 50 respectively. 16-bit wide field images were taken on a Nikon EclipseTi2 inverted spinning disk microscope equipped with 60X oil (NA 1.4) objective, an LED light source (Lumencor® from AHF analysentechnik AG, Germany), and a digital CMOS camera (ORCA-Flash4.0 V3 C13440-20CU from Hamamatsu) controlled with NIS-Elements software. The microscope imaging chamber is equipped to maintain 37 °C temperature and 5 % CO₂. Illumination, exposure and gain settings were kept the same across different conditions for imaging TMRM and MitoTracker signals.

TOMM20 staining analysis for mitochondria content in mouse cortical neurons

DIV 7-8 mouse cortical neurons cultured from *Taok2* WT, Het, and KO mouse embryos were fixed and stained for TOMM20. 16-bit Z-series images with a step size of 300 nm Images were acquired on confocal spinning disk microscope using a 60X oil (NA 1.4) objectives. Illumination, exposure and gain settings were kept the same across the conditions. The images were loaded onto ImageJ and z-projection (sum slices) for the entire cell in z-axis was performed on the confocal images. Using ImageJ, soma region was carefully delineated and total intensity, also known as integrated density, in the soma (soma area * mean intensity) was measured. For background correction, mean intensity (background mean intensity) was obtained from the neighbouring region (out of the cell). Using the following equation, we obtained the corrected values. Corrected value = total intensity in the soma – (background mean intensity * soma area).

Electron microscopy of synaptic mitochondria from mouse brain cortices

Coronal vibratome sections of the cingulate cortex (cg1 and cg2) and the prelimbic cortex (PL) of the PFC, the primary somatosensory regions S1HL, S1Fl, S1BF, and the intermediate HC were collected and prepared for electron microscopy as described in Richter *et al.* (Richter et al., 2019). Semithin sections (0.5 μm) were prepared for light microscopy mounted on glass slides and stained for 1 min with 1% Toluidine blue. Ultrathin sections (60 nm) were examined in an EM902 (Zeiss, Munich, Germany). Pictures were taken with a MegaViewIII digital camera (A. Tröndle, Moorenweis, Germany). EM images that were collected and analyzed for synapse formation on the

dendritic spines or shafts from Richter *et al.* were reanalyzed for mitochondrial morphology. Mitochondria morphology from the EM images obtained from Taok2 Wt and Taok2 KO genotypes were analyzed manually using ImageJ. based on their morphology the mitochondria are and categorized to Category 1 - Normal mitochondria with well stacked Cristae, Category 2 - mitochondria with enlarged Cristae, Category 3 - mitochondria with condensed Cristae.

Mito7-dsRed puncta analysis in human iPSC-derived neurons

TAOK2 KO, A135P and wildtype human iPSC-derived NGN2 neurons were transfected with 0.8 µg of Mito7-dsRed (Addgene #55838) and 0.2 µg of pCAG-Venus at day 5, with 2 µl of Lipofectamine 2000 (Thermo). Venus was used to trace neuron projections. 10 neurons per genotype from two separate neural inductions were used for analysis. Imaging settings were kept the same between images and Mito7-dsRed images were analyzed at the same threshold. Dendrites were traced using ImageJ and the measure tool was used to quantify the size of the puncta within the traced region

Correlation of 41 ASD-risk gene PPI networks

Corrplot (R package) was used to create the correlation plot. The normalized biotinylation score to the bait protein was used to calculate the correlation between ASD-risk gene PPI networks. The Silhouette and Within cluster sum of squares methods were used to calculate the optimal kmeans number for clustering. Genes were ordered by hierarchal clustering.

Cell type/DEG/ASD gene list enrichment analysis

Human cell type gene expression and ASD DEGs and ASD gene lists were obtained from their respective publications (Sanders et al., 2015; Nowakowski et al., 2017; Yuen et al., 2017; Feliciano et al., 2019; Ruzzo et al., 2019; Velmeshev et al., 2019; Satterstrom et al., 2020; Wilfert et al., 2021). For the enrichment analysis we used the Fisher exact test comparing each gene list with the shared ASD-risk gene PPI network in the mouse brain background protein list, which was used for pathway enrichment analysis. P-values and ODDs ratios were calculated for each comparison. To account for multiple comparisons, Bonferroni correction thresholds were calculated as $p = 0.05$ divided by the number of comparisons.

Data representation and figure generation

Networks and gene set enrichment maps were created on Cytoscape v3.8.2. Graphs were created on GraphPad Prism 7. Representative electrophysiology traces were extracted onto CorelDRAW. Microscopy images were prepared using ImageJ. Dot plots, correlation plots, and heat maps were created on R Studio. Flowcharts were created on and exported from BioRender.com (SD235B8ORF, KW235KT7TM, RZ235KTA0S). Final figures were organized and created using Adobe Illustrator CC.

Data and materials availability

Correspondence and construct material requests to Karun K. Singh. Mass spectrometry datasets consisting of raw files and results files with statistical analysis to identify PPI

networks or significant DEPs will be deposited into ProteomeXchange through the Proteomics Identification Database. Individual PPI networks and shared ASD-risk gene PPI network map protein lists and enriched pathways can be found in Tables S1-S9. The Mouse_Human_Reactome and Mouse_GO_ALL_no_GO_jea gene sets used for overrepresentation and gene set enrichment analyses were downloaded on 13 August 2021 from http://download.baderlab.org/EM_Genesets/(Merico et al., 2010). RNA sequencing raw sequence files and results files with statistical analysis to identify significant DEGs will be deposited into the Gene Expression Omnibus. ASD proband variant information and clinical scores are available through the MSSNG database (research.mss.ng)(Yuen et al., 2017) and the associated Metabase (data-explorer.mss.ng), respectively.

Statistics analysis

Data are expressed as mean \pm s.e.m, except the clinical analysis which is shown as a box and whisker plot showing the minimum, median, and maximum scores. A minimum of three biological replicates were used for all experiments, where separate HEK cell transfections, iPSC dox-inductions, mouse neuron cultures, or littermates are considered as individual biological replicates. All statistical analysis was done on GraphPad Prism 7. All comparisons were assumed to be parametric, except for the clinical score analyses. ROUT's outlier test was used to identify possible outliers, with a Q value of 0.1 %. For statistical analysis unpaired t-test, or One-Way ANOVA and Two-Way ANOVA with *post hoc* Holm-Sidak tests were used to compare all experimental

conditions to the control condition. All unpaired t-tests were two-sided, except for the one-sided t-test used for identification of BioID2 prey proteins. Clinical scores were assumed to be non-parametric and the Kruskal-Wallis H test with post hoc Dunn's test was used to compare all groups to each other. Any variation from the described statistical analyses is described and explained in the figure legends. The p-values are defined in the figure legends and $p < 0.05$ are considered statistically significant.

3.4 RESULTS

3.4.1 Development of a neuronal proximity-based proteomic system to identify PPI networks

To identify the PPI networks of 41 ASD-risk genes, we used mixed mouse cortical neurons and glia co-cultures infected with lentiviral constructs expressing BioID2 fusion proteins (pLV-hSyn-tGFP-P2A-POI-13xLinker-BioID2-3xFLAG) (Figure S1A). Neuron-specific expression of BioID2-tagged proteins was driven by a human Synapsin1 promoter, and neuron/glia co-cultures were used to promote synaptic maturation. A 13x Gly-Ser linker sequence was used to join proteins-of-interest (POIs) with the BioID2-3xFLAG, which increases the range of biotinylation to include direct and indirect interactors and proteins in the close proximity to the POI. To help monitor transduction efficiencies, TurboGFP (tGFP) was coexpressed with BioID2 fusion proteins in a bicistronic system employing a P2A “self-cleaving” peptide. We used a Luciferase-P2A-BioID2-3xFLAG construct as a negative control (Figure S1B). Since lentiviral (LV) systems can accommodate larger gene sizes than adeno-associated virus (AAV), we were

able to perform BioID2 on larger proteins such as SHANK3, GRIN2B, MET, SYNGAP1 and CNTNAP2.

Embryonic age 16-17 (E16-17) mouse pup cortices were harvested and cultured until days in vitro (DIV) 14, then infected with the BioID2 fusion constructs by using a lentivirus with an average transduction efficiency of 70% (Figure 1A). Biotin was added on DIV17 and cells were lysed after 18 hours on DIV18 to allow maximal biotinylation time. To reduce variability between mass spectrometry runs, TMT10plex isobaric-labeling was used to combine at least 3 biological replicates per gene. One additional technical TMT-labeling replicate of luciferase control sample chosen at random was used to account for differences in labeling. Two statistical cut-offs were used to identify positive hits for the PPI networks of each POI: Biotinylated proteins in the POI sample with 1) a significant increase in Log₂ abundance compared to the luciferase control (Student's t-test, $p < 0.05$) (Uezu et al., 2016) and 2) that were significant outliers when accounting for the overall protein abundance compared to the protein abundance ratio between the POI and control samples (SigB $p < 0.05$) (Cox and Mann, 2008). Protein abundances were normalized between biological replicates based on the sample with the highest total protein abundance. To reduce variability between each viral transduction, flow cytometry was used to determine the total abundance of GFP in the infected neurons between samples. The abundance levels of samples that had less total GFP (area under the curve in GFP intensity histogram) than the luciferase control was normalized by the factor required to equalize the total GFP of the POI sample to the luciferase control sample. To account for false positive hits due to variability in TMT-labeling between

samples, the ratio of protein abundances between the luciferase control technical TMT-labeling replicates were used as the minimal required ratio between POI and control sample abundances. Proteins that did not have abundance ratios (POI/Luciferase control) higher than this minimal ratio were considered false positives and removed from further analysis. Further, to promote high efficiency infections, we created an optimized lentiviral production protocol to produce high-titer virus for small and large risk genes (Figure 1B). This BioID2 screen was used in five specific experimental outputs: to identify 1) shared molecular pathways, 2) the impact of patient genetic variants on the PPI network, 3) correlation between ASD-risk genes, 4) enrichment of ASD-relevant cell types in the shared PPI network map, and 5) correlation of clinical phenotypes with the ASD-risk genes (Figure 1A).

To validate the BioID2 screening system, we used the well characterized excitatory synapse protein DLG4 (PSD95). Neurons expressing PSD95-BioID2 displayed punctate localization of BioID2-3xFLAG fusion proteins and biotinylated proteins around the dendrites, suggesting appropriate synaptic expression and biotinylation (Figure S1C). The Luciferase-P2A-BioID2 control showed non-specific localization and biotinylation throughout the neuron, which is expected (Figure S1C). We identified 74 proteins that interact with PSD95, and Reactome pathway enrichment revealed neurotransmitter receptors and glutamatergic synapses, as expected (Figure S1D and Table S1). Enriched pathways also include less directly associated networks, such as GABAergic synapses, Rho GTPase signaling, and Wnt signaling. Comparison of our PSD95 PPI network with the previously published PPI networks for PSD95, from *in vivo* mouse BioID and *in vivo*

mouse tandem affinity purification (Fernández et al., 2009; Uezu et al., 2016), revealed 25 shared proteins between all three datasets (Figure S1E), highlighting that our BioID2 system captures relevant networks. The distinct and partially-shared proteins from the other systems suggest differences between proximity-labeling and affinity purification methods and/or *in vitro* and *in vivo* approaches.

Cortical neurons are a major cell type associated with ASD (Satterstrom et al., 2020); however, scalable BioID labeling approaches have been done primarily in cell lines, such as HEK293 cells (Go et al., 2021). To determine the necessity and importance of using neurons for the BioID2 screen of ASD-risk genes, we performed BioID2 in HEK293 cells using PSD95, and a subset of ASD risk genes including *ETFB*, *SPAST*, *STXBPI*, *SYNGAP1*, and *TAOK2* (Figure S2 and Table S2). The PSD95 PPI network from HEK293 cells showed enrichment of many pathways, including EGF- and NTR-receptor signaling and cell junction organization, but there was a complete absence of synaptic pathways (Figure S2A). Furthermore, BioID2 of all six ASD risk proteins in HEK293 cells revealed a significant loss of protein interactions localized in neuron-specific compartments, and large differences in the PPI network between HEK293 cells compared to mouse neurons (Figure S2B-S2G, and Table S3). The differences in the plasma membrane proteins between the HEK293 cells and the neurons identified through BioID may be due to the differences in transient versus stable expression of the BioID constructs. Transient expression can cause a higher level of expression and result in overrepresentation of some biotinylated proteins in the sample. This could thereby result in differences in the membrane proteins between HEK293 cells and neurons. However, it

is possible that proteins found in the membrane of HEK293 cells may not be present in neurons and vice versa. While HEK293 cells yield interaction networks for ASD risk genes, they may not have relevance to pathways associated with brain-specific pathophysiology of neurodevelopmental disorders.

To further validate the specificity of our neuron-specific BioID2 screening system, we targeted proteins associated with compartments (Fazal et al., 2019), including microtubules (MAP2C), the endoplasmic reticulum network (CANX), plasma membrane (PDGFR transmembrane domain), trans-Golgi apparatus (TGOLN), the presynaptic terminal (SNCA), and the nuclei (MECP2). Cellular compartment analysis of each PPI network revealed enrichment of the compartments expected for MAP2C, MECP2, CANX, PDGFR-TM domain, and TGOLN (Figure S3, Figure S4D, Table S1 and Table S4). SNCA did not have a strong enrichment of presynaptic compartments; however, it did identify enriched pathways involving axons, growth cones and the synapse (Figure S3E). BioID2 of MECP2, a nuclear protein, indicated localization to the nucleus (Figure S4A) and interaction with proteins enriched in nucleus-specific pathways, such as transcription regulation and mRNA splicing (Figure S4B). The MECP2 PPI network in mouse neurons had differences in protein interactions compared to the MECP2 network in HEK293 cells, but there was no enrichment for neuron-specific compartments (Figure S4C and S4D). The difference in identified proteins suggests that mouse neurons have differing MECP2 interactions that are localized to the nucleus. Further, the PPI network of MECP2 did not include some of the known protein interactions in mouse neurons (e.g., ATRX, CREB1, SIN3A, NCor, and TET1), suggesting that our system may not be

optimized for nuclear proteins, possibly due to the presence of highly biotinylated endogenous proteins. Alternatively, the inaccessibility of nuclear proteins that may not easily separate from the DNA or solubilize in the lysis buffer, the availability of exposed lysine residues for biotinylation on these proteins, and the possible improper localization of MECP2-BioID2 fusion protein may prevent identification of known interactors. The enrichment of proteins specific to each compartment provides additional validation that the BioID2 screen in mouse cortical neurons can provide relevant PPI networks.

3.4.2 Identification of a shared PPI network map and common pathways of 41 ASD-risk genes in mouse cortical neurons

To develop a shared PPI network map for ASD risk genes, we selected 41 ASD-risk genes that encode proteins with a range of molecular functions, including regulation of phosphorylation and ubiquitination, enzymatic control of metabolism, protein regulation and transport, and synapse formation and function (Figure 1B). These genes were chosen from a combined list of ASD-risk genes from the SPARK, SFARI category 1, 2, and syndromic gene lists and previous sequencing studies (Sanders et al., 2015; Yuen et al., 2017; Feliciano et al., 2019; Ruzzo et al., 2019; Satterstrom et al., 2020; Wilfert et al., 2021). The final combined list was filtered for size limitations of the lentivirus, in our construct, genes that were larger than 5000bp were excluded, however we included some large genes (>4kb) such as SHANK3 and SYNGAP1, allowing us to examine the PPI network of proteins from a range of sizes. All genes that were chosen for the screen have cytoplasmic functions and nuclear genes were not selected because it has

previously been shown there is a separation in function between nuclear gene regulating proteins and cytoplasmic neural communication proteins (Satterstrom et al., 2020). The final list of genes was based on spanning a wide range of functions, including those that had no known function, and our ability to obtain or create the cDNA of the gene. For each gene, the human cDNA was cloned into a BioID2 lentiviral backbone and protein expression was confirmed with western blotting (Figure S5). All BioID2 fusion constructs were found to be the expected size through western blotting; however, some constructs showed a second larger size protein due to lower P2A efficiency or increased cleaved BioID2-FLAG (lowest band) due to increased degradation (Figure S5). We identified the individual PPI networks and enriched Reactome pathways, biological processes and cellular compartments for each of the 41 ASD-risk genes in mouse cortical neurons. These data can be found in Table S1 and Table S5, and are meant to be a resource for the research community. Validation of randomly chosen BioID protein hits verified the novel interactions between TAOK2 and Fbxo41, which coincides with a recent publication showing interaction with other innate immune regulators (Pennemann et al., 2021), and between the mitochondrial citrate synthase protein and Syngap1 (Figure S5B and S5C)

The 41 ASD-risk gene PPI network consisted of 1109 proteins (41 ASD bait proteins and 1068 prey proteins) and 2349 connections. Half of the identified prey proteins were shared between 2-15 ASD bait proteins (489 prey proteins and 1770 connections), and of these, 15 prey proteins were shared between at least 10 different ASD bait proteins. Every ASD bait protein shared at least 4 shared prey proteins with one

other ASD bait protein, with up to 38 shared interactions between DLG4 and SYNGAP1 (Figure 2A). The PPI network of 31 out of the 41 ASD bait genes included at least 1 other ASD bait protein. Reciprocal identification was observed between DLG4 and CDKL5, SYNGAP1, GRIA1, or GRIA2 and between GRIA1 and GRIA2. The most identified ASD bait proteins were GRIN2B, PPP1R9B, GRIA2, and KCNQ2. Conversely, BioID2 of TAOK2, CDKL5, DLG4, LRRC4C, and SYNGAP1 identified the most ASD bait proteins, suggesting high connectivity between a subset of ASD bait proteins. To determine the utility of creating an ASD PPI network in neurons, we compared our results with physical interactions between the 41 ASD bait genes extracted from the STRING database (greater than or equal to medium confidence, 0.4). Our BioID2 shared 41 ASD-risk gene PPI network had 245 connections (where each connection represents at least 5 shared protein hits) between 36 of the 41 ASD bait proteins. The STRING database had 33 direct interactions between 23 of the bait proteins, revealing a near 50-fold increase in the number of connections within our 41 ASD-risk gene PPI network (Figure 2A and 2B) through inclusion of indirect and close-proximity proteins. Current databases, such as STRING, are primarily derived from non-neuronal sources using gene co-expression or direct interaction data (Lage, 2014), which do not capture proteins within close proximity that have significant importance to the function of proteins-of-interest. However, our PPIs were identified in neuronal cells and include both direct interacting proteins, shared interacting proteins and close proximity proteins that highlight potential connections or protein pathways missed by traditional methods.

The most significant pathways in the shared 41 ASD-risk gene PPI network involve synaptic transmission, demonstrating that our system can identify the most frequently identified pathways in ASD pathophysiology (Figure 2C and Table S6). Other enriched pathways included TCA cycle and mitochondrial activity, Wnt signaling, potassium channel activity, and MAPK signaling (Figure 2C and Figure S6A). These enriched pathways suggest that synaptic function plays a core role among non-nuclear ASD risk proteins, but it is not the only pathway involved between the 41 genes. The majority of the shared ASD-risk PPI network localized to specific cellular compartments including axons, dendrites and synapses (Figure S6B and Table S6), while the majority of biological processes involve synaptic signaling and organization, and protein transport (Figure S6C and Table S6). Shared pathways in the ASD-risk gene PPI network reflect the major role of synaptic dysfunction in ASD, but also highlight that other, less well-studied pathways are important contributors to convergent ASD pathology.

3.4.3 The shared PPI network map identifies the tricarboxylic acid (TCA) cycle and pyruvate metabolism as a common signaling pathway in ASD

One rationale for constructing a PPI network map with ASD-risk genes was to identify novel or poorly characterized convergent signaling mechanisms. In this regard, one of the top pathways we identified was the TCA cycle and pyruvate metabolism (mitochondrial/metabolic processes), implicating dysregulation in mitochondrial function and cellular metabolism. This pathway has been associated with a few ASD associated genes, such as *Fmr1* and *Mecp2* (Shulyakova et al., 2017; Licznernski et al., 2020; Bülow

et al., 2021), but the mechanisms are not well understood, and it is unknown whether other ASD risk genes regulate mitochondrial/metabolic processes. Interestingly, previous ASD clinical studies have identified abnormal mitochondrial function in patient lymphoblastoid cells (Rose et al., 2014, 2018; Shen et al., 2019a; Frye, 2020), but whether this occurs in mammalian brain cells is unknown. TCA cycle and respiratory electron transfer chain protein Reactome pathway was highly enriched in the shared ASD-risk gene PPI network map (adj. p-value = 3.14×10^{-12}), even without the PPI network for the mitochondrial protein ETFB (adj. p-value = 1.35×10^{-7}) (Table S6). 28 out of 41 ASD-risk genes were found to be interacting with at least one TCA cycle and pyruvate metabolism associated protein (Figure 3A). Citrate synthase (CS), which is involved in turning acetyl-CoA into citrate early in the TCA cycle, was found to interact with eight ASD bait proteins (ERBIN, MET, NRXN1, SHANK3, SPAST, STXBP1, SYNGAP1, TAOK2). The TCA cycle and pyruvate metabolism are essential for proper cellular respiration. Therefore, we investigated this finding by focusing on a gene in our screen that was not previously associated with mitochondrial and metabolic processes in the brain, *TAOK2*, a gene in the 16p11.2 deletion/duplication region associated with ASD (Calderon de Anda et al., 2012; Ultanir et al., 2014; Yadav et al., 2017; Richter et al., 2019). We measured cellular respiration using live-cell metabolic assays in a *Taok2* knockout (KO) mouse model, which we previously demonstrated has deficits in synapse formation and function (Richter et al., 2019). *Taok2* heterozygous knockout (Het) cultured mouse cortical neurons showed a significant increase in maximal respiration, proton leak, non-mitochondrial respiration, and spare respiratory capacity, and a decrease in ATP coupling

efficiency (Figure 3B and 3C, and Figure S7A-S7D) compared to wildtype (WT) neurons. These changes suggested the presence of less functional mitochondria, which was corroborated by proteomic analysis of post-synaptic density fractions isolated from *Taok2* WT and KO mouse cortices (Figure S7E). *Taok2* KO mice PSD fractions had significant downregulation of proteins involved in synaptic function and activity, and also in respiratory ETC complex proteins (Figure S7F and Table S7). Analysis at the transcriptome level also revealed reduced mRNA levels of mitochondrial membrane proteins in *Taok2* KO mouse cortices (Figure S7G and S7H, and Table S7), coinciding with the reduced protein levels of mitochondrial proteins (Figure S7E and S7F). Further investigation revealed that *Taok2* Het and KO neurons have a reduced proportion of Teramethylrhodamine (TMRM)-stained mitochondria (Figure 3D and 3E, and Figure S7I), which only stains active mitochondria with an intact membrane potential. These mitochondria had an overall increase in the amount or size of mitochondria labeled by TOMM20, an outer mitochondrial membrane protein involved in translocation of mitochondrial proteins into the mitochondria (Figure 3F and 3G). These data implicate dysregulated mitochondria in the absence of *Taok2*; therefore, we examined the morphology of mitochondria *in vivo* from electron microscopy (EM) images taken from WT and *Taok2* KO mouse cortical excitatory neurons (Richter et al., 2019). We found that *Taok2* KO mouse neurons had altered mitochondrial morphology with a reduction in category 1 and 3 mitochondria, which show more typical mitochondria morphology, and an increase in category 2 mitochondria at their synapses (Figure 3H and 3I) (Da Costa et al., 2018). Category 2 mitochondria have enlarged non-contiguous mitochondrial

cristae(Da Costa et al., 2018), which can cause reduced oxidative phosphorylation and prevent proper translation and insertion of inner membrane proteins(John et al., 2005; Schmidt et al., 2010). We extended our mouse studies to examine whether *TAOK2* regulates mitochondrial/metabolic processes in human induced pluripotent stem cell (iPSC)-derived NGN2-neurons. We used CRISPR/Cas9 to generate isogenic *TAOK2* homozygous KO and heterozygous knock-in *TAOK2 A135P* iPSC lines. A135P is a *de novo* missense variant which we previously demonstrated renders *TAOK2* as kinase dead(Richter et al., 2019). We generated human neurons through direct differentiation of iPSCs via NGN2 overexpression and found altered cellular respiration in *TAOK2* KO neurons (Figure S7J-S7L) similar to mouse neurons, and significant increases in the spare respiratory capacity of *TAOK2* KO and *A135P* neurons (Figure S7M). Human neurons transfected with Mito7-DsRed also displayed an increase in mitochondrial puncta size in *TAOK2* KO and *A135P* neurons, suggesting an increase in the number or size of the mitochondria, similar to that observed in the mouse neurons (Figure S7N and S7O). To determine if these changes were due to long-term developmental deficits caused by loss of *TAOK2* function, we used acute shRNA knock-down through *in utero* electroporation and found that *Taok2* knock-down in cultured mouse neurons caused decreased mitochondrial membrane potential (Figure S8A-S8C) similar to that detected in the knockout mice (Figure 3D and 3E). Taken together, using *TAOK2* as a validation gene from the identified mitochondrial/metabolic PPI network, we determined that mouse and human models with disruption of *TAOK2* have altered cellular respiration, likely caused by altered activity, size and number of mitochondria.

To determine if other ASD risk genes converging on the mitochondrial/metabolic network regulate cellular respiration, we used the CRISPR/Cas9 system to knock out *Syngap1*, *Taok2*, and *Spast*. The CRISPR/Cas9 knockout of *Taok2* in mouse neurons was used as a validation of the system and to determine if the effect on mitochondrial cellular respiration is due to acute or chronic loss of *Taok2*. We also targeted *Etfb* and *Rheb*, which are both ASD risk genes known to localize to the mitochondrion or regulate neuronal mitochondrial function (Yang et al., 2021). Combined gRNAs against BFP and Luciferase were used as a negative control (Hart et al., 2015; Richardson et al., 2016), and we used 1-3 gRNAs targeting different genomic regions of the ASD-risk genes (Figure S8D). Mouse cortical neurons were infected with Cas9-EGFP and gRNA-mCherry lentiviral constructs. Western blots of neurons infected with *Taok2* gRNAs and Cas9 showed decreased expression by approximately 50%, suggesting a partial knockout (Figure S8E). CRISPR/Cas9 knockout of *Etfb*, a subunit of riboflavin required for proper electron transfer in the ETC, showed increased basal and maximal respiration, proton leakage, and no change in ATP synthase-dependent cellular respiration (Figure 3J and 3K, and Figure S8F and S8G). These changes may correspond to increased cellular respiration to counteract faulty ETC electron transfer. Mouse neurons with CRISPR knockout of *Taok2*, *Syngap1*, and *Rheb* also showed significant or trending changes in many aspects of cellular respiration (Figure 3J and 3K, and Figure S8F-S8I). CRISPR KO of *Spast* did not cause significant changes in cellular respiration, suggesting that it does not play a role in cellular respiration. However, it is possible that *Spast* could be involved in other aspects of the TCA cycle and pyruvate metabolism, such as glycolysis.

The increase in basal respiration in *Taok2*, *Syngap1*, and *Etfb* KO neurons (Figure S8F) may be indicative of an acute effect, where altered cellular respiration has not yet reached homeostasis within the neuron (Ruggiero et al., 2021). CRISPR/Cas9 knockout of *Ap2s1*, *Gria1*, *Ppp1r9b*, and *Ppp2r5d*, which did not have mitochondrial or metabolic proteins in their PPI network showed no changes in cellular respiration based on the seahorse assay (Figure S8J-S8L). These findings suggest that a subset of ASD risk genes regulate cellular respiration in neurons, and highlight the relevance of TCA cycle and pyruvate metabolism pathways in the developing brain as a risk factor for ASD when dysregulated.

3.5 Discussion

ASD is a heterogeneous group of neurodevelopmental disorders that are largely caused by genetic variants in multiple risk genes (Gaugler et al., 2014; Iossifov et al., 2014; Robinson et al., 2016; Weiner et al., 2017). A long-standing question in the field is how different risk genes contribute to ASD, and whether there are convergent signaling mechanisms that explain how a multitude of genes lead to a common, albeit heterogeneous, developmental brain disorder. Specific disease cell types or signaling pathways have been proposed as convergent mechanisms in ASD (Voineagu et al., 2011; O’Roak et al., 2012b; Chang et al., 2015; Velmeshev et al., 2019; Ramaswami et al., 2020), but the bulk of these data are based on RNA expression, which does not take into account signaling at the protein level. To address this gap, we devised an *in vitro* neuron-specific proteomic screen to identify individual and shared PPI networks between 41

ASD-risk genes. Our screen identified links between risk genes and multiple convergent signaling pathways.

While other approaches for identifying PPI networks exist, such as Y2H or affinity purification coupled with mass spectrometry in cells lines, these methods can miss weak and transient interactions and signaling networks specific to neurons (Sakai et al., 2011; Corominas et al., 2014; Li et al., 2015). Our use of BioID2 for the 41 ASD-risk genes revealed shared protein interactions which include direct and indirect interactions and close-proximity proteins between ASD-risk proteins in a neuronal cell type, providing detailed insight into the relationship between the 41-risk proteins. However, it is not possible to distinguish between proximal proteins or direct interaction proteins without method such as co-immunoprecipitation or co-localization staining, especially due to the larger biotinylation area with a 35nM radius created by the 13x linker, which increases biotinylation of proteins within close proximity of the POI. Regardless, even with the increased radius of biotinylation and long 18-hour incubation, post-synaptic density proteins DLG4, GRIN2B, GRIA1 had protein-specific interactions that were related to their known function (Supplementary Table 1). Furthermore, unlike the Y2H or affinity purification studies we used single canonical isoforms of each gene and therefore some PPI networks may not encompass the full scope of possible interactions in the neuron. Further, the mouse system possesses glial cells required for synaptic maturation, and it is scalable; therefore, the system could be used to screen hundreds of genes. Some caveats of BioID2 include possible biotinylation inefficiencies, protein function impairments, and protein biotinylation selection biases, however, newer proximity-

labeling tools could be used to extend the identification of PPI networks (Branon et al., 2018; May et al., 2020; Minde et al., 2020; Niinae et al., 2020).

It should be noted that our cell culturing system utilized a neuron-glia co-culture which allowed for increased maturation of the mouse cortical neuron. PPI networks identified in this co-culture will not be able to identify interactions induced by contact between glia and neurons. A possible alternative is to use membrane insert cell culture plates that allow separated growth of glia and neurons, while still allowing exchange of factors between glia and neurons required for survival and maturation. Furthermore, these neurons were grown in 5% CO₂ and approximately 21% O₂ at 37 degrees in high humidity, which is not reflective of the early neonatal brain (Nalivaeva et al., 2018). At early developmental time points the brain is a hypoxic environment that gradually receives more oxygen as vascularization increases, therefore future studies should determine the PPI networks of ASD-risk genes in cells cultured at lower oxygen concentrations.

Previous genetic screening platforms have identified shared pathways between ASD-risk genes. CRISPR/Cas9 knockout screens have identified cell types and processes associated with groups of ASD-risk genes (Cederquist et al., 2020; Jin et al., 2020; Willsey et al., 2021). Since these knockout screens disrupted genes early in development, this may skew results towards neurogenesis deficits. Our BioID2 screen complement CRISPR/Cas9 approaches, given that they can be used to study earlier or later time points, and can be used to study disease-relevant variants. BioID2 can also help to understand the function or role of poorly characterized ASD risk genes using our PPI

network pipeline and statistical cut-offs, where most previous studies rely on known compartment localization. Future studies could also be used to study changes in disease-relevant PPI networks in genetic mouse models or patient-derived iPSC neurons and organoids. Since changes in protein interaction complexes or synaptic networks in multiple ASD mouse models have been observed (Sakai et al., 2011; Brown et al., 2018), this suggests that core ASD networks can reveal risk gene clusters or identify hub genes.

One of the main findings from our study is the identification of multiple convergent and shared pathways between 41 ASD-risk genes that are non-nuclear, which fall into categories pertaining to synaptic transmission, TCA cycle and mitochondrial activity, Wnt signaling, potassium channel activity, MAPK signaling, and other specific signaling pathways. Synaptic transmission and function is widely known in ASD pathophysiology, and Wnt and MAPK signaling have also been disrupted in ASD patient cell lines (Gazestani et al., 2019). We focused on validating the TCA cycle and mitochondrial activity pathways because its dysfunction is indirectly associated with neurodevelopmental disorders (Iwata et al., 2020) and our screen identified many uncharacterized ASD-risk genes associated with this pathway (Figure 3A). Clinical studies have found mitochondrial and metabolic dysfunction or changes in metabolites in primary lymphocytes or brain tissue in individuals with ASD (Frye and Rossignol, 2011; Anitha et al., 2012; Rossignol and Frye, 2012; Frye et al., 2013; Rose et al., 2014, 2018; Wang et al., 2016b; Kurochkin et al., 2019), but whether this is direct or indirect is not known. A mouse model expressing an mtDNA variant was shown to display autism associated behavioral deficits (Yardeni et al., 2021), but the variant is weakly associated

with ASD. Some ASD associated syndromic disorders, co-morbid disorders and genetic ASD models have shown deficits in mitochondrial and metabolic processes, however the specific proteins involved were unknown (Ebrahimi-fakhari et al., 2016; Shulyakova et al., 2017; Fernandez et al., 2019; Jagtap et al., 2019; Li et al., 2019; Shen et al., 2019b; Kanellopoulos et al., 2020; Licznanski et al., 2020; Bülow et al., 2021; Madison et al., 2021; Menzies et al., 2021). Our findings indicate that TCA cycle and mitochondrial activity proteins are interacting with multiple ASD-risk genes, including genes that were not previously connected to metabolic processes. However, we also identified two ASD-risk genes (*RHEB* and *BCKDK*) that had been previously implicated directly in metabolic processes (Heinemann-Yerushalmi et al., 2021; Yang et al., 2021). This highlights that our screen can identify relevant protein interactions and may even suggest a more direct connection between mitochondrial/metabolic processes and some genetic models (*CDKL5* and *KCTD13*) (Jagtap et al., 2019; Madison et al., 2021). However, whether these proteins are interacting outside or inside of the mitochondria due to dual localization and shuttling, is currently unknown and requires further investigation.

Our CRISPR/Cas9 KO studies revealed that multiple ASD-risk genes are important for proper cellular respiration. Interestingly, these genes were all found to interact with citrate synthase (Figure 3A), suggesting that upstream or downstream regulation may occur between ASD risk genes and TCA cycle function. Deficits in the TCA cycle can cause overreliance on glutaminolysis to produce energy and cause a decrease in synaptic vesicle glutamate levels (Fendt and Verstreken, 2017; Divakaruni et al., 2018; Namba et al., 2021). This shift may help explain the deficits in synaptic

transmission in neurons with disruption of synaptic ASD-risk genes, such as *Syngap1* and *Taok2*. Interestingly, the abnormal mitochondria morphology and numbers in neurons lacking *Taok2* may be related to its function in tethering the endoplasmic reticulum to microtubules, where ER contact regulates mitochondrial fusion and fission, suggesting another possibly link between mitochondrial function and TAOK2 (Abrisch et al., 2020; Nourbakhsh et al., 2021).

The shared PPI network provides an important link between metabolic processes and ASD pathology. These data underscore the value of using PPI networks to map ASD-risk gene connectivity, and to pinpoint which risk genes are involved in convergent mitochondrial/metabolic dysregulation in ASD.

3.6 CONCLUSION

In conclusion, our neuron-specific 41 ASD-risk gene PPI network map demonstrates that protein interactions and their associated pathways and sub-cellular compartments are relevant to ASD disease pathology, and are missing from transcriptome-based approaches. Our approach is scalable and to our knowledge, represents one of the largest protein network mapping studies for ASD risk genes. This resource containing the individual PPI networks of 41 ASD-risk genes will be valuable for future in-depth study of the genes, and has the potential to grow larger with PPI networks of additional risk genes.

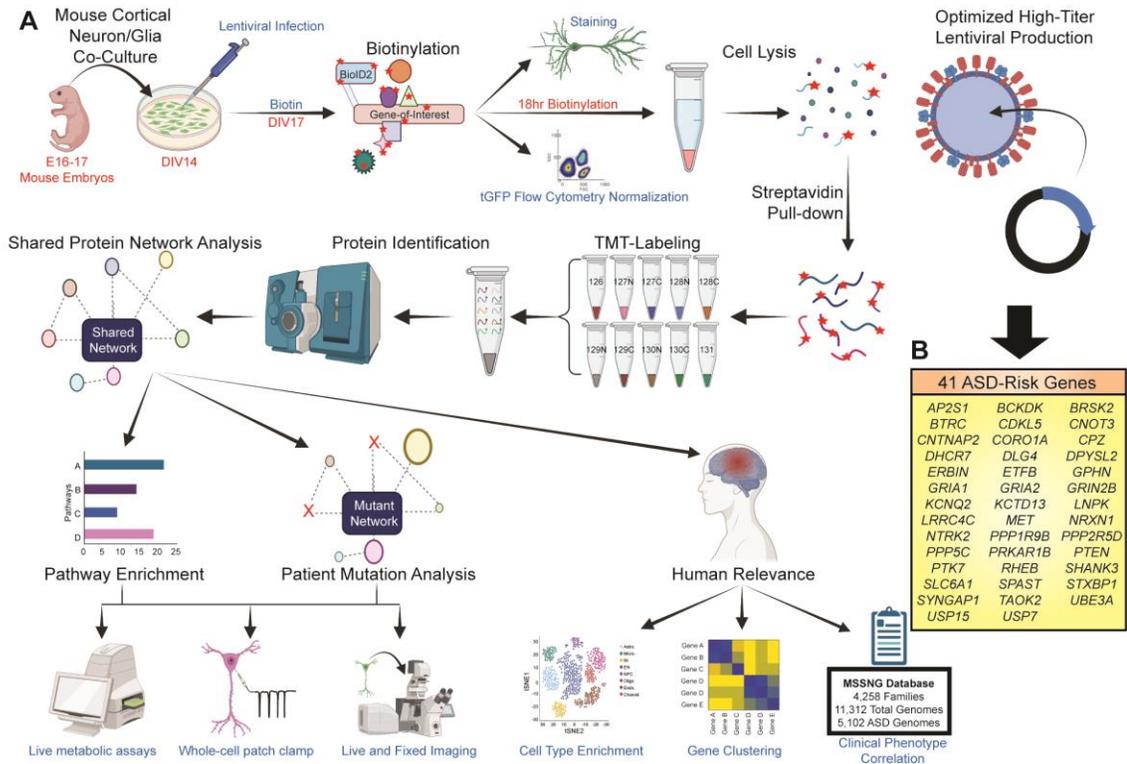


Figure 1. Development of a proximity-based proteomics screen to identify neuronal PPI networks for ASD-risk genes. (A) Workflow of neuron-specific BioID2 screen for identification of ASD-risk gene PPI networks. (B) List of 41 ASD-risk gene used in the BioID2 screen

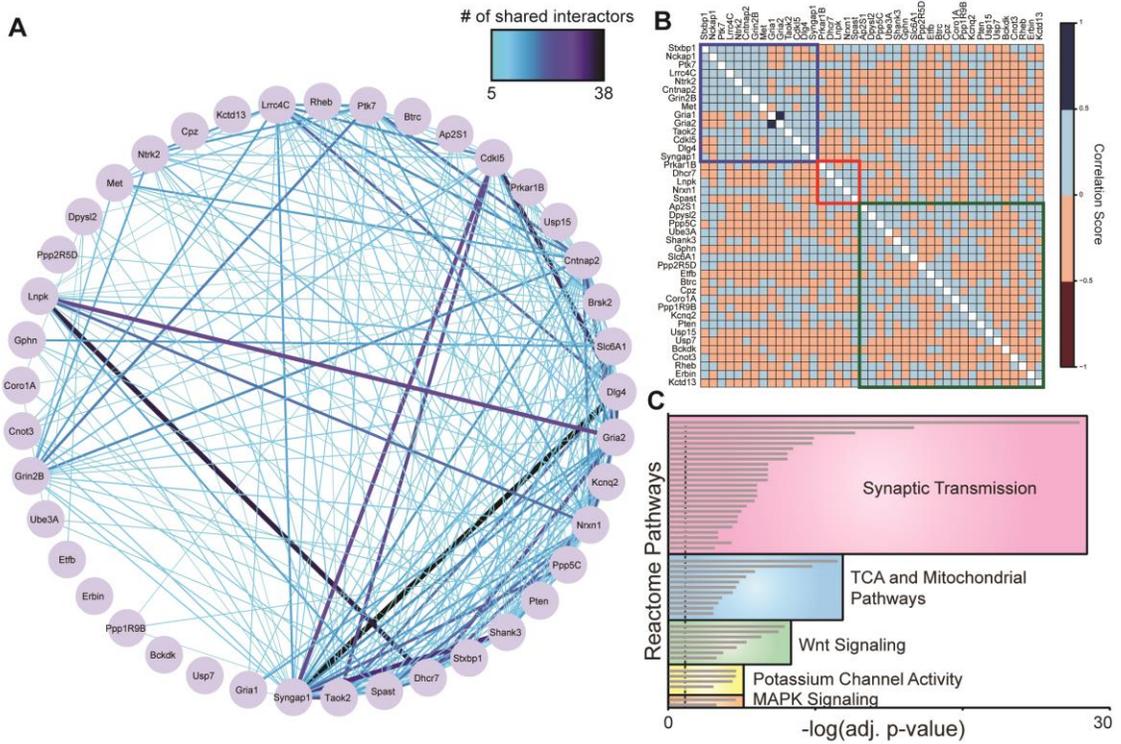


Figure 2. Neuron-specific PPI network map identifies convergent signaling

pathways between 41 ASD-risk genes. (A) Shared PPI network map of 41 ASD-risk

bait genes. Large red nodes represent bait proteins. Color and increased thickness of connecting lines represent number of interactions (direct or shared prey proteins) between

bait genes. (B) Correlation plot of 41 ASD-risk genes through individual PPI networks.

Genes were ordered by hierarchical clustering and clustered using kmeans (k = 3). (C)

Top 50 enriched Reactome pathways of the shared 41 ASD-risk gene PPI network map.

Individual pathways are grouped by functional similarity (g: Profiler, Bejamini-Hochberg

FDR adj. $p < 0.05$).

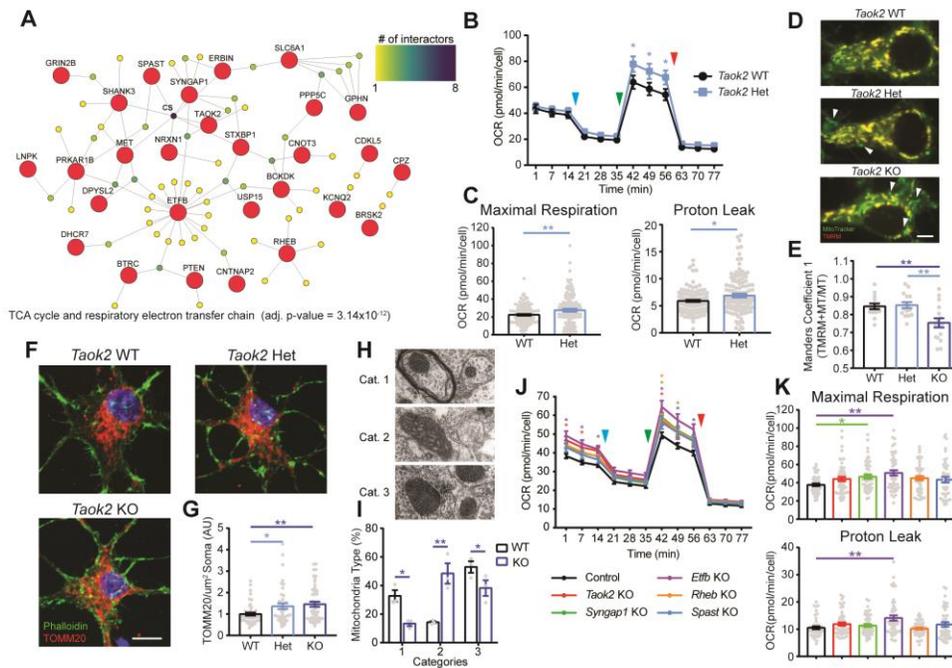


Figure 3. Regulation of cellular respiration and mitochondrial activity by ASD-risk

genes. (A) Network of TCA cycle and pyruvate metabolism proteins enriched in the shared ASD-risk gene PPI network map (g: Profiler, Benjamini-Hochberg FDR adj. $p < 0.05$). (B) Loss of *Taok2* alters cellular respiration in DIV14 mouse cortical neurons (Two-Way ANOVA, $F(1,336) = 18.22$, $p < 0.0001$ between genotypes, post hoc Holm-Sidak test; WT = 118 wells and Het = 122 wells from three separate cultures). Arrows indicate the addition of oligomycin (blue), FCCP (green), and Antimycin A and Rotenone (red). (C) Loss of *Taok2* increases maximal respiration (*left*) and proton leakage (*right*) in DIV14 mouse cortical neurons (ROUT outlier test, $Q = 0.1\%$; Two-tailed unpaired t-test, maximal respiration: $t = 3.015$, $df = 33$, $p = 0.0029$, proton leak: $t = 2.374$, $df = 225$, $p = 0.0184$; WT = 118 wells, Het = 122 wells from three separate cultures). (D) Representative images of *Taok2* WT, Het, and KO neurons stained with MitoTrack (green) and TMRM (red). Arrows indicate Mitotracker-labeled mitochondria with no

TMRM staining. Scale bar is 5 μm . (E) Taok2 Het and KO neurons have decreased number of active (TMRM) Mitotracker-labeled mitochondria (One-Way ANOVA, $F(2, 42) = 7.47$, $p = 0.0017$, post hoc Holm-Sidak test; WT = 14, Het = 15, KO = 15 neurons from 1-3 separate pups per genotype from one culture). (F) Representative images of Taok2 WT, Het, and KO neurons stained with TOMM20 (red) and Phalloidin (green). Scale bar 5 μm . (G) Taok2 Het and KO neurons have increased TOMM20 staining (One-Way ANOVA, $F(2,132) = 4.633$, $p = 0.0111$; WT = 45, Het = 42, KO = 48 neurons from 1-4 separate pups per genotype from two cultures). (H) Representative images of synaptic mitochondria morphological categories. (I), Taok2 KO neurons have increased proportion of category 2 mitochondria with enlarged cristae. (Two-way ANOVA, $F(1, 12) < 0.0001$, $p = 0.9998$ between genotypes, post hoc Holm-Sidak test. 19-25 images per animal and three animals per genotype). (J) CRISPR/Cas9 KO of Taok2, Syngap1, Etfb, Rheb, and Spast differentially alters cellular respiration in DIV14 mouse cortical neurons (Two-Way ANOVA, $F(5,280) = 3.492$, $p = 0.0044$, post hoc Holm-Sidak test; Control = 45 wells, Taok2 KO = 51 wells, Syngap1 KO = 50 wells, Etfb KO = 47 wells, Rheb KO = 48 wells, Spast KO = 45 wells from five separate cultures). Arrows indicate the addition of oligomycin (blue), FCCP (green), and Antimycin A and Rotenone (red). (K) Significant increases in maximal respiration (*top*) and proton leakage (*bottom*) show significant increase in Syngap1 and Etfb KO neurons (ROUT outlier test, $Q = 0.1\%$; maximal respiration: One-Way ANOVA, $F(5, 280) = 2.927$, $p = 0.0136$, post hoc Holm-Sidak test, proton leak: One-Way ANOVA, $F(5, 277) = 4.138$, $p = 0.0012$, post hoc Holm-Sidak test). Mean \pm s.e.m. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

3.7 SUPPLEMENTAL DATA

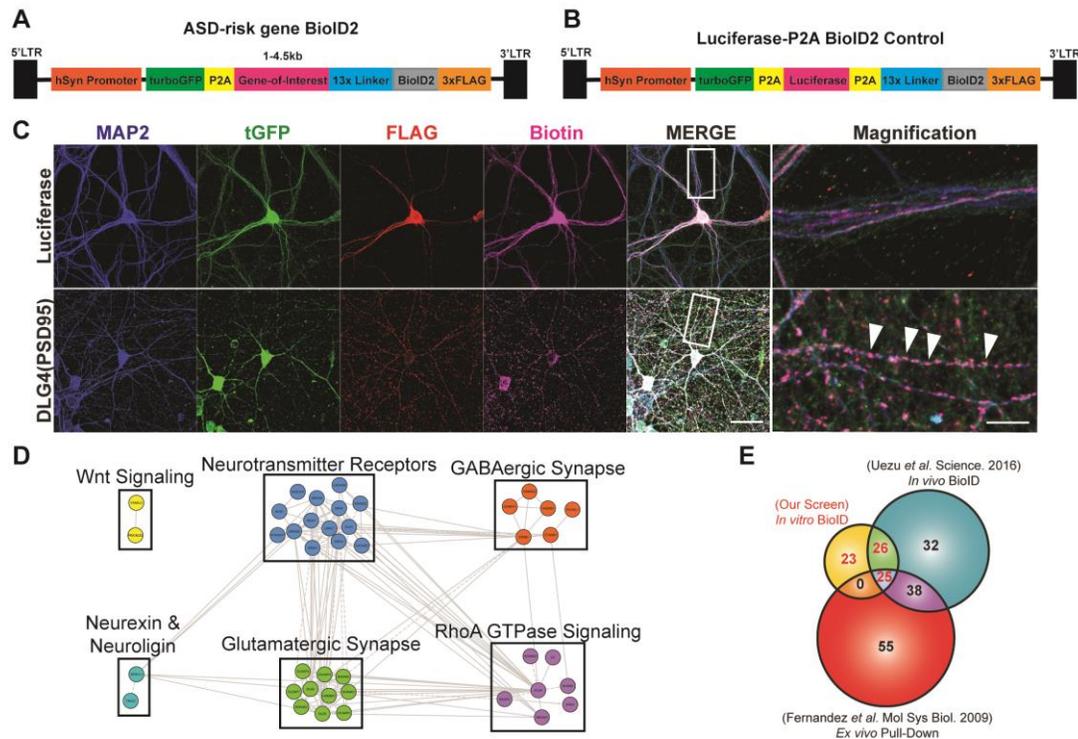


Figure S1. BioID2 of DLG4 in mouse cortical neurons. (A) Diagram of the BioID2 fusion construct for the 41 ASD-risk genes. (B) Diagram of the control Luciferase control fusion construct. (C) Representative images of cortical neurons infected with the PSD95-BioID2 and Luciferase-P2A-BioID2 constructs. Scale bar is 20 μ m. Magnified images are shown on the right. White arrows point to synaptic localization of PSD95-BioID2. Scale bar is 5 μ m. (D) Reactome pathways enriched in the PSD95 PPI network. Clusters created using the Reactome FI app on Cytoscape and labeled with most significantly enriched pathways for each cluster (adj. $p < 0.05$). (E) Venn diagram of shared protein interactors between our in vitro PSD95 PPI network and proteins identified by mouse in vivo PSD95 BioID (Uezu et al. 2016) or mouse in vivo tandem affinity purification of PSD95 (Fernandez et al. 2009).

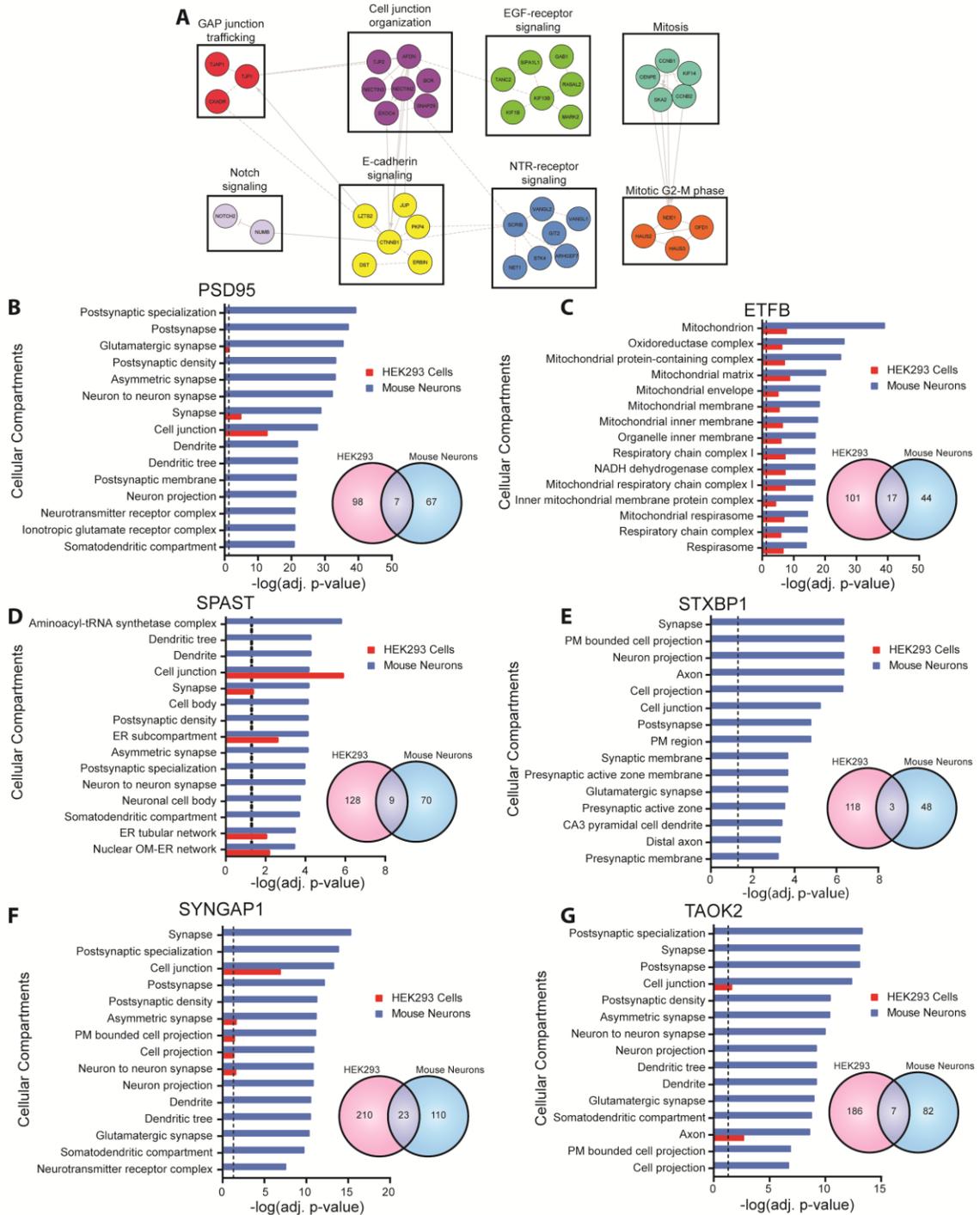


Figure S2. Comparison of ASD-risk gene PPI networks from BioID2 in HEK293 cells and mouse cortical neurons. (A) Reactome pathways enriched in the PSD95 PPI

network from HEK293 cells. Clusters created using the Reactome FI app on Cytoscape and labeled with the most significantly enriched pathways (adj. $p < 0.05$). Top 15 cellular compartments from mouse cortical neurons (blue) enriched in the PPI networks of PSD95 (B), ETFB (C), SPAST (D), STXBP1 (E), SYNGAP1 (F), and TAOK2 (G) compared to enrichment in HEK293 cells (red) (g: Profiler, Benjamini-Hochberg FDR adj. $p < 0.05$). Adjacent Venn diagrams show shared protein interactors identified by BioID2 in HEK293 cells vs mouse cortical neurons. PM: Plasma membrane, OR: Outer membrane, ER: endoplasmic reticulum.

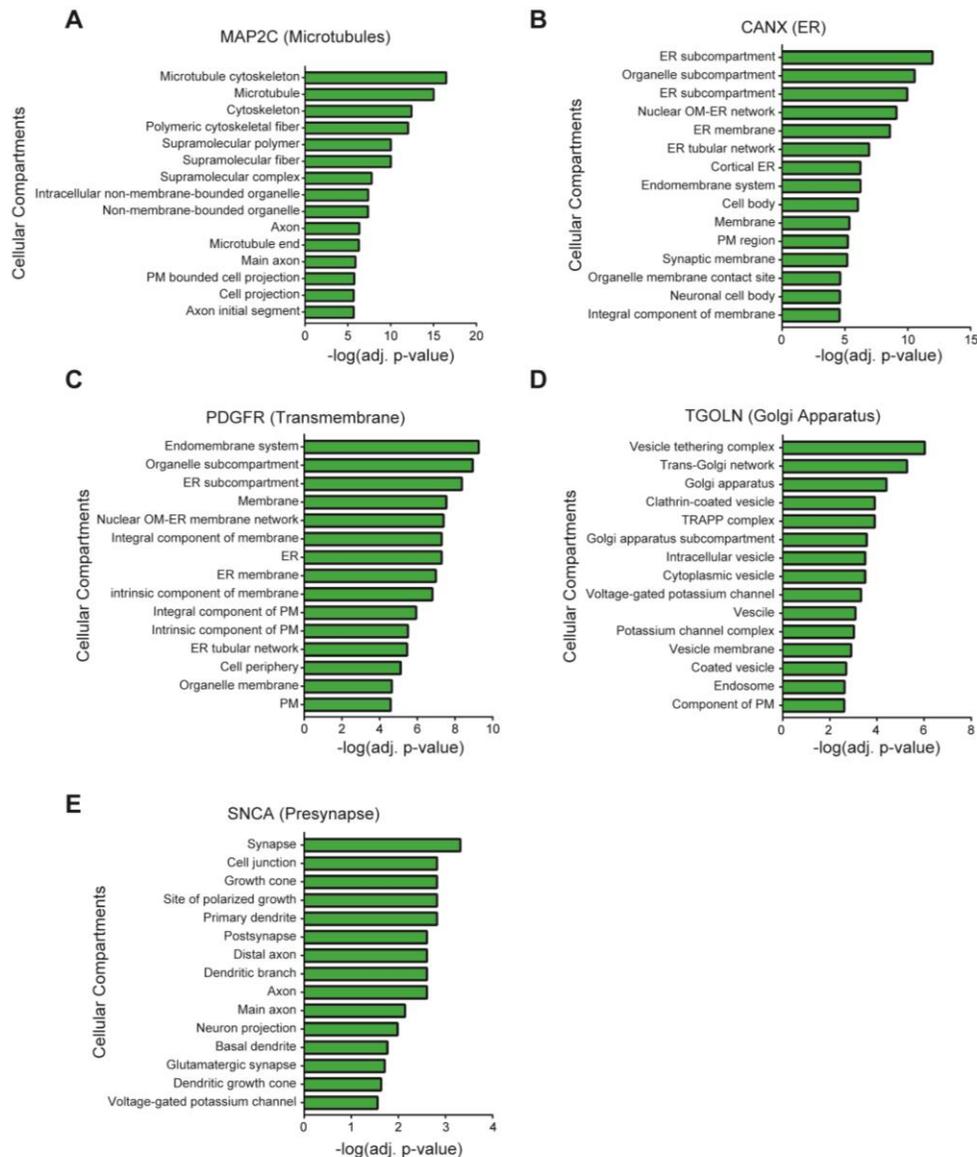


Figure S3. Validation of BioID2 using genes localized to specific compartments in mouse cortical neurons. BioID2 of cellular compartment proteins MAP2C (A), CANX (B), PDGFR (transmembrane domain) (C), TGOLN (D), and SNCA (E). g: Profiler pathway enrichment was used to identify significantly enriched cellular compartments (g: Profiler, Benjamini-Hochberg FDR adj. $p < 0.05$). PM: plasma membrane, OM: outer membrane, ER: endoplasmic reticulum.

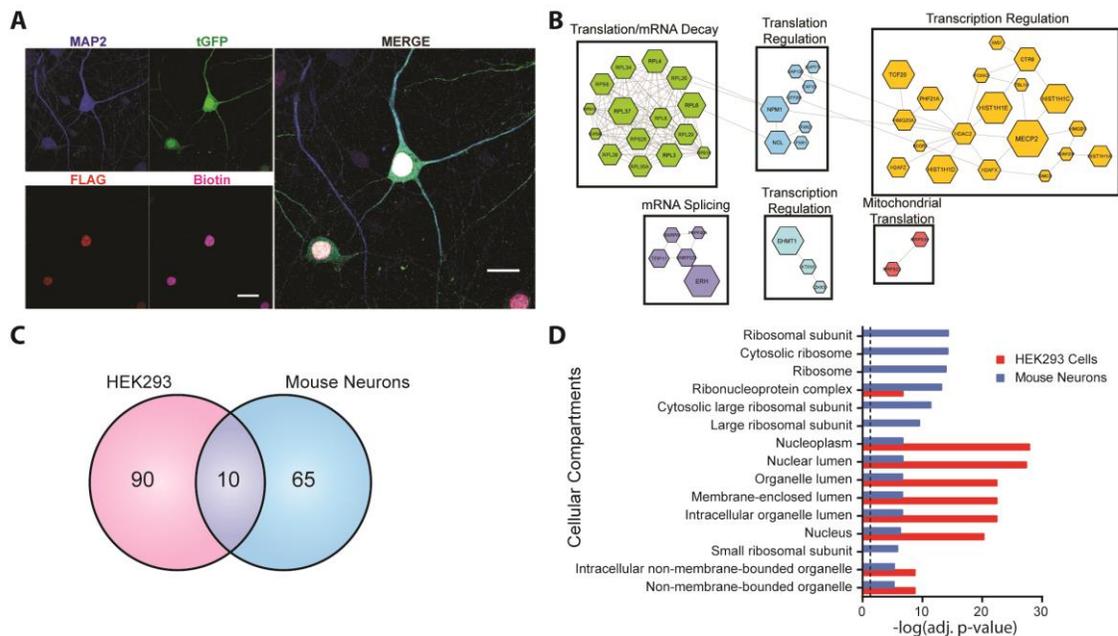


Figure S4. BioID2 of MECP2 in mouse cortical neurons. (A) Representative images of cortical neurons infected with the MECP2-BioID2 construct. Scale bar is 20 μ m. (B) Reactome pathways enriched in the MECP2 PPI network. Clusters created using the Reactome FI app on Cytoscape and labeled with most significantly enriched pathways (adj. $p < 0.05$). (C) Venn diagram shows shared protein interactors identified by BioID2 in HEK293 cells vs mouse cortical neurons. (D) Top 15 cellular compartments from mouse cortical neurons (blue) enriched in the MECP2 PPI network compared to enrichment in HEK293 cells (red) (g: Profiler, Benjamini-Hochberg FDR adj. $p < 0.05$).

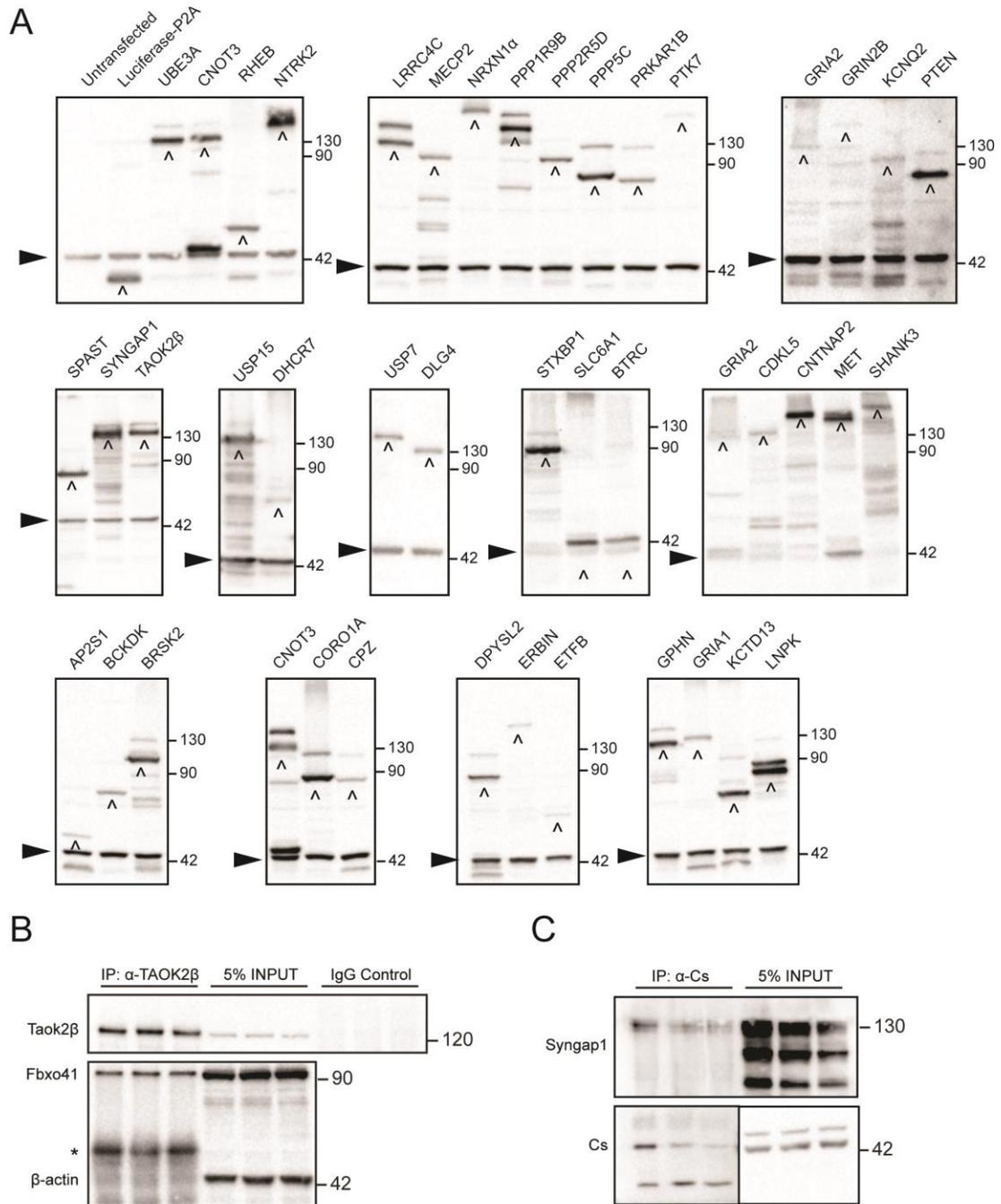


Figure S5. Western blots of 41 ASD-risk gene BioID2 constructs. (A) Western blots of ASD-risk gene BioID2 constructs transfected in HEK293 cells and immunoblotted for FLAG and β -actin as the loading control. ^ denotes expected BioID2 fusion protein size.

Bands higher than the caret indicated bands are possible tGFP fusion proteins due to P2A inefficiency. Arrow denotes β -actin loading control. Bands lower than β -actin are possible degraded BioID2-FLAG proteins. (B) Coimmunoprecipitation of Fbxo41 with TAOK2 β from three separate CD1 mouse cortices. (C) Coimmunoprecipitation of Syngap1 with citrate synthase from three separate CD1 mouse cortices.

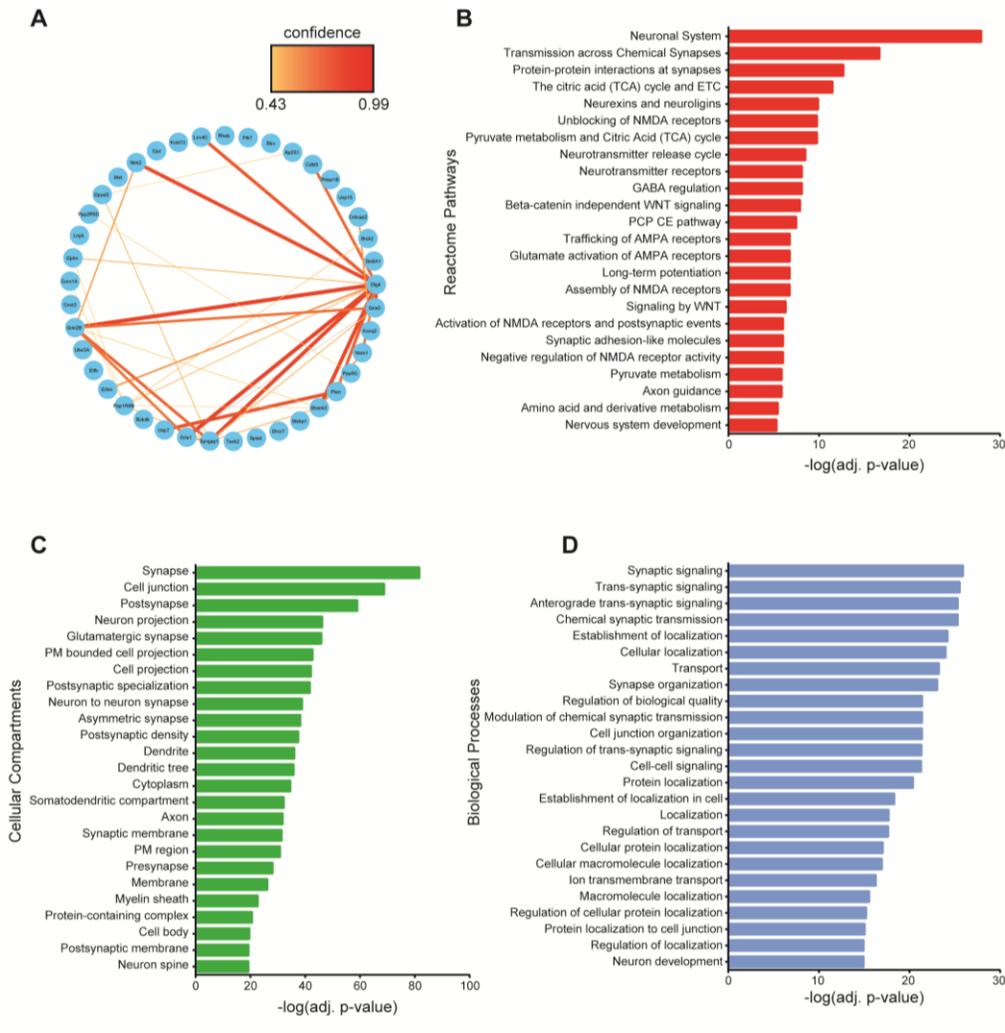


Figure S6. Enriched pathways in the shared ASD-risk gene PPI network map identified using BioID2. (A) Known physical interactions of the 41 ASD-risk genes from the STRING database. Color and increasing thickness of the line represents the confidence of the interaction starting at medium interaction confidence (0.4). Pathway enrichment was used to identify significantly enriched Reactome pathways (B), cellular compartments (C), and biological processes (D) (g: Profiler, Benjamini-Hochberg FDR adj. $p < 0.05$). The top 25 pathways are shown for each graph. PM: plasma membrane, OM: outer membrane, ER: endoplasmic reticulum.

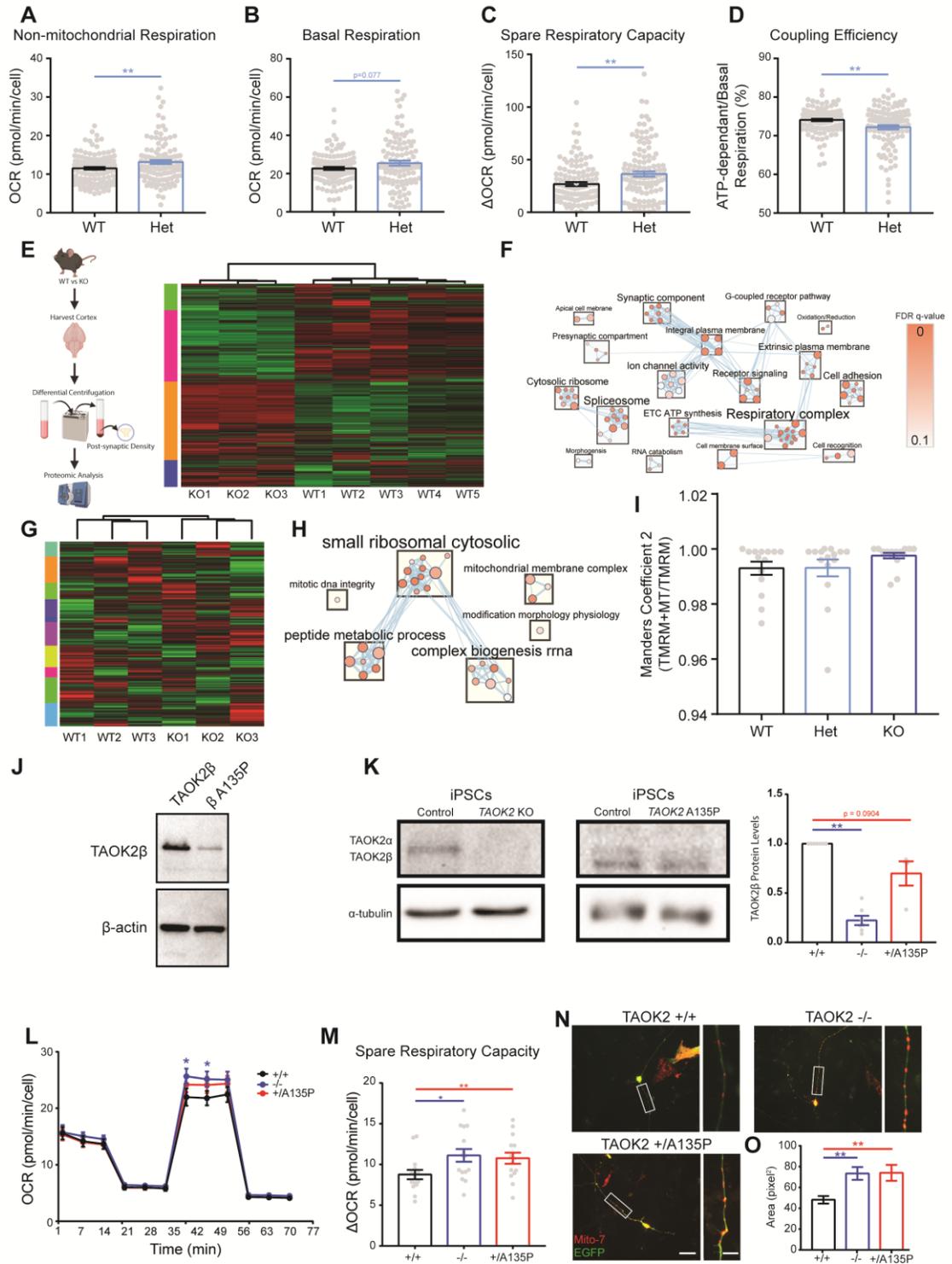


Figure S7. Loss or disruption of ASD-risk gene TAOK2 causes alterations in cellular respiration and mitochondrial proteins. (A-D) Taok2 Het neurons have increased non-mitochondrial respiration and spare respiratory capacity, and decreased coupling efficiency (ROUT outlier test, $Q = 0.1\%$; Two-tailed unpaired t-test, non-mitochondrial respiration: $t = 2.668$, $df = 228$, $p = 0.0082$, basal respiration: $t = 1.776$, $df = 227$, $p = 0.0770$, spare respiratory capacity: $t = 3.007$, $df = 234$, $p = 0.0029$, coupling efficiency: $t = 3.146$, $df = 220$, $p = 0.0019$; WT = 118 wells, Het = 122 wells from three separate cultures). (E) Shotgun proteomics of post-synaptic density fraction from Taok2 WT and KO mouse cortices. (F) Taok2 KO PSD fractions have significant decrease in synaptic and mitochondrial protein gene sets (GSEA, $FDR < 0.1$; five Taok2 WT and three Taok2 KO mice littermates). Size of nodes represents number of proteins and color represents FDR q-value. (G) RNA sequencing of Taok2 WT and KO mouse cortices. (H), TAOK2 KO mouse cortices have altered mRNA levels of mitochondrial membrane proteins. (GSEA, $FDR < 0.1$, three Taok2 WT and KO mice littermates each). Size of nodes represents number of proteins and color represents FDR q-value. (I), All active mitochondria stained by MitoTracker (One-Way ANOVA, $F(2, 42) = 1.355$, $p = 0.2689$, post hoc Holm-Sidak test; WT = 14, Het = 15, KO = 15 neurons from three separate cultures). (J) Western blot of TAOK2 WT and A135P BioID2 constructs. (K) Western blot of CRISPR/Cas9-edited iPSCs and neurons showing loss of TAOK2 expression in TAOK2 KO (-/-) and A135P (+/A135P) lines (One-sample t-test, WT vs KO: $t = 16$, $df = 6$, $p < 0.0001$, WT vs A135P: $t = 2.465$, $df = 3$, $p < 0.0904$; WT = 7, KO = 7, A135P = 4 wells from separate iPSC cultures). (L-M) DIV7 TAOK2 KO human neurons have

significantly increased maximal respiration (Two-Way ANOVA, $F(2,42) = 0.659$, $p = 0.5226$ between genotypes, post hoc Holm-Sidak test; WT = 118 wells and Het = 122 wells from three separate cultures). TAOK2 KO and A135P neurons have significantly increased spare respiratory capacity (One-Way ANOVA, $F(2,42) = 3.409$, $p = 0.0424$, post hoc Holm-Sidak test; WT = 118 wells, Het = 122 wells from three separate cultures).

(N) Representative images of human neurons transfected with Mito7-DsRed constructs at DIV7 and fixed and imaged at DIV9. Scale bar is 20 μm . Magnification of boxed areas shown on the right. Scale bar is 5 μm . (O) TAOK2 KO and A135P neurons have larger Mito7-DsRed punctae size compared to wildtype neurons (One-Way ANOVA, $F(2, 1309) = 5.032$, $p = 0.0067$, post hoc Holm-Sidak test; WT = 71, KO = 520, A135P = 421 punctae from 15-16 neurons per genotype). Mean \pm s.e.m. $p < 0.05$, $**p < 0.01$.

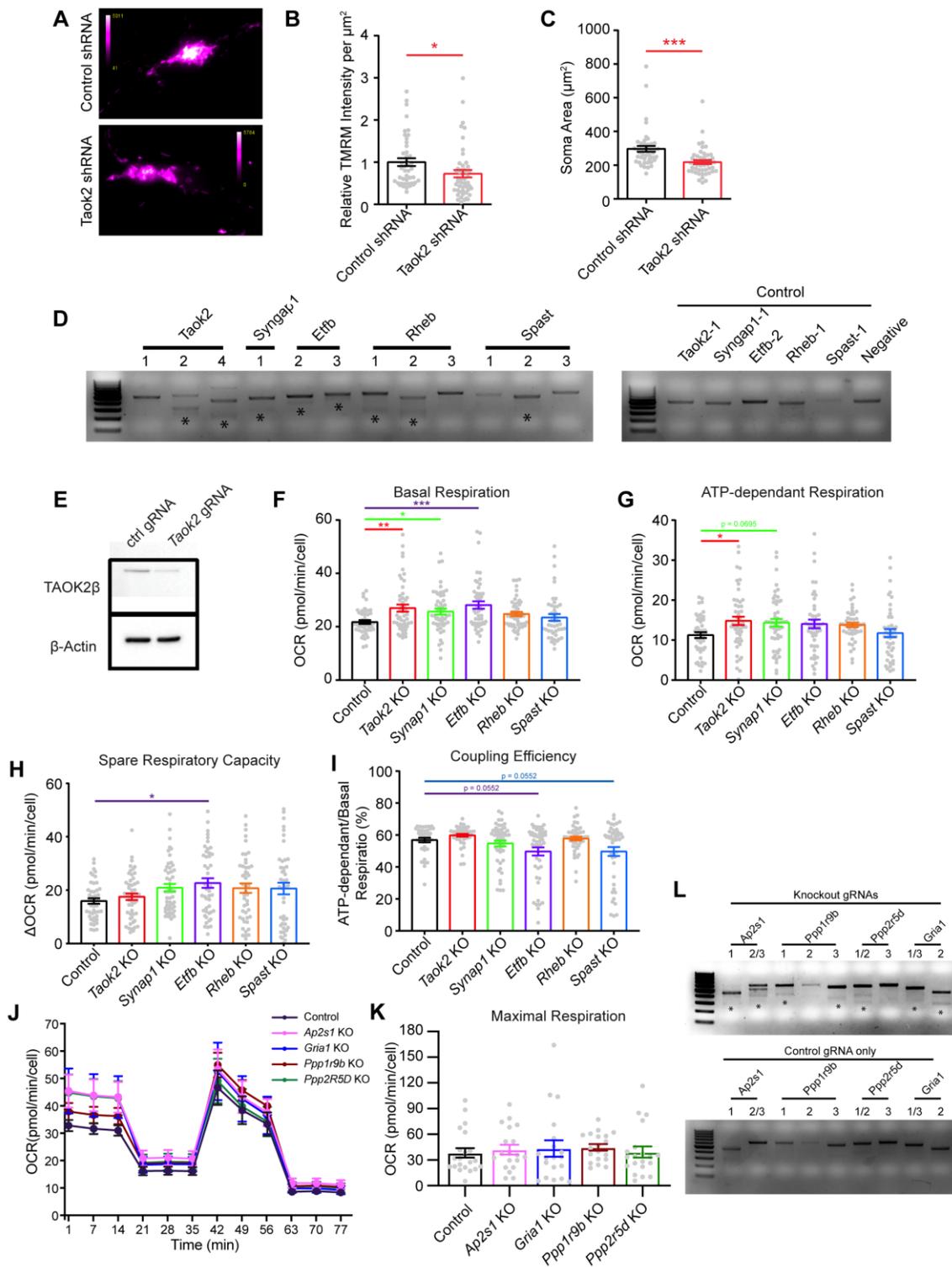


Figure S8. Acute knockout of ASD-risk genes in mouse cortical neurons alters

mitochondrial activity and cellular respiration. (A) Representative images of mouse cortical neurons infected with control shRNA or Taok2 shRNA showing TMRM intensity. Scale bar is 10 μ m. (B) Mouse cortical neurons with acute knockout of Taok2 have decreased relative TMRM activity, even with decreased soma size (C) (TMRM activity: Unpaired t-test, $t = 2.156$, $df = 91$, $p = 0.0337$; WT = 44, Taok2 KD = 49 neurons from two separate cultures). (D) Indel cleavage assay for Taok2, Syngap1, Etfb, Rheb, and Spast KO shows at least at least one disrupted gRNA target region in mouse cortical neurons. *Indicates secondary band due to digested indel. (E) Reduced Taok2 protein expression in mouse cortical neurons infected with Cas9 and Taok2 KO gRNAs. Infected at DIV7 and harvested at DIV18 for western blot of TAOK2 β . Significant changes in different aspects of cellular respiration (Basal respiration (F), ATP- dependent respiration (G), spare respiratory capacity (H), and coupling efficiency (I) in mouse cortical neurons with CRISPR/Cas9 KO of Taok2, Syngap1, or Etfb (ROUT Outlier Test, $Q = 0.1\%$; basal respiration: One-Way ANOVA, $F(5,279) = 3.994$, $p = 0.0016$, ATP- dependent respiration: One-Way ANOVA, $F(5,279) = 2.517$, $p = 0.0300$, spare respiratory capacity: One-Way ANOVA, $F(5,276) = 2.381$, $p = 0.0389$, coupling efficiency: One-Way ANOVA, $F(5,263) = 4.894$, $p = 0.0003$, post hoc Holm-Sidak test; Taok2 KO = 51 wells, Syngap1 KO = 50 wells, Etfb KO = 47 wells, Rheb KO = 48 wells, Spast KO = 45 wells from five separate cultures). Mean \pm s.e.m. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. (J) CRISPR/Cas9 KO of *Ap2s1*, *Gria1*, *Ppp1r9b*, and *Ppp2r5d* show no changes in cellular respiration or maximal respiration (K) in DIV14 mouse

cortical neurons (Two-Way ANOVA, $F(4,93) = 0.4318$, $p = 0.7853$, post hoc Holm-Sidak test; maximal respiration: One-Way ANOVA, $F(4, 93) = 0.1827$, $p = 0.9469$; Control = 21 wells, Ap2s1 KO = 19 wells, Gria1 KO = 19 wells, Ppp1r9b KO = 19 wells, Ppp2r5d KO = 20 wells from five separate cultures). (L) Indel cleavage assay for Ap2s1, Gria1, Ppp1r9b, Ppp2r5d KO shows at least at least one disrupted gRNA target region in mouse cortical neurons. *Indicates secondary band due to digested indel.

Tables are available at <https://doi.org/10.1101/2022.01.17.476220>

Table S1. BioID2 PPI networks of 41 ASD-risk genes and cellular compartment genes

Table S2. Comparison of BioID2 PPI networks identified in HEK293 cells and mouse cortical neurons

Table S3. Comparison of BioID2 PPI network enriched cellular components identified in HEK293 cells and mouse cortical neurons

Table S4. BioID2 PPI network enriched cellular components of compartment specific genes

Table S5. BioID2 PPI network enriched pathways of 41 ASD-risk genes

Table S6. 41 ASD-risk gene PPI network map enriched pathways

Table S7. Differentially expressed genes and proteins and dysregulated pathways in *Taok2* KO mouse cortices

**CHAPTER 4: PROXIMITY-LABELING OF AUTISM RISK GENES
IDENTIFIES DISEASE RELEVANT IMPACT ON PROTEIN FUNCTION AND
CLINICAL PATHOLOGIES**

Nadeem Murtaza^{1,2}, Annie A. Cheng², Chad O. Brown^{1,2}, Jarryll A. Uy^{2,3}, Joelle El-Hajjar¹, Neta Pipko¹, Brianna K. Unda², Sansi Xing¹, Bhooma Thiruvahindrapuram⁴, Worrawat Engchuan⁴, Brett Trost⁴, Eric Deneault⁵, Bradley W. Doble⁶, James Ellis^{4,7}, Evdokia Anagnostou^{8,9}, Gary D. Bader⁷, Stephen W. Scherer^{4,7}, Yu Lu¹, Karun K. Singh^{2,3*}

¹Department of Biochemistry and Biomedical Sciences, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada

²Krembil Research Institute, University Health Network, Toronto, ON, Canada

³Department of Laboratory Medicine and Pathobiology, Faculty of Medicine, University of Toronto, Toronto, ON, Canada

⁴The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, ON, Canada

⁵Centre for Biologics Evaluation, Biologic and Radiopharmaceutical Drugs Directorate, Health Products and Food Branch, Health Canada, Ottawa, Ontario, Canada

⁶Department of Paediatrics, Department of Biochemistry and Medical Genetics, University of Manitoba, Winnipeg, MB, Canada

⁷Department of Molecular Genetics, Faculty of Medicine, University of Toronto,
Toronto, ON, Canada

⁸Department of Neurology, The Hospital for Sick Children, Toronto, ON, Canada.

⁹Bloorview Research Institute, Holland Bloorview Kids Rehabilitation Hospital, Toronto,
ON, Canada.

***Correspondence:** karun.singh@uhnresearch.ca

ACKNOWLEDGMENTS

We thank C.I. and P.dG. for proof reading the manuscript. Work in the Singh Lab was supported by the Canadian Institute of Health Research (CIHR), the Ontario Brain Institute (OBI), the Network for European Funding for Neuroscience Research (NEURON ERA-NET), and the Donald K. Johnson Eye Institute at University Health Network.

PREFACE

The combined manuscript from this chapter and Chapter 3 is currently in the first round of revisions at Cell Reports as of July 2022. It has been modified and kept separate to match the formatting of this thesis. References for Chapters 3 and 4 are combined with the thesis references. Chapter 3 uses the same methods as Chapter 4, and only methods specific to this chapter were kept for space.

The aim of this chapter, was to use BioID as a proof-of-principle for investigating the impact of patient *de novo* mutations on unbiased ASD-risk gene function at the

protein level. Screening *de novo* mutations are an important way to identify how ASD-risk genes cause the pathologies associated with the disorder, however current screens are specified towards specific functions. Our study in Chapter 2, utilized confocal microscopy, electrophysiology, and protein assays to elucidate the impact of TAOK2 mutations, however BioID was able to pinpoint these deficits in one experiment. Therefore, the use of BioID to screen genetic variants associated with ASD, has the potential to streamline studies by providing data to direct functional assays or prioritize specific mutations.

Dr. Karun Singh and I designed this study and all the experiments. Annie Cheng and Dr. Brianna Unda created the initial BioID2 lentiviral construct backbone. I generated all of subsequent DNA constructs and performed all experiments and data analysis unless otherwise specified. Annie Cheng assisted in the generation of all lentiviruses. Sansi Xing and Yu Lue ran samples through the mass spectrometer and assisted with data acquisition. Chad Brown performed all electrophysiology recordings. Jarryll Uy, Joelle El-Hajjar, and Neta Pipko assisted with human neuron morphology analysis and western blots. Eric Deneault, Dr. James Ellis, and Dr. Steve Scherer assisted in CRISPR/Cas9 editing the human TAOK2 KO and A135P iPSC lines. Evdokia Anagnostou advised on clinical score analysis. Dr. Gary Bader advised on pathway analyses used in the project. Dr. Karun Singh and I wrote the manuscript with input from Annie Cheng, Dr. Brett Trost, Dr. Bhooma Thiruvahindrapuram, Dr. Worrawat Engchuan and Dr. Bradley Doble.

4.1 ABSTRACT

Autism spectrum disorder (ASD) is a genetically heterogeneous disorder. Sequencing studies have identified hundreds of risk genes for autism spectrum disorder (ASD), but the signaling networks of actively expressed genes at the protein level remain largely unexplored. To address this gap, we used neuron-specific proximity-labeling proteomics (BioID) to identify protein-protein interaction (PPI) networks of 41 ASD-risk genes. Neuron-glia co-cultures provided neuron-specific PPI networks that were not largely identified in frequently used HEK293 cells. These neuron-specific networks included synaptic transmission and dendritic arborization proteins, for which we showed that patient *de novo* missense variants disrupt the PPI network of synaptic ASD-risk genes and cause deficits in synaptic transmission. Network analysis revealed that the combined 41 risk gene PPI network map had a 50-fold increase in shared connectivity between distantly related ASD-risk genes than those in existing public databases. Through the PPI network map, we identified convergent pathways between established and uncharacterized risk genes, including synaptic transmission, mitochondrial/metabolic processes, Wnt signaling pathways, ion channel activity and MAPK signaling. Using CRISPR gene knockouts, we further investigated metabolic dysfunction, which has been peripherally associated with ASD, and revealed a functional association between mitochondrial activity and multiple ASD-risk genes not previously connected with this pathway. The network also showed an enrichment of 112 additional ASD-risk genes and brain region-specific differentially expressed genes identified in post-mortem ASD patients. Clustering of the ASD-risk genes based on their PPI network identified gene

groups that correspond to clinical behavior score severity depending on the type of mutation, highlighting a potential link between molecular processes and complex clinical behavior phenotypes. Together, our data reveal that using cell type-specific PPI networks to map ASD risk genes can identify previously unknown individual and convergent neuronal signaling networks, provide a scalable method to assess the impact of patient variants, and reveal biological insight into disease mechanisms and sub-cohorts in ASD.

4.2 INTRODUCTION

1 in 65 children under the age of 8 (Of et al., 2018) are diagnosed with Autism spectrum disorder (ASD). ASD is defined by two core symptoms, including deficits in social communication and repetitive behaviors and restricted interests. However, due to the heterogeneity of ASD, individuals with the disorder have many co-occurring symptoms ranging from anxiety to severe epilepsy (Doyle and McDougle, 2012). Currently, treatment of the disorder is based on attenuating these secondary symptoms or teach these individuals behavioral coping mechanisms (Ghanizadeh et al., 2014; Tonge et al., 2014). Monozygotic twin studies have shown that there is a large genetic component towards the risk of ASD (Gaugler et al., 2014). Multiple monogenic syndromic forms of ASD have been studied; the majority of ASD risk is estimated to be from common genetic, rare-inherited, or *de novo* variants (Neale et al., 2012; Sanders et al., 2012; Gaugler et al., 2014; Koiré et al., 2021). As such, numerous large scale WES and WGS studies have identified hundreds of genes associated with ASD risk (Glessner et al., 2009; Iossifov et al., 2012; Neale et al., 2012; O’Roak et al., 2012b, 2012a; Sanders et al., 2012,

2015; Yuen et al., 2017; Feliciano et al., 2019; Ruzzo et al., 2019; Satterstrom et al., 2020; Wilfert et al., 2021). The large numbers of genes that have been identified in the last decade have not been matched with research into understanding their role in the pathology of ASD. Studying the hundreds of genes and their associated ASD variants one at a time, is both resource and time intensive.

Studies have used lower complex organisms to multiple mutations through function assays or protein-protein interactions (PPI) (Sakai et al., 2011; Corominas et al., 2014; Wong et al., 2019; Post et al., 2020; Marcogliese et al., 2022). However, a major caveat is the correlation of behavior phenotypes in animal models to complex human behaviors and the relevance of interactions identified in non-neuronal cells. To address this issue, we developed a lentiviral *in vitro* proximity-labeling proteomics (BioID2) system that uses mouse primary neurons. Although proximity labeling has not been used to study the impact of genetic variants on PPI networks and their correlated protein function, affinity proteomics has been used to identify novel pathways altered by mutations in the Tau protein, which is associated with Alzheimer disease (Tracy et al., 2021). Our system uses cortical neurons, because of their importance in ASD pathology (Nowakowski et al., 2017; Velmeshev et al., 2019), to identify changes in the PPI networks of three ASD-risk genes, TAOK2, GRIA1, and PPP2R5D as a proof-of-principle in using proximity labeling to study ASD associated genetic variants in a high-throughput and unbiased manner.

The majority of *de novo* missense mutations have been found to disrupt protein interaction in genes that play a central role in biological processes, many of which are

convergent molecular pathways (Chen et al., 2015, 2018). Our large BioID screen identified multiple convergent pathways through the shared 41 ASD-risk gene PPI network (Murtaza et al., 2022), and provide a backdrop to study the connection between clinical behavioral phenotypes and genetic variants.

In the current study, we examined rare and *de novo* missense variants in synaptic or poorly characterized ASD-risk genes. We found disruption of key PPIs that led to functional deficits in synaptic transmission. Our PPI network in mouse cortical neurons was cross-referenced with human data to demonstrate its relevance to ASD pathology. Comparing the shared 41 ASD-risk gene PPI network map from our large ASD-risk gene study combined with human sequencing data revealed an enrichment of an additional 112 ASD risk genes and expression in human brain cell types associated with ASD pathology (Sanders et al., 2015; Yuen et al., 2017; Feliciano et al., 2019; Ruzzo et al., 2019; Satterstrom et al., 2020). More importantly, comparing the PPI network to human clinical data from the MSSNG database (genome sequencing and clinical data from over 5,000 individuals with ASD) (Yuen et al., 2017; Trost et al., 2020), we found that individuals with variants in risk genes with a high degree of shared interactions have similar adaptive behavior scores.

Taken together, we demonstrate that neuron-specific PPI networks provide a powerful approach to reveal novel individual and convergent disease mechanisms in ASD. Given the scalability of our method and its underutilization in ASD research, we believe our PPI network resource and screening system can be applied more broadly to

additional autism risk genes to identify previously unknown or overlooked disease mechanisms that are not captured with current approaches.

4.3 MATERIALS AND METHODS

Human iPSC to neuron differentiation via NGN2 induction

Human iPSCs were cultured on Matrigel (Corning) coated plates using mTeSR1 media (Stem Cell Technologies) and passaged every 3-4 days using ReLeSR (Stem Cell Technologies) until neural induction. A modified NGN2 induction protocol (Zhang *et al.*, 2013) was used to differentiate human iPSCs into excitatory NGN2 neurons (Zhang *et al.*, 2013). Human iPSCs were dual infected with pTet-O-NGN2-P2A-EGFP and FUW-M2rtTA lentiviruses for dox-inducible expression and were titered for > 90% infection efficiency. On Day -1 iPSCs were singularized using Accutase (Stem Cell Technologies) and plated with mTeSR1 media (supplemented with 10 μ M Y-27632) on Matrigel at 400,000 cells per well in a 6-well plate. On Day 0, media exchanged and supplemented with Doxycycline (1 μ g/ml). On Day 1 and 2, media was replaced with iNPC media (DMEM/F12 media (Gibco) supplemented with N2 (Gibco), MEM NEAA (Thermo), 2mM GlutaMAX, and Pen-Strep) with Doxycycline and Puromycin (2 μ g/ml). On Day 3, media was then replaced with iNi media (Neurobasal media with SM1 (Stem Cell Technologies), 2mM GlutaMAX, Pen-Strep, 20 ng/ml BDNF, 20 ng/ml GDNF, and 1 μ g/ml Laminin) with Doxycycline. On day 4, differentiated neurons were singularized using Accutase and re-plated at 100,000 cells per well in a 24-well plate in only iNi media. Plates were pre-coated with 20 μ g/ml Laminin and 67 μ g/ml Poly-ornithine

(Sigma). Mouse glial cells were plated on top of the differentiated neurons after 24 hours at a density of 50,000 cells per well. Half-media changes were carried out every other day, and iNi media was supplemented with 2.5 % FBS on Day 9 and onwards. Neurons were grown until day 28 post NGN2-induction.

Analysis for the identification of ASD-risk and cellular compartment protein PPI networks

Only proteins identified with two unique peptides were used for analysis. Flow cytometry was used to calculate the total GFP in infected neuron samples. If the POI sample had less GFP than the luciferase control sample, the factor needed to equalize the amount of GFP was applied to the protein abundances of the POI samples. Protein abundances were also normalized to the highest total protein count sample for each set of biological replicates. Unpaired one-tailed student's test was used to determine significantly enriched biotinylated proteins in the POI sample using the Log₂ abundances of the three biological replicates of the POI samples compared to the luciferase control samples ($p < 0.05$) (Uezu et al., 2016). Significance B outlier test was used to identify significantly biotinylated proteins in the POI sample compared to the luciferase control sample using the average abundance and protein abundance ratio between POI and luciferase samples (SigB $p < 0.05$). Only proteins that were found to be significant from both analyses were included in the PPI network. The protein abundance ratio between the luciferase control replicate samples, which were labeled with different TMT labels, was considered to be the minimal ratio required for significance. Any protein that did not

surpass this ratio was considered to be a false positive, even if statistically significant, and not included in the PPI network.

Synapsin1 puncta analysis in human iPSC-derived neurons

Synapsin1 stained images were processed and analyzed with ImageJ software. The Synapsin1 antibody was co-immunostained with MAP2 to determine dendrites with presynaptic puncta. Five biological replicates, which represent five separate neural inductions, were used for synaptic analysis. 5-10 neurons per genotype per replicate were used. Imaging settings were kept the same between images and synapsin1 images were analyzed at the same threshold. Dendrites were traced using ImageJ and the measure tool was used to quantify the number and size of the puncta within the traced region.

Cell type/DEG/ASD gene list enrichment analysis

Human cell type gene expression and ASD DEGs and ASD gene lists were obtained from their respective publications (Sanders et al., 2015; Nowakowski et al., 2017; Yuen et al., 2017; Feliciano et al., 2019; Ruzzo et al., 2019; Velmeshev et al., 2019; Satterstrom et al., 2020; Wilfert et al., 2021). For the enrichment analysis we used the Fisher exact test comparing each gene list with the shared ASD-risk gene PPI network in the mouse brain background protein list, which was used for pathway enrichment analysis. P-values and ODDs ratios were calculated for each comparison. To account for multiple comparisons, Bonferroni correction thresholds were calculated as $p = 0.05$ divided by the number of comparisons.

Clinical score analysis

Rare variants of individuals diagnosed with ASD were extracted from the MSSNG database (research.mss.ng)(Yuen et al., 2017), which has whole genome sequences of 4,258 families and 5,102 ASD-affected individuals. Only variants with estimated high or medium impact strengths were used for analysis, and variants were categorized into three groups (missense variants, splicing variants, and frame shift/premature stop codon variants). Adaptive behavior and socialization standard scores of affected individuals was extracted from the MSSNG associated Metabase (data-explorer.mss.ng). Individuals were grouped based on the presence of mutations in the 41 ASD-risk genes that were clustered into three groups. Individuals that had variants in genes between multiple groups were not included in the analysis. Separate analyses were carried out between individuals grouped by missense, splicing or frame shift/premature stop codon variants. Clinical data was considered as non-parametric and the Kruskal-Wallis ranked test with post hoc Dunn's test was used for comparison between the adaptive behavior and socialization standard scores of each group.

4.4 RESULTS

4.4.1 PPI networks identify differences in signaling between missense variants in ASD risk genes

We hypothesized that PPI networks could be used to study missense variants, which are a large and important class of genetic risk factors for ASD that have less obvious functional impacts compared to loss-of-function (LoF) variants. Sequencing of

ASD individuals have identified many missense variants of unknown significance (VUS) and therefore, the biological impact of variants in the majority of risk genes remain unknown. Understanding the impact of a variant is important because it provides the affected individual and family with a possible causal explanation and, in some cases, it could help to assess clinical trajectory or treatments. Missense variants have been suggested to impact protein stability and protein-protein interaction networks (Chen et al., 2020); however, these data were imputed from databases using primarily non-neuronal datasets, and were not tested in neurons. We used BioID2 to identify differences in severity and pathogenic mechanisms of *de novo* missense variants identified in individuals diagnosed with ASD. Due to the strong link between synaptic functional deficits and ASD pathophysiology, we chose two known synaptic genes (TAOK2 β which is the synapse-specific isoform of TAOK2 and GRIA1) and a less well-characterized risk gene with no specific cellular localization (PPP2R5D) (Figure 1A-1C, Table S8, and Table S9).

We used BioID2 to determine the change in the TAOK2 β PPI network due to the A135P *de novo* missense variant, which was identified in an individual with ASD. The TAOK2 β A135P PPI network had a reduced number of proteins associated with the synaptic compartment, and simultaneously had increased dendritic and ribosomal proteins (Figure 1D). The latter changes may be due to the loss of PPI network proteins in dendritic spines where TAOK2 β localizes, and an increase in dendritic and ribosome translation complex protein interactions specific to the TAOK2 β A135P (Figure 1E), combined with the decreased expression of the A135P mutant. To corroborate the

possible synaptic deficits caused by the A135P variant, we performed patch-clamp electrophysiology on the isogenic iPSC-derived NGN2-neurons (Figure S1A-S1C)(Zhang et al., 2013; Deneault et al., 2018, 2019). *TAOK2* KO and *TAOK2* A135P neurons had decreases in frequency and amplitude of spontaneous excitatory post-synaptic currents (sEPSCs) (Figure 1F and 1G), which coincides with the reduced biotinylation of synaptic proteins in neurons expressing the *TAOK2* A135P-BioID2 protein. The lack of change in the intrinsic firing properties or Synapsin1-positive punctae density in *TAOK2* A135P neurons, as opposed to the *TAOK2* KO neurons (Figure S1C and S1D), suggest that the shift in interaction and localization for the heterozygous A135P line has dissimilar phenotypes compared to the complete loss of *TAOK2*. In fact, *TAOK2* A135P neurons displayed increased size of Synapsin1 punctae, suggesting possible changes in the synaptic structure (Figure S1C and S1D). Taken together, the *TAOK2* β A135P variant showed significant decreases in synaptic pathway protein interactions, demonstrating that changes in PPI networks can be predictive of functional deficits.

We also asked whether PPI networks can distinguish the impact of missense variants based on their location within functional domains of a gene. We investigated *GRIA1* and two *de novo* missense variants, R208H and A636T(de Ligt et al., 2012; Iossifov et al., 2014; Geisheker et al., 2017), located in the extra-cellular ligand binding domain and the transmembrane domain, respectively (Fig. 1B). The *GRIA1* variants showed strong differential effects in their enriched cellular compartments (Figure 1H) and the number of shared interacting proteins compared to the wildtype (Figure 1I).

GRIA1 R208H had a significant loss of proteins localizing to the AMPA receptor and post-synaptic density, which suggests functional changes in synapse function. GRIA1 A636T had a less severe impact, with small increases in the number of compartment-specific protein interactions and gains in membrane junction and ER proteins (Figure 1H, Table S8, and Table S9), suggesting possible trafficking issues. There were no changes in expression between the two variants (Figure S1E). No changes in expression for the GluA2 subunit indicated that overall neuron translation of other AMPA receptor subunits were unchanged (Figure S1F). To identify functional changes that coincide with the changes in PPI networks, we infected mouse cortical neurons with the GRIA1 WT and both variants to obtain whole-cell voltage clamp recordings. This revealed a trend towards decreased sEPSC frequency in neurons expressing the R208H variant, but not the A636T variant (Figure 1J and 1K). Although the A636T mutant had no change in sEPSCs, we did observe large sEPSC bursts (Figure 1J), which may be indicative of altered trafficking of AMPA receptors through the ER network and longer turnover rates (Pick and Ziff, 2018; Schwenk et al., 2019). Together, the stronger loss of interactions for R208H compared to A636T coincide with the electrophysiology results, demonstrating the potential use of BioID2 PPI networks to streamline identification of functional differences in missense variants for receptor proteins.

Finally, we used BioID2 to test missense variants in the risk gene PPP2R5D, a regulatory subunit of phosphatase-2A (Shang et al., 2016). This protein is not known to have multiple functional domains or a specific localization; therefore, BioID2 could help to first understand where it functions in neurons and then the impact of ASD missense

variants. We selected three *de novo* PPP2R5D variants, P53S, E198K, and E420K, which are spread throughout the protein(Houge et al., 2015; Shang et al., 2016). The PPI networks for the variants had both common and dissimilar effects (Figure 1C, 1L, and S1G), with all three variants reducing interactions with synaptic and dendritic proteins enriched in the wildtype PPI network (Figure 1L). This suggests that PPP2R5D has a potential role in dendrites and synapses based on PPI network. Additionally, all of the variants caused a loss and gain of diverse interactions (Figure S1G), with no change in expression levels (Figure S1H). Interestingly, both the E198K and the E420K variants gained trans-Golgi compartment proteins (Figure 1L, Table S8 and, Table S9), suggesting altered localization. Previous studies have described an overactive AKT pathway caused by the PPP2R5D E420K variant(Papke et al., 2021). However, measurement of phospho-AKT levels in HEK293 cells expressing the variants revealed no difference (Figure S1I), suggesting that specific molecular assays may miss functional deficits. To probe the E198K and E420K variants further, we infected mouse neurons with Flag-tagged PPP2R5D wildtype and variants, and stained for a trans-Golgi apparatus-specific protein, GALNT2 (Figure 1M). The Golgi-apparatus of neurons expressing the E198K variant was significantly smaller and spread throughout the primary dendrites, which coincides with enrichment of apical dendrite and Golgi associated vesicles in its PPI network (Figure 1L, 1M, and Table S8 and S9). Neurons expressing the E410K variants showed no change in size, however the Golgi-apparatus also spread into the apical dendrites, which may be explained by the increased presence trans-Golgi network and vesicle

tethering complex protein in its PPI network (Figure 1L, 1M, and Table S8 and S9). The overexpression of PPP2R5D variants had no effect on neuron size (Figure S1J).

Together, the BioID2 approach revealed dendritic and synaptic localization of PPP2R5D, which is lost in multiple missense variants that have their own subtle differences. The differences in the PPI network of wildtype proteins and their ASD-associated variants highlight the utility of the system to screen multiple disease variants within a gene.

4.4.2 The 41 ASD-risk gene PPI network map enriches for additional ASD risk genes, human disease cell types, and correlates with human behavioral phenotypes from clinical datasets

To further demonstrate the utility of the 41 ASD-risk gene PPI network map resource, we used enrichment analysis to determine relevance to human ASD. We found a significant enrichment of 112 additional ASD-risk genes (Fisher's Exact test $p = 2.69 \times 10^{-30}$, OR = 3.45), highlighting the strong functional connectivity between ASD-risk genes at the protein level (Figure 2A). Along with enrichment of ASD-risk genes from the original 41 ASD-risk protein baits, we found that gene lists reported from individual sequencing studies were enriched, especially when examining cytoplasmic (non-nuclear) proteins (Figure S10A). This suggests strong connectivity of ASD protein signaling outside the nucleus. Gene lists with only nuclear proteins were not enriched (Figure S10B), providing evidence that there is less interaction between proteins localized to the nucleus and those in the cytoplasm. Of the 153 ASD-risk proteins in the network, 69 are

interacting with 2 or more ASD bait proteins. *Slitrk5*, *Gria2*, *Dlg4*, *Grin2b*, and *Shank2* were identified by more than eight of the ASD bait proteins, suggesting a potential central role for these genes in ASD pathology. Enrichment of multiple cytoplasmic ASD-risk proteins in the PPI network indicates functional connectivity between intracellular signaling proteins.

While the PPI network from 41 ASD-risk genes was generated using human genes, it was obtained in a background of mouse cortical neuron and glia co-cultures; therefore, it is unknown whether this network map is applicable to human brain cell types or differentially expressed genes (DEGs) implicated in ASD pathology. To address this, we examined the enrichment of specific cell types based on their single cell RNA-sequencing profiles (Nowakowski et al., 2017; Velmeshev et al., 2019). We found that the 41 ASD-risk gene PPI network map strongly enriches for excitatory and inhibitory neuron cell types, along with neural progenitor cells, astrocytes and microglia (Figure S10C), which have been associated with ASD pathophysiology (Parikshak et al., 2013; Willsey et al., 2013; Tang et al., 2014; Velmeshev et al., 2019; Xu et al., 2020). When examining the ASD-specific DEGs of different cell types from human post-mortem brain samples (Velmeshev et al., 2019), the shared PPI network was enriched for DEGs in layer 2/3 and 4 neurons, parvalbumin and VIP interneurons, and protoplasmic astrocytes (Figure 2B). The enrichment of ASD DEGs of specific cell types highlights the human disease relevance of the 41 ASD-risk gene PPI network map.

Finally, we hypothesized a potential relationship between highly connected genes within the 41 ASD-risk gene PPI network map and human ASD behavioral phenotypes.

This would link gene clusters to human phenotypes, and provide additional insight into the biological basis of ASD. We took the individual PPI networks of the 41 ASD-risk genes and identified 3 groups (labeled Group 1, 2 and 3) of highly connected ASD-risk genes, using the correlation between their individual PPI networks (Murtaza et al., 2022). Groups 1 and 2 showed high connectivity between the ASD-risk genes within each group, whereas connectivity was lower in Group 3. To determine if grouping the 41 ASD-risk genes is correlated with clinical ASD behavioral scores based on shared PPI networks, we obtained clinical data of individuals with rare variants in the 41 ASD-risk genes from the MSSNG database. The database contained the sequenced genomes of a total of 4,258 families and 5,102 ASD-affected individuals at the time of data extraction (Yuen et al., 2017). We obtained the adaptive behavior and socialization scores from between 112-879 individuals, depending on the availability of scores, who possess at least one rare missense/splicing/LoF variant in the 41 ASD-risk genes (data-explorer.mss.ng). Remarkably, we found that individuals with missense variants in Group 1 genes had lower adaptive behavior standard scores compared to Groups 2 or 3, suggesting that missense variants impact the function of Group 1 genes in regards to adaptive behavior (Figure 5C, 5D, and S10D). However, individuals with variants impacting mRNA splicing in Groups 1 had significantly higher standard adaptive behavior and socialization scores compared to Group 2 or 3 (Figure 5E and 5F). Interestingly, the NRXN1 gene that is part of group 2 has been found to have alternative splicing in individuals with neuropsychiatric disorders (Flaherty et al., 2019). This suggests that splice variants may play a more prominent role in these groups with respect

to their effect on adaptive behavior and socialization scores. No significant differences were seen between individuals with frame shift or stop gain variants in genes from any group (Figure S10E and S10F), possibly due to the lower number of individuals in the analyses, or an equally detrimental impact of these variants on all ASD risk genes. The differences between Group 1 and Groups 2 or 3 suggest that PPI networks can be used to cluster ASD-risk genes, and individuals with variants in those genes. Group 1 genes were found to have the largest enrichment of ASD-risk genes (Figure S10G), suggesting that the highly interconnected PPI networks and shared pathways for this group of genes may be a core driver for the affected clinical phenotypes (Figure S10G). The functional grouping of ASD-risk genes highlights the potential of using PPI networks to correlate biological function with clinical phenotype. This could lead to a better approach in subdividing individuals with ASD and understanding the biological basis of these subgroups.

4.5 DISCUSSION

ASD-associated *de novo* missense mutations are enriched in hub genes of known protein interaction networks (Chen et al., 2018, 2020). However, few studies have used proximity-labeling to study the impact of disease-relevant mutations on the PPI network of genes associated with neurodevelopmental or neurological disorders (Chou et al., 2018; Pintacuda et al., 2021; Tracy et al., 2021). Our BioID2 screen provides functional evidence of the impact ASD-associated *de novo* missense variants have on the PPI network of three ASD-risk genes, as examples. PPP2R5D is the regulatory subunit of

PP2A, which has been associated with Golgi assembly through the beta subunit (Lowe et al., 2000; Schmitz et al., 2010). The impact of PPP2R5D variants on the Golgi apparatus morphology and localization highlights a potential novel role of the ASD-risk gene, PPP2R5D, and Golgi apparatus function ASD pathology. One caveat to our technique was the variability between separate runs that required reduced stringency in our comparisons between wildtype and mutant proteins, therefore although possible changes in the PPI networks were identified, functional validation is important to filter out changes caused by background noise. Time- and resource-intensive studies have also investigated multiple variants in single genes in various animal models (Post et al., 2020; Meili et al., 2021). Additional bioinformatic approaches have been used to determine the pathogenicity of rare missense variants; however, the impact on biological pathways remains to be determined (Koire et al., 2021; Wu et al., 2021). Using neuron-specific PPI networks allows the use of a relevant cell type, while being able to scale up the screen to test multiple single-gene variants in a nonbiased manner and reduced period of time. This approach could have potential applications for variants of unknown significance by providing important information on the severity of a given genetic variant.

The enrichment of an additional 112 ASD risk genes in the shared ASD-risk gene PPI network map and the enrichment of the network in ASD-associated cell types further emphasizes the interconnectedness of ASD risk proteins. Mid-fetal deep cortical projection neurons and superficial cortical glutamatergic neurons are enriched for ASD-risk genes and are associated with autism pathology (Parikshak et al., 2013; Willsey et al., 2013). The ASD-shared PPI network was highly enriched for genes expressed in

excitatory and inhibitory neurons, and for DEGs in individuals with ASD specific to Layers 2/3 excitatory neurons and VIP interneurons. The high connectivity and enrichment of ASD-risk genes within the network reflect its relevance to shared pathways associated with ASD pathology. Future studies will need to distinguish which ASD PPI networks are specific to each cell type, or possible subpopulations, to understand the subtle network changes that impact disease mechanisms.

Of great interest was the ability to group the 41 ASD-risk genes based on their PPI network, and the correlation of these groups to clinical scores in adaptive behavior and socialization. Although we focused specifically on missense/LoF variants, we found that the type of variant in each group of genes influenced the average score of the individuals within the group. However, the overall effect size of the analysis was small, suggesting that while we identified some differences in ASD-risk genes grouped by their BioID PPI network, further BioID of ASD-risk genes or methods to consider the genetic background of individuals is required before any strong classification can be made. To work through the complexity, it may be important to combine our analysis with other methods of categorizing mutations (e.g. gnomAD pLI, Polyphen-2) as higher or lower impact (Adzhubei et al., 2010; Lek et al., 2016), which would reduce the number of people shared between groups. Based on our findings, we highlight the ability to group ASD-risk genes based on their PPIs and correlate the groups to differences in clinical scores related to ASD.

4.6 CONCLUSION

In conclusion, our proximity-labeling proteomic method to study the impact of ASD-risk genes variants identified in individuals with ASD, reveals an unbiased potential method of high-throughput screening at the protein level. Furthermore, the comparison of PPI networks to large-scale human clinical and genetic datasets demonstrates a step towards grouping ASD individuals and risk genes based on biological evidence. Ultimately, the hope is that this approach may translate into a better understanding of wide-ranging ASD clinical phenotypes and the development of targeted therapies.

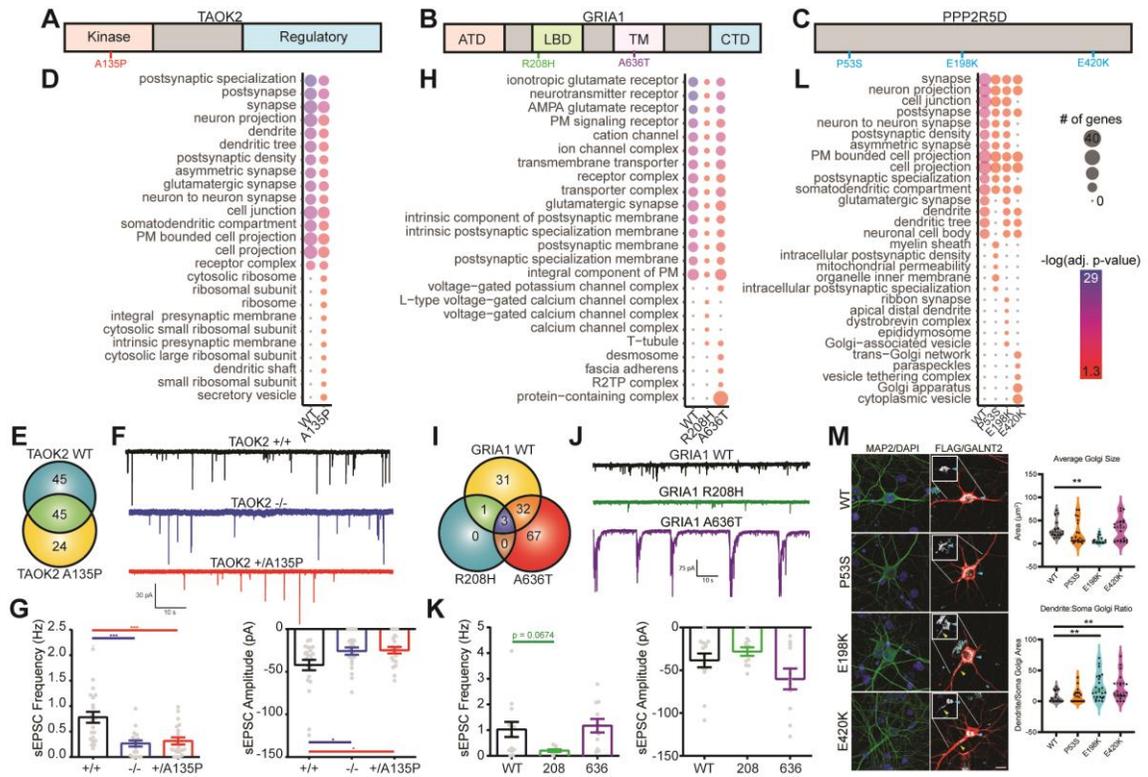


Figure 1. *De novo* missense variants alter the PPI networks of uncharacterized and established ASD-risk genes that correspond to functional deficits. (A-C), Diagram of TAOK2, GRIA1, and PPP2R5D and location of *de novo* missense variants. ATD: amino-terminal domain, LBD: ligand-binding domain, TM: transmembrane domain, CTD: carboxy-terminal domain. (D) Dot plot of top 15 cellular compartment gene sets and top 10 variant-specific gene sets identified in TAOK2 variants (g: Profiler, Benjamini-Hochberg FDR adj. $p < 0.05$). PM: plasma membrane. Size of dots indicate protein number and the color represents the significance. (E) Venn diagram of PPI network proteins of TAOK2 WT and A135P. (F) Representative traces of sEPSC recordings of DIV21 TAOK2 WT, KO, and A135P human iPSC-derived NGN2 neurons. (G) TAOK2 KO and A135P neurons show decreased sEPSC frequency (*left*) and amplitude (*right*)

(ROUT Outlier Test, $Q = 0.1\%$; frequency: One-Way ANOVA, $F(2,57) = 11.63$, $p < 0.0001$, amplitude: One-Way ANOVA, $F(2,63) = 4.027$, $p = 0.0226$, post hoc Holm-Sidak test; WT = 23, KO = 22, and A135P = 21 neurons from three separate transductions). (H) Dot plot of top 15 cellular compartment gene sets and top five variant-specific gene sets identified in GRIA1 variants (g: Profiler, Benjamini-Hochberg FDR adj. $p < 0.05$). (I) Venn diagram of PPI network proteins of GRIA1 WT and variants. (J) Representative traces of sEPSC recordings of mouse cortical neurons expressing GRIA1 or variants. (K) R208H variant shows trend in decrease sEPSC frequency (*left*) and no change in amplitude (*right*). (ROUT Outlier Test, $Q = 0.1\%$; frequency: One-Way ANOVA, $F(2,31) = 3.506$, $p = 0.0424$, amplitude: One-Way ANOVA, $F(2,33) = 3.147$, $p = 0.0561$, post hoc Holm-Sidak test; WT = 14, R208H = 11, and A636T = 11 neurons from three separate infections). (L) Dot plot of top 15 cellular compartment gene sets and top five variant-specific gene sets identified in PPP2R5D variants (g: Profiler, Benjamini-Hochberg FDR adj. $p < 0.05$). (M) Representative images of neurons infected with PPP2R5D WT and variants show altered Golgi morphology, stained by GALNT2 and localization to the dendrite (white arrow) and soma (blue arrow) (*right*). Inlet shows Golgi staining alone. Scale bar is $20\mu\text{m}$. Neurons expressing the E198K variant have decreased Golgi size, while both the E198K and E420K variants cause mislocalization of the Golgi into the dendrite (*left*). (Average Golgi size: One-Way ANOVA, $F(3,76) = 5.773$, $p = 0.0013$, Dendrite: Soma Golgi ratio: One-Way ANOVA, $F(3,76) = 6.095$, $p = 0.0009$, post hoc Holm-Sidak test; 20 neurons from 5 separate infections per condition). Mean \pm s.e.m. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

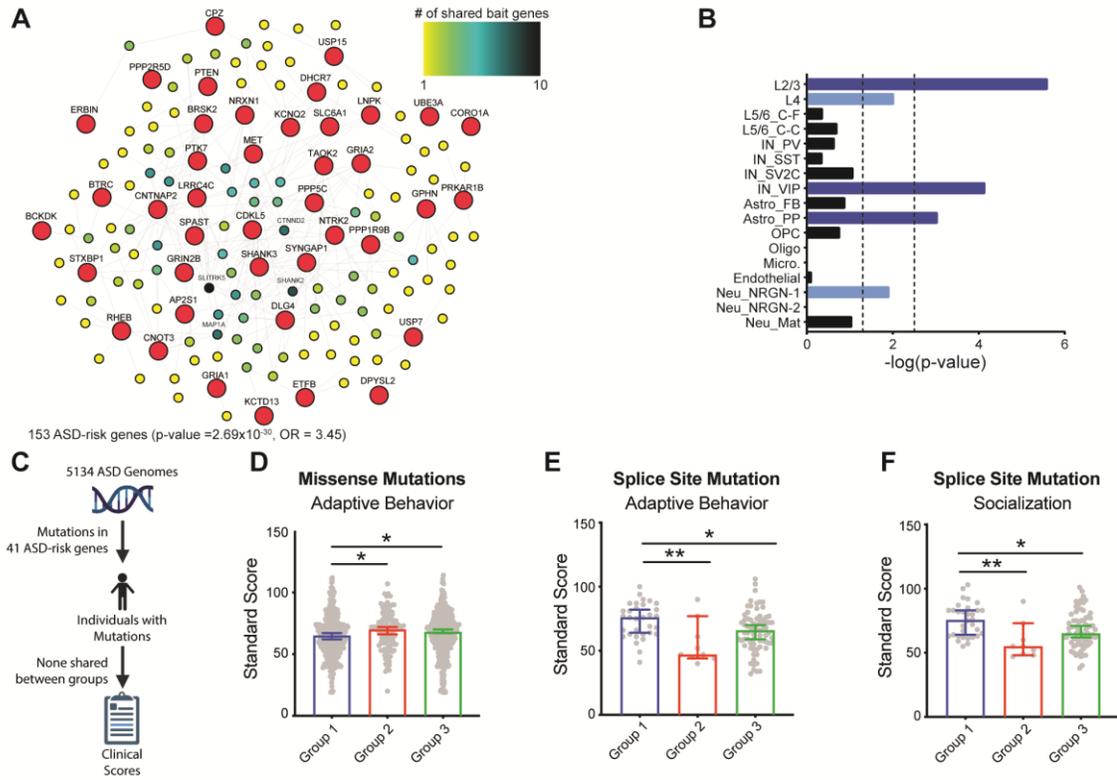


Figure 2. Shared ASD PPI network map correlates to human brain development

and disease pathology. (A) Network of ASD-risk genes enriched in the shared ASD-risk

gene PPI network. Large red nodes represent bait genes, while smaller colored nodes

represent sharedness of ASD-risk genes between bait genes, between 1 (yellow) and 10

(blue) shared bait genes (g: Profiler, Benjamini-Hochberg FDR adj. $p < 0.05$). (B)

Enrichment of ASD differentially expressed genes (DEGs) in specific cell types.

(Fisher's exact test). Dashed lines represent nominal ($p = 0.05$, left) and Bonferroni

corrected ($p = 0.05/\text{number of cell types}$, right) significance thresholds. EN = excitatory

neurons, IN = inhibitory neurons, RGC = radial glial cells, MGE RGC = medial

ganglionic eminence, IPC = intermediate progenitor cells, Astro. = astrocyte, OPC =

oligodendrocyte progenitor cells, Micro. = microglia, Endo. = endothelial cells, CP =

choroid plexus cells, C-F = cortico-fugal, C-C = cortico-cortico, PV = parvalbumin, SST = somatostatin, VIP = vasoactive intestinal peptide, FB = fibrous, PP = protoplasmic, Neu_NRG1 = neurogranin-expressing. Light blue bars have nominal p-value significance, while dark blue bars have Bonferroni corrected significance. (C) Flow chart of clinical data extraction (D) Significant decrease in the average standard scores of individuals diagnosed with ASD, who have rare inherited missense mutations in Group 1 genes compared to Groups 2 and 3 (Non-parametric Kruskal-Wallis test, $p = 0.0103$, post hoc Dunn's test, Group 1 = 350, Group 2 = 113, and Group 3 = 416 probands). Individuals with splice site mutations in Cluster 1 have significantly higher adaptive behavior (E) and socialization standard (F) scores than Cluster 2 and 3 (Non-parametric Kruskal-Wallis test, adaptive behavior: $p = 0.0036$, Group 1 = 32, Group 2 = 9, and Group 3 = 72 probands; socialization: $p = 0.0021$, Group 1 = 32, Group 2 = 9, and Group 3 = 71 probands; post hoc Dunn's test). Box and whisker plot (minimum, 1st quartile, median, 3rd quartile, maximum).

4.7 SUPPLEMENTAL DATA

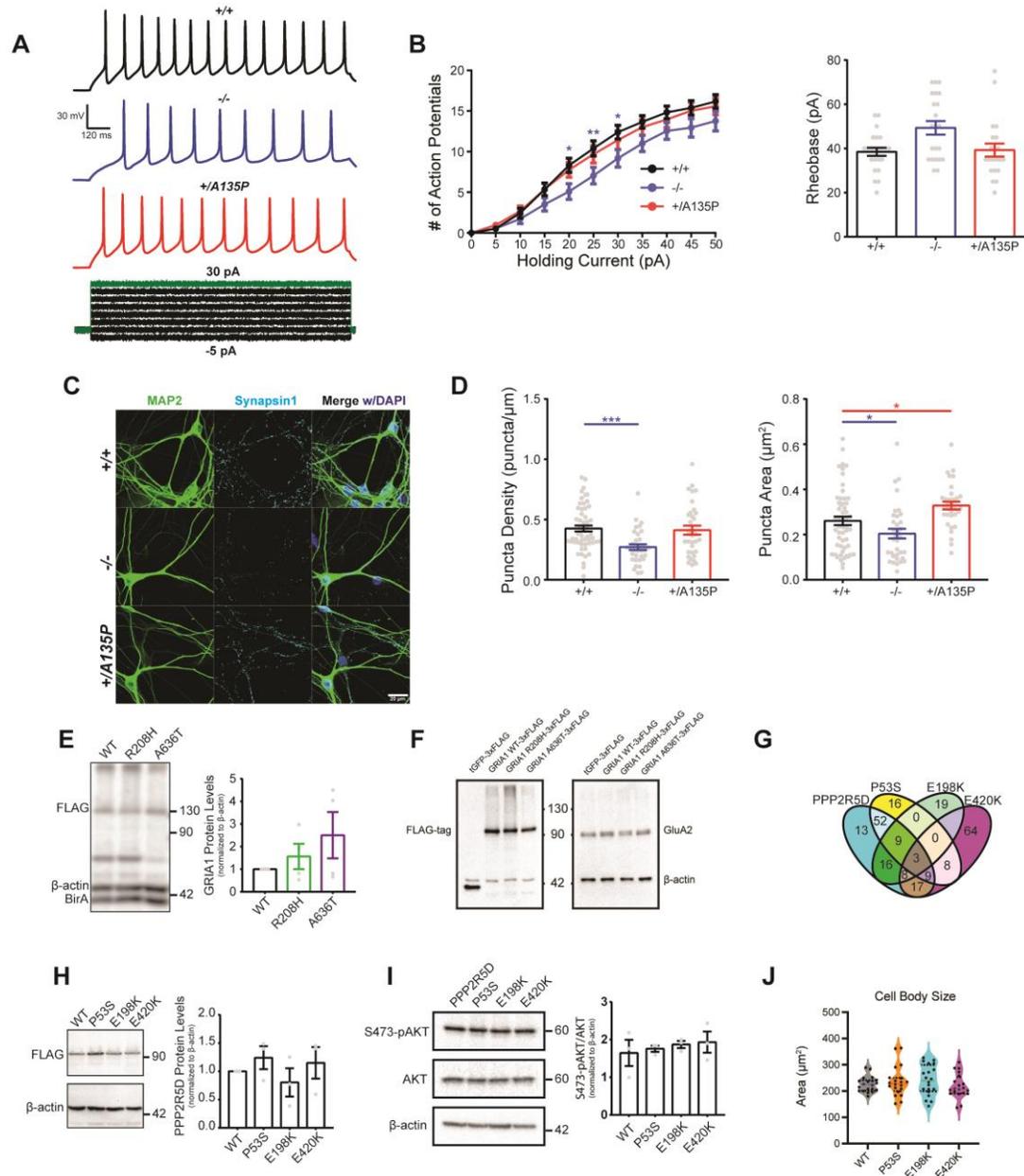


Figure S1. *De novo* mutations in TAOK2 caused altered synaptic transmission and neuron firing. (A) Representative traces of repetitive firing (*top*) and current injection (*bottom*). (B) TAOK2 KO neurons have reduced repetitive firing (*left*) (Two-Way

ANOVA, $F(2, 693) = 17.6$, $p < 0.0001$ between genotypes; WT = 23, KO = 22, A135P = 21 neurons from 3 separate transductions) increased rheobase (*right*) (One-Way ANOVA, $F(2, 63) = 5.229$, $p = 0.0079$; WT = 23, KO = 22, A135P = 21 neurons from 3 separate transductions). (C) Representative images of TAOK2 WT (+/+), KO, and A135P human neurons, stained with MAP2, Synapsin1, and DAPI 21 days after NGN2 induction. (D) Reduced synapsin1 punctae density and size in TAOK2 KO neurons (*left*) and increased synapsin1 puncta size in TAOK2 A135P neurons (*right*) (ROUT outlier test, $Q = 0.1\%$, One-Way ANOVA, density: $F(2, 120) = 8.104$, $p = 0.0005$, area: $F(2, 119) = 8.207$, $p = 0.0005$, post hoc Holm-Sidak test; WT = 55, KO = 35, A135P = 34 neurons from five separate transductions). (E) Representative western blot of GRIA1 WT, R208H, and A636T BioID2 constructs expressed in HEK293 cells (*left*) and quantification (*right*) showing no significant difference (One-sample t-test, WT vs R208H: $t = 1.008$, $df = 3$, $p = 0.3877$, WT vs A636T: $t = 1.466$, $df = 3$, $p = 0.2388$; four separate transfections). (F) Western blot of GluA2 subunit in DIV16 cortical neurons infected with GRIA1 WT, R208H, and A636T-FLAG lentiviral constructs for 4 days (G) Venn diagram of PPI network proteins of PPP2R5D WT and variants. (H) Representative western blot of PPP2R5D WT, P53S, E198K, and E420K BioID2 constructs expressed in HEK293 cells (*left*) and quantification (*right*) showing no difference in expression (One-sample t-test, WT vs P53S: $t = 1.185$, $df = 2$, $p = 0.3577$, WT vs E198K: $t = 0.7838$, $df = 2$, $p = 0.5152$, WT vs E420K: $t = 0.5371$, $df = 2$, $p = 0.6449$; three separate transfections). (I) Representative western blot of p-AKT and AKT in HEK293 cells expressing PPP2R5D WT or variant constructs expressed in HEK293 cells (*left*) and quantification

(*right*) showing no difference in the pAKT/AKT ratio (One-way ANOVA, $F(3,8) = 0.2948$, $p = 0.8283$; three separate transfections). Quantification of western blots in (E), (H), and (I) were done with the volume intensity of the Flag-tag band. (J) Neurons expressing the PPP2R5D variants show no changes in neuron size (One-Way ANOVA, $F(3,76) = 0.825$, $p = 0.484$, post hoc Holm-Sidak test; 20 neurons from 5 separate infections per condition). Mean \pm s.e.m. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

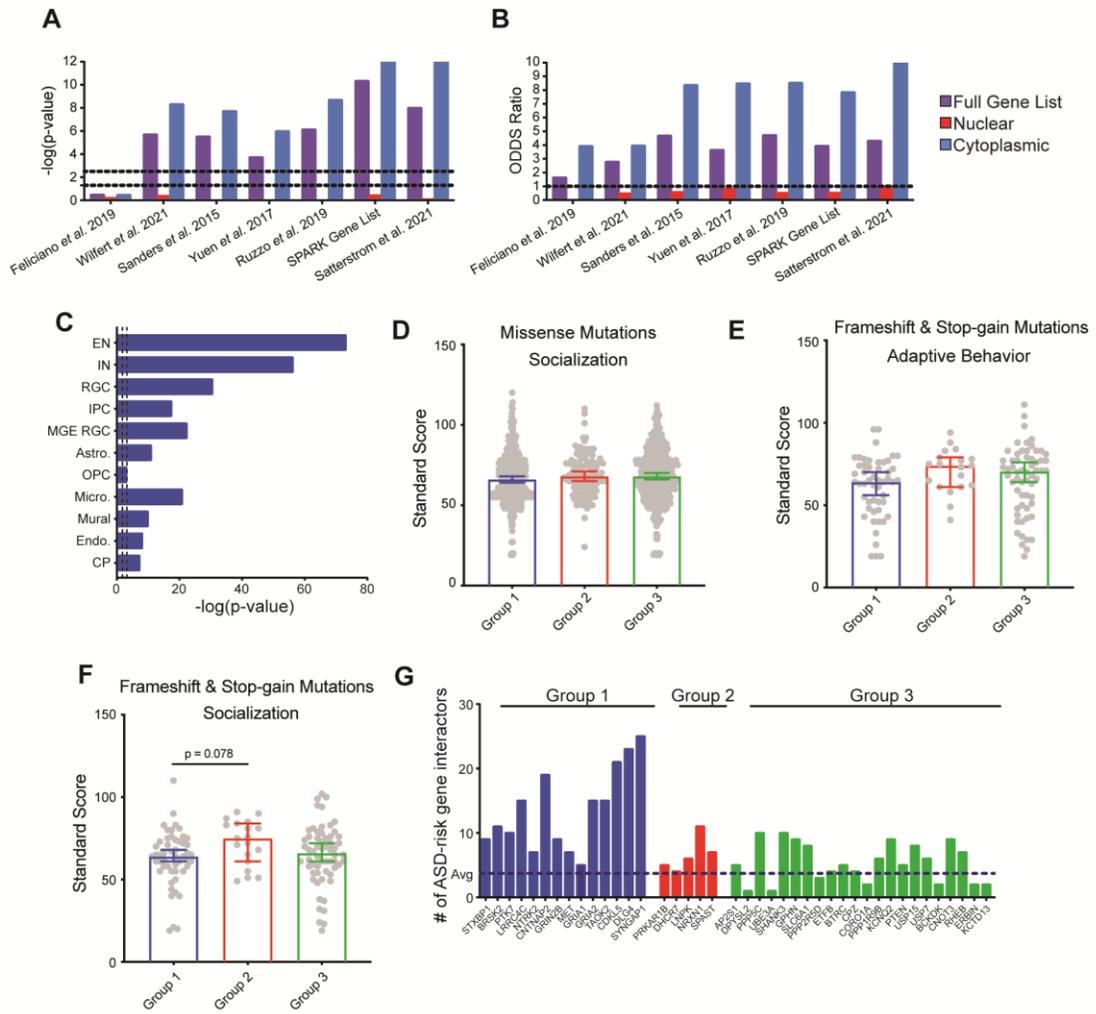


Figure S2. Enrichment of ASD risk-genes in the shared ASD PPI network map and

grouping of clinical phenotypes. (A) Enrichment of full gene list, cytoplasmic gene only

lists, and nuclear gene only lists from published works and SPARK, in the shared ASD-

risk gene PPI network shown by significance. (Fisher’s exact test). Dashed lines represent

nominal ($p = 0.05$, *left*) and Bonferroni corrected ($p = 0.05/\text{number of cell types}$, *right*)

significance thresholds. (B) ODDs ratio of full gene list, cytoplasmic gene only lists, and

nuclear gene only lists enriched in the shared ASD-risk gene PPI network. (C) The shared

ASD-risk gene PPI network enriches for human neuron cell types (Fisher's exact test). Dashed lines represent nominal ($p = 0.05$, *left*) and Bonferroni corrected ($p = 0.05/\text{number of cell types}$, *right*) significance thresholds. EN = excitatory neurons, IN = inhibitory neurons, RGC = radial glial cells, MGE RGC = medial ganglionic eminence, IPC = intermediate progenitor cells, Astro. = astrocyte, OPC = oligodendrocyte progenitor cells, Micro. = microglia, Endo. = endothelial cells, CP = choroid plexus cells, C-F = cortico-fugal, C-C = cortico-cortico, PV = parvalbumin, SST = somatostatin, VIP = vasoactive intestinal peptide, FB = fibrous, PP = protoplasmic, Neu_NRGN = neurogranin-expressing. Light blue bars have nominal p-value significance, while dark blue bars have Bonferroni corrected significance. (D) Individuals with missense mutations in Cluster 1, 2 and 3 genes show no significant differences in socialization standard scores (Non-parametric Kruskal-Wallis test, $p = 0.1765$, post-hoc Dunn's test; Group 1 = 351, Group 2 = 114, and Group 3 = 416 probands). Individuals with frameshift or stop gain mutations in Cluster 1, 2 and 3 genes show no significant differences in adaptive behavior (E) and socialization standard (F) scores (Non-parametric Kruskal-Wallis test, adaptive behavior: $p = 0.1069$, Group 1 = 51, Group 2 = 19, and Group 3 = 60 probands; socialization: $p = 0.0803$, Group 1 = 51, Group 2 = 19, and Group 3 = 60 probands; post hoc Dunn's test). (G) Number of ASD-risk genes identified in each of the 41 ASD-risk gene protein-protein interactions. Dashed line represents average expected risk genes. Box and whisker plot (minimum, 1st quartile, median, 3rd quartile, maximum). * $p < 0.05$, ** $p < 0.01$.

Tables are available at <https://doi.org/10.1101/2022.01.17.476220>

Table S8. Comparison of BioID2 PPI networks between ASD-risk genes and their variants

Table S9. BioID2 PPI network enriched pathways of ASD-risk genes and their variants

Table S10. List of sources for 41 ASD-risk genes and cellular compartment genes

CHAPTER 5: CONCLUSIONS AND FUTURE DIRECTIONS

5.1 Elucidation of *TAOK2* function in neurodevelopment and using it to study ASD pathology

In our first chapter we characterized the role of *TAOK2* as a novel ASD-risk gene. Prior to our published work, few studies showed the importance of *TAOK2* during brain development and its potential disruption in ASD (Calderon de Anda et al., 2012; Ultanir et al., 2014; Yadav et al., 2017). However, our work was the first to highlight the function of *TAOK2* in neuron maturation, activity, and connectivity, and show how disruption of *TAOK2* cause deficits associated with ASD pathophysiology and behaviors. Studies have identified significant changes in cortical neurons of individuals with ASD (Velmeshev et al., 2019; Willsey et al., 2022), which coincided with the increased brain size, decreased relative cortex size, and reduced neuron dendrite arborization seen in the *Taok2* KO mice. In fact, we found that neuron morphological and functional deficits caused by loss of *Taok2* were localized to the cortex. These gross changes across the brain coincided with the ASD-related anxiety and social interaction behaviors demonstrated by the mice, which have only been seen in a few genetic mouse models of ASD-risk genes. As described in Chapter 1, E/I imbalance is a major hypothesis underlying ASD pathology. Here we saw that at the cellular level, loss of *TAOK2* caused reduced dendritic spine formation and activity in the PFC and somatosensory cortex (SSC), with specific excitatory post-synaptic current impairment. However, patch-clamp recordings were only

done on the PFC and SSC regions, and changes in the inhibitory neuron inputs may be localized to other regions.

Prior to our study the genetic risk of *TAOK2* for ASD, was solely based on its location in the 16p11.2 locus, the most common CNV region associated with ASD (Weiss et al., 2008). Through our collaboration with the Scherer lab and their Autism Speaks MSSNG project, which has sequenced over 10,000 individuals, multiple genetic variants in the *TAOK2* gene in individuals with ASD were identified and not published in their initial release (Yuen et al., 2017). Furthermore, another study identified two additional *de novo* variants in *TAOK2*, highlighting the its disruption as a risk for NDDs (Deciphering et al., 2016). Our study highlights the use of biological assays to test the functional impact of genetic variants, as we found that two out of five functionally tested variants had significant functional outcomes. Here we elucidated the impact of two *de novo* mutations, A135P and P1022*. The A135P mutation creates a kinase-dead *TAOK2* that acted as a dominant negative mutant and caused similar deficits as the loss of *TAOK2*. While, the P1022* mutation creates an overactive and unstable form of *TAOK2*, which increased dendritic growth with more complex dendritic spine phenotypes. The contrasting effects of these two *de novo* mutations, underscore the necessity and benefits of studying individual mutations, due to the possibility of differing functional impacts. However, our investigation into the impact of *de novo* and rare-inherited variants underscores the time and resources required to determine the functional changes caused multiple variants. Although we obtained disease-relevant information of how these mutations altered *TAOK2* function and their potential effect on neuron function and

activity, it would not be feasible to use the same methodology to study many genes simultaneously. This set the stage for Aim 3 (Chapter 4), which focused on creating a method to study multiple ASD-risk genes in a functionally unbiased manner.

We also identified novel signaling pathway for *TAOK2* not previously observed. Disruption of *TAOK2* caused deficits in active-RhoA levels in both mouse cortical neurons and patient-derived LCLs. RhoA is a GTPase protein involved in regulating F-actin stability (Sit and Manser, 2011). Aberrant RhoA activity caused altered dendritic spine motility in *Taok2* KO mice, showcasing a novel signaling pathway regulated by *TAOK2*. Interestingly, *KCTD13*, another 16p11.2 gene, has been associated with RhoA expression, which may highlight potential cross-talk between two genes in the same CNV (Lin et al., 2015). We rescued the spine motility deficits by introducing a RhoA activator peptide, highlighting the potential for a future therapeutic target for ASD pathologies. Together, our investigation into the function of *TAOK2* during brain development is an example of an in-depth single gene study to unravel the pathology underlying a neurodevelopmental disorder, such as ASD. Our study provided a comprehensive understanding of *TAOK2* that could be used by others in the field to further investigate other aspects of neurodevelopment.

Since publication, other groups have identified possible links of *TAOK2* to novel biological pathways, such as endoplasmic reticulum function and localization (Nourbakhsh et al., 2021). We have also identified novel functions of *TAOK2* regards its connection to the TCA cycle and mitochondrial cellular respiration, as described in Chapter 3. This shows that *TAOK2*, while only one ASD-risk gene, can be a model for multiple ASD-

risk genes that share similar pathophysiology, and could be a potential mouse model to test therapeutic targets at the cellular and behavioral level.

5.2 Identification of shared ASD-risk gene PPI network and convergent pathways between ASD-risk genes using proximity-labeling proteomics

As described in Chapter 1, many signaling pathways and biological processes are involved in the neuropathology of ASD. The large number of ASD-risk genes identified in the last decade fall within these broad categories, yet how they work together is not fully understood. We addressed this gap in knowledge using proximity-labeling proteomics, which can be used to identify the proteins that interact directly or are in close proximity to a POI.

Proximity-labeling proteomics has primarily been used in cell lines to study the interaction networks of proteins, with many studies focusing on identifying the protein composition of specific complexes, organelles, or cellular compartments (Markmiller et al., 2018; Youn et al., 2018; Fazal et al., 2019; Antonicka et al., 2020; Go et al., 2021). Few studies have used proximity-labeling in relevant brain cell-types, and those that have, only identified the PPI networks of a few proteins (Loh et al., 2016; Uezu et al., 2016; Chung et al., 2017; Spence et al., 2019; Hamdan et al., 2020; Takano et al., 2020). Importantly, we showed that while BioID of ASD-risk genes in HEK293 cells will identify a PPI network, neuron specific proteins and enriched pathways will not be identified. Therefore, to determine the PPI network relevant to a disease, cell types associated with the disease should be used, which in our case was cortical neurons.

Our work was the first to screen many ASD-risk genes through proximity-based biotin labeling proteomics in neurons. It utilized the least number of cells and protein compared to previous studies, and used multiplex isobaric-tag labeling to allow direct comparison in abundance of biotinylated proteins between samples (Loh et al., 2016). We identified the PPI networks of 41 ASD-risk genes, which had a range of known and unknown functions, including protein modification (ubiquitination and phosphorylation), synaptic scaffolding and transmission, cell adhesion, and cytoskeleton organization. A major part of this aim was to create a database of PPI networks identified in neurons for ASD-risk genes, which can even be applied to other NDDs because of the overlap of risk genes. This database will provide other researchers with a novel starting point when studying a specific ASD-risk gene and our system was designed to be easily reproduced in labs that are not specialized in proteomics.

Our system showed variability between the size of PPI networks between ASD-risk genes, which could be due to biological differences. However, these differences are likely due to the lack of gene-specific optimization. The large number of ASD-risk genes used in this study prevented true optimization for each gene and using gene specific controls was infeasible with the resources we had. Another caveat of BioID is that it cannot distinguish between direct interactions and proteins within close proximity of the POI. However, the importance of close proximity proteins should not be ignored. We were able to show through metabolic seahorse assays and immunofluorescence staining that less strongly enriched pathways in the PPI networks of ASD-risk genes do have strong impacts on neuronal processes (Chapter 2) and organelle morphology (Chapter 3).

Nonetheless, our identification of convergent pathways enriched in the shared PPI network, included known ASD-associated pathways, such as synaptic transmission pathways, WNT signaling, MAPK signaling and potassium channel activity. Most studies have focused on studying deficits in synaptic transmission in regards to ASD pathophysiology, therefore we focused on mitochondria and metabolism related pathways. Mitochondrial function and metabolism have been associated with multiple neurodevelopmental disorders and ASD; however, these were not directly tied to any ASD-risk genes (Cheng et al., 2017; Frye, 2020). In fact, mitochondrial genes are minimally identified through sequencing studies and are not strongly associated with ASD. The identification of TCA cycle and ETC proteins as part of the PPI network of 28 out of 41 ASD-risk genes, which were not previously associated with these pathways was unexpected and highlight the potential novel function that these genes have on cell metabolism. Future studies should focus on determining the potential role of other ASD-risk genes on mitochondrial function and metabolism. The study of these pathways opens up new potential avenues for treatments through modulation of metabolic processes or mitochondrial activity through the use of natural metabolites at important periods during development (Neal et al., 2008).

5.3 Proximity-labeling proteomics as a screening method to study the functional impact of genetic variants associated with ASD

The significant enrichment of *de novo* mutations in ASD-risk genes and their large effect on predicted interactions, highlight the importance of investigating the impact

of genetic variants on PPIs (Chen et al., 2018). We found that through proximity-labeling proteomics, we could identify significant changes in the PPI network of mutant TAOK2, GRIA1, and PPP2R5D proteins, that coincided with measurable functional changes. The benefit of using BioID was that it did not focus on any one type of assay. For example, in Chapter 2, we overexpressed the TAOK2 mutant proteins in HEK293 cells to measure protein levels, simultaneously with the JNK protein to measure the phosphorylation capabilities, and used *in utero* electroporation to study the effect of the TAOK2 mutants on neuron dendritic arborization and dendritic spine formation. Although we observed many deficits caused by the *de novo* mutations, we could only identify changes in the assays we performed and with two completely unknown variants. Proximity-labeling provided a starting point to streamline functional assays and prioritize the variants that are most likely to disruption protein interactions.

ASD-risk genes have long been hypothesized to function in convergent pathways, however this has not been directly shown at the protein interaction level. Most studies focus on co-expression modules and have found regulation of ASD-risk genes by transcription factors or chromatin remodeling proteins (Rylaarsdam and Guemez-Gamboa, 2019). Our study found that ASD-risk genes are enriched in the shared PPI network of 41 ASD-risk genes, more than expected even in a brain-specific protein background. This highlights the functional associations of ASD-risk genes at the protein level, especially between neural communication and cytoskeleton organization proteins.

The complex clinical phenotypes exhibited by individuals with ASD, has made it difficult to reliably diagnose or group specific individuals based on their clinical

symptoms. Many studies have attempted to correlate genetic mutations and altered molecular pathways to complex clinical phenotypes (Bruining et al., 2014; Cederquist et al., 2020; Warriier et al., 2022). These correlations are easier in smaller, less complex organisms (Wong et al., 2019; Post et al., 2020). In Chapter 4, we showed that BioID of ASD-risk genes has the potential to group these genes based on the shared PPI networks, and also group individuals that have mutations in these genes. These grouped individuals showed significant differences in adaptive behavior and socialization test scores, which were found to depend on the type of mutation they possessed, such as exonic missense mutations, splice site mutations, or LoF frameshift/premature stop codon mutations. Although our effect size was small, with BioID of more ASD-risk genes, larger datasets, and analysis tools that can control for the presence of multiple genetic variants in the same individuals, we could potentially identify distinct categories of ASD based on the association of ASD-risk genes. Grouping ASD-risk genes based on their protein interactions shows potential in identifying subsets of ASD, which currently do not exist and has created difficulties in studying the heterogeneous disorder.

5.4 Future studies using proximity-label proteomics to elucidate disease-relevant mechanisms

Future studies that utilize proximity-labeling to study ASD or other NDD risk genes should focus on using human iPSCs to better recapitulate and identify human disease-relevant PPI networks. Although we found similar phenotypes in *TAOK2* KO mouse and human neurons, this may not be the case for other genes. As mentioned, in

Chapter 1, mouse brain development has many significant differences compared to human brain development.

Our system identified the PPI network at a very specific time point in mouse cortical neurons, and will only reflect the function of the protein at those time points. During neurodevelopment, the function and importance of proteins change quickly. With technologies such as the Tet-On doxycycline inducible system, and cell-type specific promoters, PPI networks of genes can be identified in a spatial and temporally regulated manner (Das et al., 2016). Our study focused on identifying the PPI network of ASD-risk genes in neurons, however multiple other cells types have been associated with ASD, including astrocytes and microglia (Velmeshev et al., 2019).

Furthermore, newer versions of BioID, such as TurboID and Split-TurboID, can be used to identify PPI networks after neural stimulation to study how the PPI network of proteins change due to neural activity (Branon et al., 2018; Cho et al., 2020). Split-TurboID can be used to study the interactions between cells, organelles, and even complexes. Split-TurboID requires the fusing of the two parts of the TurboID protein to function and can be used to identify proteins under specific circumstances, such as interaction between different brain cell types or at synaptic connections, which are altered in ASD. This system was recently used *in vivo* to identify the proteins involved at sites of contact between astrocytes and neurons (Takano et al., 2020).

Furthermore, overexpression of the BioID constructs may affect the health of the cell and introduce artifacts due to the overabundance of the protein. CRISPR/Cas9 editing techniques can be used to fuse the BioID proteins to genes endogenously, so the

expression is kept at levels typical for the cell. This was recently accomplished in a paper that studied the function of Tau in human iPSC-derived neurons, which also identified novel interactions of Tau with mitochondrial proteins (Tracy et al., 2022).

Future studies could use this system to identify changes in the protein composition of distinct compartments or PPI networks of important proteins in patient iPSC-derived neurons. For example, using the BioID to determine changes in the PPI network of PSD95 could elucidate differences in the synaptic compartment of idiopathic ASD cases, which have no specific genetic cause and make up the majority of ASD cases. Human iPSCs present a useful tool that can be combined with proximity-labeling proteomics to study disease relevant mechanisms.

5.5 Using 3D human stem cell derived organoids to study ASD pathophysiology

Future studies using single gene knockout or mutant models or discovery-based “omics” techniques should focus on using 3D human stem cell derived organoids to study disease relevant pathophysiology. Although we used human iPSCs to study the deficits caused by mutation or knockout of TAOK2, these neurons were directly differentiated from iPSCs using NGN2 and did not go through the NPC stage or have the type organization seen in a human brain (Hulme et al., 2022). Brain organoids recapitulate the developmental trajectories and organizations of brain, and exhibit cortical layer and cross region connectivity, although at a very simplistic level (Lancaster et al., 2013; Birey et al., 2017; Kelley and Paşca, 2022). Multiple region specific organoid generation protocols exist, including for the cerebral cortex, striatum, hippocampus, thalamus,

hypothalamus, midbrain, and cerebellum (Paşca et al., 2011; Lancaster et al., 2013; Kelley and Paşca, 2022). Recent studies have also studied connectivity between brain regions through fusions of dorsal-ventral, thalamo-cortical, cortico-striatal, hypothalamic-pituitary, and cortico-spinal-muscle organoids (Birey et al., 2017; Kelley and Paşca, 2022). They have been used to study multiple NDD models related to ASD (Timothy syndrome, 22q11.2, 16p11.2), and therefore future studies for *TAOK2* should focus on using brain organoids to determine if the disease pathologies we saw in the mouse model is also seen in human cell types (Khan et al., 2020; Urresti et al., 2021; Birey et al., 2022). The ability to study multiple cell-types and different time points during organoid development in a human background, combined with CRISPR/Cas-editing technologies and proximity-labeling proteomics, will reveal new avenues to study ASD-risk genes and signaling pathways associated with NDDs in human cell types that were previously not possible.

5.6 Significance

My thesis project characterized a novel ASD-risk gene, *TAOK2*, through the use of mouse and human models and investigation of the impact of patient mutations, highlighting the importance of single gene studies to elucidate disease-relevant pathophysiology. However, the advancements in DNA sequencing have created a bottle neck in the ability to study all ASD-risk genes. Proximity-labeling proteomics, which is a discovery-based technique, offered a method to identify convergent pathways and mechanisms of multiple ASD-risk genes and study the impact of patient mutations in a

less biased manner. A major issue in the ASD research field is the ability to categorize and group individuals with the disorder, our system was able to group individuals based on the ASD-risk gene interactions, and showed that ASD-risk genes do function in shared protein networks and pathways. Our work will aid in identifying therapeutic targets that can rescue the deficits caused by multiple ASD-risk genes, as well as provide a method to categorize the vast number of ASD-risk genes based on functional data to complement current diagnostic tools.

REFERENCES

- Abraham JR, Szoko N, Barnard J, Rubin RA, Schlatter D, Lundberg K, Li X, Natowicz MR (2019) Proteomic Investigations of Autism Brain Identify Known and Novel Pathogenetic Processes. *Sci Rep*:1–11.
- Abrisch RG, Gumbin SC, Wisniewski BT, Lackner LL, Voeltz GK (2020) Fission and fusion machineries converge at ER contact sites to regulate mitochondrial morphology. *J Cell Biol* 219.
- Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR (2010) A method and server for predicting damaging missense mutations. *Nat Methods* 7:248–249.
- Afgan E et al. (2018) The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res* 46:W537–W544.
- Alford SC, Ding Y, Simmen T, Campbell RE (2012) Dimerization-dependent green and yellow fluorescent proteins. *ACS Synth Biol* 1:569–575.
- Amir RE, Van Den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY (1999) Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet* 23:185–188.
- Amiri A, Cho W, Zhou J, Birnbaum SG, Sinton CM, McKay RM, Parada LF (2012) Pten deletion in adult hippocampal neural stem/progenitor cells causes cellular abnormalities and alters neurogenesis. *J Neurosci* 32:5880–5890.
- Anagnostou E, Taylor MJ (2011) Review of neuroimaging in autism spectrum disorders: What have we learned and where we go from here. *Mol Autism* 2:1–9.

- Ananiev G, Williams EC, Li H, Chang Q (2011) Isogenic pairs of wild type and mutant induced pluripotent stem cell (iPSC) lines from rett syndrome patients as In Vitro disease model. *PLoS One* 6.
- Aneichyk T et al. (2018) Dissecting the Causal Mechanism of X-Linked Dystonia-Parkinsonism by Integrating Genome and Transcriptome Assembly A Mendelian form of dystonia arises from altered splicing and intron retention within a general transcription factor. *Dissecting the Causal* . *Cell* 172:897–909.
- Ang CE, Olmos VH, Zhou B, Lee QY, Sinha R, Narayanaswamy A, Mall M, Chesnov K, Sudhof T, Wernig M (2020) the Dynamic Interplay Between Homeodomain Transcription Factors and Chromatin Environment Regulates Proneural Factor Outcomes. *bioRxiv:2020.12.02.398677*.
- Anitha A, Nakamura K, Thanseem I, Yamada K, Iwayama Y, Toyota T, Matsuzaki H, Miyachi T, Yamada S, Tsujii M, Tsuchiya KJ, Matsumoto K, Iwata Y, Suzuki K, Ichikawa H, Sugiyama T, Yoshikawa T, Mori N (2012) Brain region-specific altered expression and association of mitochondria-related genes in autism. *Mol Autism* 3:1–12.
- Antonicka H, Lin ZY, Janer A, Aaltonen MJ, Weraarpachai W, Gingras AC, Shoubridge EA (2020) A High-Density Human Mitochondrial Proximity Interaction Network. *Cell Metab* 32:479-497.e9.
- Arikkath J, Reichardt LF (2008) Cadherins and catenins at synapses: roles in synaptogenesis and synaptic plasticity. *Trends Neurosci* 31:487–494.
- Bader GD, Isserlin R, Merico D, Voisin V (2014) Enrichment Map - a Cytoscape app to

- visualize and explore OMICs pathway enrichment results. *F1000Research* 3:1–8.
- Bailey DB, Mesibov GB, Hatton DD, Clark RD, Roberts JE, Mayhew L (1998) Autistic Behavior in Young Boys with Fragile X Syndrome. 28.
- Barger CJ, Branick C, Chee L, Karpf AR (2019) Pan-cancer analyses reveal genomic features of FOXM1 overexpression in cancer. *Cancers (Basel)* 11:251.
- Barres BA (2008) The Mystery and Magic of Glia: A Perspective on Their Roles in Health and Disease. *Neuron* 60:430–440.
- Bateman A et al. (2021) UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Res* 49:D480–D489.
- Baxter AJ, Brugha TS, Erskine HE, Scheurer RW, Vos T, Scott JG (2015) The epidemiology and global burden of autism spectrum disorders. *Psychol Med* 45:601–613.
- Belmonte MK, Allen G, Beckel-Mitchener A, Boulanger LM, Carper R a, Webb SJ (2004) Autism and Abnormal Development of Brain Connectivity. *J Neurosci* 24:9228–9231.
- Ben-Shalom R, Keeshen CM, Berrios KN, An JY, Sanders SJ, Bender KJ (2017) Opposing Effects on NaV1.2 Function Underlie Differences Between SCN2A Variants Observed in Individuals With Autism Spectrum Disorder or Infantile Seizures. *Biol Psychiatry* 82:224–232.
- Bergen SE, Ploner A, Howrigan D, O’Donovan MC, Smoller JW, Sullivan PF, Sebat J, Neale B, Kendler KS (2019) Joint contributions of rare copy number variants and common SNPs to risk for schizophrenia. *Am J Psychiatry* 176:29–35.

Billstedt E, Gillberg C, Gillberg C (2005) Autism after Adolescence: Population-based 13- to 22-year Follow-up Study of 120 Individuals with Autism Diagnosed in Childhood. *J Autism Dev Disord* 35:351–360.

Billups B, Forsythe ID (2002) Presynaptic mitochondrial calcium sequestration influences transmission at mammalian central synapses. *J Neurosci* 22:5840–5847.

Birey F et al. (2017) Assembly of functionally integrated human forebrain spheroids. *Nature* 545:54–59.

Birey F, Li MY, Gordon A, Thete M V., Valencia AM, Revah O, Paşca AM, Geschwind DH, Paşca SP (2022) Dissecting the molecular basis of human interneuron migration in forebrain assembloids from Timothy syndrome. *Cell Stem Cell* 29:248-264.e7.

Bissonnette CJ, Lyass L, Bhattacharyya BJ, Belmadani A, Miller RJ, Kessler JA (2011) The controlled generation of functional basal forebrain cholinergic neurons from human embryonic stem cells. *Stem Cells* 29:802–811.

Blizinsky KD, Diaz-Castro B, Forrest MP, Schürmann B, Bach AP, Martin-de-Saavedra MD, Wang L, Csernansky JG, Duan J, Penzes P (2016) Reversal of dendritic phenotypes in 16p11.2 microduplication mouse model neurons by pharmacological targeting of a network hub. *Proc Natl Acad Sci*:201607014.

Bonnefont J, Tiberi L, van den Aemele J, Potier D, Gaber ZB, Lin X, Bilheu A, Herpoel A, Velez Bravo FD, Guillemot F, Aerts S, Vanderhaeghen P (2019) Cortical Neurogenesis Requires Bcl6-Mediated Transcriptional Repression of Multiple Self-Renewal-Promoting Extrinsic Pathways. *Neuron* 103:1096-1108.e4.

Boyer LF, Campbell B, Larkin S, Mu Y, Gage FH (2012) Dopaminergic Differentiation

of Human Pluripotent Cells. *Curr Protoc Stem Cell Biol* 22:1H.6.1-1H.6.11.

Braccioli L, Vervoort SJ, Adolfs Y, Heijnen CJ, Basak O, Pasterkamp RJ, Nijboer CH, Coffey PJ (2017) FOXP1 Promotes Embryonic Neural Stem Cell Differentiation by Repressing Jagged1 Expression. *Stem Cell Reports* 9:1530–1545.

Braganza A, Li J, Zeng X, Yates NA, Dey NB, Andrews J, Clark J, Zamani L, Wang XH, St Croix C, O’Sullivan R, Garcia-Exposito L, Brodsky JL, Sobol RW (2017) UBE3B is a calmodulin-regulated, mitochondrion-associated E3 ubiquitin ligase. *J Biol Chem* 292:2470–2484.

Branon TC, Bosch JA, Sanchez AD, Udeshi ND, Svinkina T, Carr SA, Feldman JL, Perrimon N, Ting AY (2018) Efficient proximity labeling in living cells and organisms with TurboID. *Nat Biotechnol* 36:880–898.

Braun JE, Huntzinger E, Fauser M, Izaurralde E (2011) GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets. *Mol Cell* 44:120–133.

Brennan KJ, Simone A, Jou J, Gelboin-Burkhardt C, Tran N, Sangar S, Li Y, Mu Y, Chen G, Yu D, McCarthy S, Sebat J, Gage FH (2011) Modelling schizophrenia using human induced pluripotent stem cells. *Nature* 473:221.

Brentani H, de Paula CS, Bordini D, Rolim D, Sato F, Portolese J, Pacifico MC, McCracken JT (2013) Autism spectrum disorders: An overview on diagnosis and treatment. *Rev Bras Psiquiatr* 35:62–72.

Broek JA, Guest PC, Rahmoune H, Bahn S (2014) Proteomic analysis of post mortem brain tissue from autism patients: evidence for opposite changes in prefrontal cortex and cerebellum in synaptic connectivity-related proteins. *Mol Autism* 5:41.

- Brot S, Auger C, Bentata R, Rogemond V, Ménigoz S, Chounlamountri N, Girard-Egrot A, Honnorat J, Moradi-Améli M (2014) Collapsin response mediator protein 5 (CRMP5) induces mitophagy, thereby regulating mitochondrion numbers in dendrites. *J Biol Chem* 289:2261–2276.
- Brown EA, Lutz JD, Davis TR, Gniffke EP, Vanschoiack AAW, Neier SC, Tashbook N, Nicolini C, Fahnestock M, Schrum AG, Smith SEP (2018) Clustering the autisms using glutamate synapse protein interaction networks from cortical and hippocampal tissue of seven mouse models. *Mol Autism* 9:1–16.
- Bruining H, Eijkemans MJ, Kas MJ, Curran SR, Vorstman JA, Bolton PF (2014) Behavioral signatures related to genetic disorders in autism. *Mol Autism* 5:1–12.
- Buiting K, Williams C, Horsthemke B (2016) Angelman syndrome-insights into a rare neurogenetic disorder. *Nat Rev Neurol* 12:584–593.
- Bülow P, Zlatic SA, Wenner PA, Bassell GJ, Faundez V (2021) FMRP attenuates activity dependent modifications in the mitochondrial proteome. *Mol Brain* 14:1–11.
- Butko MT, Yang J, Geng Y, Kim HJ, Jeon NL, Shu X, MacKey MR, Ellisman MH, Tsien RY, Lin MZ (2012) Fluorescent and photo-oxidizing TimeSTAMP tags track protein fates in light and electron microscopy. *Nat Neurosci* 15:1742–1751.
- Buxhoeveden DP, Semendeferi K, Buckwalter J, Schenker N, Switzer R, Courchesne E (2006) Reduced minicolumns in the frontal cortex of patients with autism. *Neuropathol Appl Neurobiol* 32:483–491.
- Calderon de Anda F, Rosario AL, Durak O, Tran T, Gräff J, Meletis K, Rei D, Soda T, Madabhushi R, Ginty DD, Kolodkin AL, Tsai L-H (2012) Autism spectrum disorder

- susceptibility gene TAOK2 affects basal dendrite formation in the neocortex. *Nat Neurosci* 15:1022–1031.
- Cantó C, Menzies KJ, Auwerx J (2015) NAD⁺ Metabolism and the Control of Energy Homeostasis: A Balancing Act between Mitochondria and the Nucleus. *Cell Metab* 22:31–53.
- Casanova EL, Switala AE, Dandamudi S, Hickman AR, Vandenbrink J, Sharp JL, Feltus FA, Casanova MF (2019) Autism risk genes are evolutionarily ancient and maintain a unique feature landscape that echoes their function. *Autism Res* 12:860–869.
- Causier B, Davies B (2002) Analysing protein-protein interactions with the yeast two-hybrid system. *Plant Mol Biol* 50:855–870.
- Cederquist GY, Tchieu J, Callahan SJ, Ramnarine K, Ryan S, Zhang C, Rittenhouse C, Zeltner N, Chung SY, Zhou T, Chen S, Betel D, White RM, Tomishima M, Studer L (2020) A Multiplex Human Pluripotent Stem Cell Platform Defines Molecular and Functional Subclasses of Autism-Related Genes. *Cell Stem Cell* 27:35-49.e6.
- Cetica V, Chiari S, Mei D, Parrini E, Grisotto L, Marini C, Pucatti D, Ferrari A, Sicca F, Specchio N, Trivisano M, Battaglia D, Contaldo I, Zamponi N, Petrelli C, Granata T, Ragona F, Avanzini G, Guerrini R (2017) Clinical and genetic factors predicting Dravet syndrome in infants with SCN1A mutations. *Neurology* 88:1037–1044.
- Chailangkarn T et al. (2016) A human neurodevelopmental model for Williams syndrome. *Nature In press*:1–25.
- Chamberlain SJ, Chen P, Ng KY, Bourgois-rocha F, Lemtiri-chlieh F, Levine ES, Lalande M (2010) Induced pluripotent stem cell models of the genomic imprinting

- disorders Angelman and Prader – Willi syndromes. *Pnas* *107*:17668–17673.
- Chambers SM, Fasano CA, Papapetrou EP, Tomishima M, Sadelain M, Studer L (2009) Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *27*.
- Chang J, Gilman SR, Chiang AH, Sanders SJ, Vitkup D (2015) Genotype to phenotype relationships in autism spectrum disorders. *Nat Neurosci* *18*:191–198.
- Chavali VRM, Haider N, Rathi S, Vratasha V, Alapati T, He J, Gill K, Nikonov R, Duong TT, McDougald DS, Nikonov S, O'Brien J, Mills JA (2020) Dual SMAD inhibition and Wnt inhibition enable efficient and reproducible differentiations of induced pluripotent stem cells into retinal ganglion cells. *Sci Rep* *10*:1–14.
- Chen JA, Peñagarikano O, Belgard TG, Swarup V, Geschwind DH (2015) The Emerging Picture of Autism Spectrum Disorder: Genetics and Pathology. *Annu Rev Pathol Mech Dis* *10*:111–144.
- Chen S, Fragoza R, Klei L, Liu Y, Wang J, Roeder K, Devlin B, Yu H (2018) An interactome perturbation framework prioritizes damaging missense mutations for developmental disorders. *Nat Genet* *50*:1032–1040.
- Chen S, Wang J, Cicek E, Roeder K, Yu H, Devlin B (2020) De novo missense variants disrupting protein-protein interactions affect risk for autism through gene co-expression and protein networks in neuronal cell types. *Mol Autism* *11*:1–16.
- Chen Z, Cobb MH (2001) Regulation of Stress-responsive Mitogen-activated Protein (MAP) Kinase Pathways by TAO2. *J Biol Chem* *276*:16070–16075.
- Chen Z, Hutchison M, Cobb MH (1999) Isolation of the Protein Kinase TAO2 and

- Identification of Its Mitogen-activated Protein Kinase/Extracellular Signal-regulated Kinase Kinase Binding Domain. *J Biol Chem* 274:28803–28807.
- Cheng N, Rho JM, Masino SA (2017) Metabolic dysfunction underlying autism spectrum disorder and potential treatment approaches. *Front Mol Neurosci* 10:1–12.
- Cherra SJ, Steer E, Gusdon AM, Kiselyov K, Chu CT (2013) Mutant LRRK2 elicits calcium imbalance and depletion of dendritic mitochondria in neurons. *Am J Pathol* 182:474–484.
- Cheung AYL, Horvath LM, Grafodatskaya D, Pasceri P, Weksberg R, Hotta A, Carrel L, Ellis J (2011) Isolation of MECP2-null Rett Syndrome patient hiPS cells and isogenic controls through X-chromosome inactivation. *Hum Mol Genet* 20:2103–2115.
- Cho KF, Branon TC, Rajeev S, Svinkina T, Udeshi ND, Thoudam T, Kwak C, Rhee HW, Lee IK, Carr SA, Ting AY (2020) Split-TurboID enables contact-dependent proximity labeling in cells. *Proc Natl Acad Sci U S A* 117.
- Chou C et al. (2018) TDP-43 pathology disrupts nuclear pore complexes and nucleocytoplasmic transport in ALS/FTD. *Nat Neurosci* 21.
- Chung CY, Khurana V, Yi S, Sahni N, Loh KH, Auluck PK, Baru V, Udeshi ND, Freyzon Y, Carr SA, Hill DE, Vidal M, Ting AY, Lindquist S (2017) In Situ Peroxidase Labeling and Mass-Spectrometry Connects Alpha-Synuclein Directly to Endocytic Trafficking and mRNA Metabolism in Neurons. *Cell Syst* 4:242-250.e4.
- Clement JP, Aceti M, Creson TK, Ozkan ED, Shi Y, Reish NJ, Almonte AG, Miller BH, Wiltgen BJ, Miller CA, Xu X, Rumbaugh G (2012) Pathogenic SYNGAP1

mutations impair cognitive development by disrupting maturation of dendritic spine synapses. *Cell* 151:709–723.

Colvert E, Tick B, McEwen F, Stewart C, Curran SR, Woodhouse E, Gillan N, Hallett V, Lietz S, Garnett T, Ronald A, Plomin R, Rijdsdijk F, Happé F, Bolton P (2015) Heritability of Autism Spectrum Disorder in a UK Population-Based Twin Sample. *JAMA Psychiatry* 72:415.

Comery TA, Harris JB, Willems PJ, Oostra BA, Irwin SA, Weiler IJ, Greenough WT (1997) Abnormal dendritic spines in fragile X knockout mice: Maturation and pruning deficits. *Proc Natl Acad Sci U S A* 94:5401–5404.

Constantino JN, Zhang Y, Frazier T, Abbacchi AM, Law P (2010) Sibling recurrence and the genetic epidemiology of autism. *Am J Psychiatry* 167:1349–1356.

Corominas R et al. (2014) Protein interaction network of alternatively spliced isoforms from brain links genetic risk factors for autism. *Nat Commun* 5:1–12.

Cortelazzo A, Guerranti R, Felice C De, Signorini C, Leoncini S, Pecorelli A, Landi C, Bini L, Montomoli B, Sticozzi C, Ciccoli L, Valacchi G, Hayek J (2013) A Plasma Proteomic Approach in Rett Syndrome : Classical versus Preserved Speech Variant. 2013.

Courchesne E, Karns CM, Davis HR, Ziccardi R, Carper RA, Tigue ZD, Chisum HJ, Moses P, Pierce K, Lord C, Lincoln AJ, Pizzo S, Schreibman L, Haas RH, Akshoomoff NA, Courchesne RY (2001) Unusual brain growth patterns in early life in patients with autistic disorder: an MRI study. *Neurology* 57:245–254.

Cox J, Mann M (2008) MaxQuant enables high peptide identification rates,

- individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 26:1367–1372.
- Crawley JN (2007) Mouse behavioral assays relevant to the symptoms of autism. *Brain Pathol* 17:448–459.
- Creighton BA, Afriyie S, Ajit D, Casingal CR, Voos KM, Reger J, Burch AM, Dyne E, Bay J, Huang JK, Anton ES, Fu MM, Lorenzo DN (2021) Giant ankyrin-B mediates transduction of axon guidance and collateral branch pruning factor sema 3A. *Elife* 10.
- Cummins JM, Vogelstein B (2004) HAUSP is required for p53 destabilization. *Cell Cycle* 3:687–690.
- Da Costa RM, Karmirian K, Rehen SK (2018) Deformation of mitochondrial cristae in human neural progenitor cells exposed to valproic acid. *An Acad Bras Cienc* 90:2223–2232.
- Daly MJ, Patterson N, Mesirov JP, Golub TR, Tamayo P, Spiegelman B (2003) PGC-1 α -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* 34:267–273.
- Darnell JC et al. (2011) FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell* 146:247–261.
- Das AT, Tenenbaum L, Berkhout B (2016) Tet-On Systems For Doxycycline-inducible Gene Expression. *Curr Gene Ther* 16:156–167.
- de Ligt J, Willemsen MH, van Bon BWM, Kleefstra T, Yntema HG, Kroes T, Vulto-van Silfhout AT, Koolen DA, de Vries P, Gilissen C, del Rosario M, Hoischen A,

- Scheffer H, de Vries BBA, Brunner HG, Veltman JA, Vissers LELM (2012) Diagnostic Exome Sequencing in Persons with Severe Intellectual Disability. *N Engl J Med* 367:1921–1929.
- De Rubeis S et al. (2014) Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature* 515:209–215.
- Deciphering T, Disorders D, Dnm PV, Dd M, Ddd DD, Disorders DD, Disease D, Disability I, Faces A (2016) Prevalence , phenotype and architecture of developmental disorders caused by de novo mutation Key Words Main text. :1–39.
- Delhaye S, Bardoni B (2021) Role of phosphodiesterases in the pathophysiology of neurodevelopmental disorders. *Mol Psychiatry* 26:4570–4582.
- Delpire E (2009) The mammalian family of sterile 20p-like protein kinases. *Pflugers Arch* 458:953–967.
- Deneault E et al. (2018) Complete Disruption of Autism-Susceptibility Genes by Gene Editing Predominantly Reduces Functional Connectivity of Isogenic Human Neurons. *Stem Cell Reports* 11:1211–1225.
- Deneault E, Faheem M, White SH, Rodrigues DC, Sun S, Wei W, Piekna A, Thompson T, Howe JL, Chalil L, Kwan V, Walker S, Pasceri P, Roth FP, Yuen RK, Singh KK, Ellis J, Scherer SW (2019) CNTN5-/+or EHMT2-/+human iPSC-derived neurons from individuals with autism develop hyperactive neuronal networks. *Elife* 8:1–26.
- Deshpande A, Yadav S, Dao DQ, Wu Z-Y, Hokanson KC, Cahill MK, Wiita AP, Jan Y-N, Ullian EM, Weiss LA (2017) Cellular Phenotypes in Human iPSC-Derived Neurons from a Genetic Model of Autism Spectrum Disorder. *Cell Rep* 21:2678–

2687.

Devine MJ, Kittler JT (2018) Mitochondria at the neuronal. *Nat Publ Gr* 19:63–80.

Dimos JT, Rodolfa KT, Niakan KK, Weisenthal LM, Mitumoto H, Chung W, Croft GF,

Saphier G, Leibel R, Goland R, Wichterle H, Henderson CE, Eggan K (2008)

Induced pluripotent stem cells generated from patients with ALS can be

differentiated into motor neurons. *Science* (80-) 321:1218–1221.

Divakaruni SS, Dyke AM Van, Chandra R, Lobo MK, Thompson SM, Blanpied TA,

Divakaruni SS, Dyke AM Van, Chandra R, Legates TA, Contreras M (2018) Long-

Term Potentiation Requires a Rapid Burst of Dendritic Mitochondrial Fission during

Induction. *Neuron* 100:1–16.

Djuric U, Cheung AYL, Zhang W, Mok RS, Lai W, Piekna A, Hendry JA, Ross PJ,

Pasceri P, Kim D, Salter MW, Ellis J (2015) Neurobiology of Disease MECP2e1

isoform mutation affects the form and function of neurons derived from Rett

syndrome patient iPS cells. *Neurobiol Dis* 76:37–45.

Doers ME, Musser MT, Nichol R, Berndt ER, Baker M, Gomez TM, Zhang S-C,

Abbeduto L, Bhattacharyya A (2014) iPSC-Derived Forebrain Neurons from FXS

Individuals Show Defects in Initial Neurite Outgrowth. *Stem Cells Dev* 23:1777–

1787.

Douvaras P, Sun B, Wang M, Kruglikov I, Lalloos G, Zimmer M, Terrenoire C, Zhang B,

Gandy S, Schadt E, Freytes DO, Noggle S, Fossati V (2017) Directed

Differentiation of Human Pluripotent Stem Cells to Microglia. *Stem Cell Reports*

8:1516–1524.

- Douvaras P, Wang J, Zimmer M, Hanchuk S, O'Bara MA, Sadiq S, Sim FJ, Goldman J, Fossati V (2014) Efficient generation of myelinating oligodendrocytes from primary progressive multiple sclerosis patients by induced pluripotent stem cells. *Stem Cell Reports* 3:250–259.
- Doyle CA, McDougle CJ (2012) Pharmacologic treatments for the behavioral symptoms associated with autism spectrum disorders across the lifespan. *Dialogues Clin Neurosci* 14:263–279.
- Durak O, Gao F, Kaeser-Woo YJ, Rueda R, Martorell AJ, Nott A, Liu CY, Watson LA, Tsai L-HH (2016) Chd8 mediates cortical neurogenesis via transcriptional regulation of cell cycle and Wnt signaling. *Nat Neurosci* 19:1–15.
- Ebrahimi-fakhari D, Saffari A, Wahlster L, Nardo A Di, Turner D, Lewis TL, Hoffmann GF, Han MJ, Di Nardo A, Turner D, Lewis TL, Conrad C, Rothberg JM, Lipton JO, Kölker S, Hoffmann GF, Han MJ, Polleux F, Sahin M (2016) Impaired Mitochondrial Dynamics and Mitophagy in Neuronal Models of Tuberous Sclerosis Complex. *Cell Rep* 17:1053–1070.
- Emerson RW et al. (2017) Functional neuroimaging of high-risk 6-month-old infants predicts a diagnosis of autism at 24 months of age. *Sci Transl Med* 9.
- Fazal FM, Han S, Parker KR, Kaewsapsak P, Xu J, Boettiger AN, Chang HY, Ting AY (2019) Atlas of Subcellular RNA Localization Revealed by APEX-Seq. *Cell* 178:1–18.
- Feliciano P, Zhou X, Astrovskaya I, Turner TN, Muzny DM, Sabo A, Consortium TS, Gibbs RA, Evan E (2019) Exome sequencing of 457 autism families recruited online

provides evidence for novel ASD genes. *Genomic Med*

4:<https://doi.org/10.1038/s41525-019-0093-8>.

Fendt SM, Verstreken P (2017) Neurons eat glutamate to stay alive. *J Cell Biol* 216:863–865.

Fernandez A, Meechan DW, Karpinski BA, Rothblat LA, Maynard TM, Lamantia A, Fernandez A, Meechan DW, Karpinski BA, Paronett EM, Bryan CA (2019) Mitochondrial Dysfunction Leads to Cortical Under-Connectivity and Cognitive Impairment 22q11. *Neuron* 102:1127-1142.e3.

Fernández E et al. (2009) Targeted tandem affinity purification of PSD-95 recovers core postsynaptic complexes and schizophrenia susceptibility proteins. *Mol Syst Biol* 5.

Filamin-a I, Remodeling F, Vliet AR Van, Giordano F, Gerlo S, Samali A, Mizuno H, Agostinis P, Vliet AR Van, Giordano F, Gerlo S, Segura I, Eygen S Van (2017) The ER Stress Sensor PERK Coordinates ER-Plasma Membrane Contact Site Formation through Article The ER Stress Sensor PERK Coordinates ER-Plasma Membrane Contact Site Formation through Interaction with Filamin-A and F-Actin Remodeling. *Mol Cell* 65:885-899.e6.

Flaherty E et al. (2019) Neuronal impact of patient-specific aberrant NRXN1 α splicing. Springer US.

Folstein S, Rutter M (1977) Genetic influences and infantile autism. *Nature* 265:726–728.

Frye RE (2020) Mitochondrial Dysfunction in Autism Spectrum Disorder: Unique Abnormalities and Targeted Treatments. *Semin Pediatr Neurol* 11:1071–9091.

Frye RE, Delatorre R, Taylor H, Slattery J, Melnyk S, Chowdhury N, James SJ (2013)

Redox metabolism abnormalities in autistic children associated with mitochondrial disease. *Transl Psychiatry* 3:1–12.

Frye RE, Rossignol DA (2011) Mitochondrial dysfunction can connect the diverse medical symptoms associated with autism spectrum disorders. *Pediatr Res* 69:41–47.

Furlong RA, Narain Y, Rankin J, Wytttenbach A, Rubinsztein DC (2000) α -Synuclein overexpression promotes aggregation of mutant huntingtin. *Biochem J* 346:577–581.

Gao R, Penzes P (2015) Common mechanisms of excitatory and inhibitory imbalance in schizophrenia and autism spectrum disorders. *Curr Mol Med* 15:146–167.

Garcia-Forn M, Boitnott A, Akpinar Z, Rubeis S De (2020) Linking Autism Risk Genes to Disruption of Cortical Development. *Cells* 9:2500.

Gaugler T, Klei L, Sanders SJ, Bodea C a, 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. *Nat Genet* 46:881–885.

Gazestani VH, Pramparo T, Nalabolu S, Kellman BP, Murray S, Lopez L, Pierce K, Courchesne E, Lewis NE (2019) A perturbed gene network containing PI3K-AKT, RAS-ERK and WNT-B-catenin pathways in leukocytes is linked to ASD genetics and symptom severity. *Nat Neurosci* 22.

Gdalyahu A, Lazaro M, Penagarikano O, Golshani P, Trachtenberg JT, Gescwind DH (2015) The autism related protein contactin-associated protein-like 2 (CNTNAP2) stabilizes new spines: An in vivo mouse study. *PLoS One* 10:1–7.

Geisheker MR et al. (2017) Hotspots of missense mutation identify neurodevelopmental disorder genes and functional domains. *Nat Neurosci* 20:1043–1051.

Geschwind DH, Levitt P (2007) Autism spectrum disorders: developmental disconnection syndromes. *Curr Opin Neurobiol* 17:103–111.

Ghanizadeh A, Sahraeizadeh A, Berk M (2014) A head-to-head comparison of aripiprazole and risperidone for safety and treating autistic disorders, a randomized double blind clinical trial. *Child Psychiatry Hum Dev* 45:185–192.

Girirajan S et al. (2010) A recurrent 16p12.1 microdeletion supports a two-hit model for severe developmental delay. *Nat Genet* 42:203–209.

Glessner JT et al. (2009) Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature* 459:569–573.

Go CD et al. (2021) A proximity-dependent biotinylation map of a human cell. *Nature* 595.

Goodspeed K et al. (2020) Current knowledge of SLC6A1-related neurodevelopmental disorders. *Brain Commun* 2:1–10.

Graef JD, Wu H, Ng C, Sun C, Villegas V, Qadir D, Jesseman K, Warren ST, Jaenisch R, Cacace A, Wallace O (2020) Partial FMRP expression is sufficient to normalize neuronal hyperactivity in Fragile X neurons. *Eur J Neurosci* 51:2143–2157.

Griesi-Oliveira K et al. (2021) Transcriptome of iPSC-derived neuronal cells reveals a module of co-expressed genes consistently associated with autism spectrum disorder. *Mol Psychiatry* 26:1589–1605.

Grove J et al. (2019) Identification of common genetic risk variants for autism spectrum

disorder. *Nat Genet.*

Guard SE, Poss ZC, Ebmeier CC, Pagratis M, Simpson H, Taatjes DJ, Old WM (2019)

The nuclear interactome of DYRK1A reveals a functional role in DNA damage repair. *Sci Rep* 9:1–12.

Guo W, Chen N, Chen Y, Xia Q, Shen Y (2005) Activation of mitochondrial ATP-sensitive potassium channel contributes to protective effect in prolonged myocardial preservation. *Annu Int Conf IEEE Eng Med Biol - Proc 7 VOLS:4027–4030.*

Hamdan H, Lim BC, Torii T, Joshi A, Konning M, Smith C, Palmer DJ, Ng P, Leterrier C, Oses-Prieto JA, Burlingame AL, Rasband MN (2020) Mapping axon initial segment structure and function by multiplexed proximity biotinylation. *Nat Commun* 11.

Harfterkamp MA, van de Loo-Neus G, Minderaa RB (2012) Atomoxetine reduces ADHD symptoms in children with autism spectrum disorder. *Evid Based Ment Health* 15:96.

Harris JJ, Attwell D (2012) The energetics of CNS white matter. *J Neurosci* 32:356–371.

Hart T, Chandrashekhar M, Aregger M, Steinhart Z, Brown KR, MacLeod G, Mis M, Zimmermann M, Fradet-Turcotte A, Sun S, Mero P, Dirks P, Sidhu S, Roth FP, Rissland OS, Durocher D, Angers S, Moffat J (2015) High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. *Cell* 163:1515–1526.

Hazlett HC et al. (2017) Early brain development in infants at high risk for autism spectrum disorder. *Nat Publ Gr* 542:348–351.

- He Y, Zhou Y, Ma W, Wang J (2019) An integrated transcriptomic analysis of autism spectrum disorder. *Sci Rep* 9:1–9.
- Hedstrom KL, Ogawa Y, Rasband MN (2008) AnkyrinG is required for maintenance of the axon initial segment and neuronal polarity. *J Cell Biol* 183:635–640.
- Heinemann-Yerushalmi L, Bentovim L, Felsenthal N, Vinestock RC, Michaeli N, Krief S, Silberman A, Cohen M, Ben-Dor S, Brenner O, Haffner-Krausz R, Itkin M, Malitsky S, Erez A, Zelzer E (2021) BCKDK regulates the TCA cycle through PDC in the absence of PDK family during embryonic development. *Dev Cell* 56:1182-1194.e6.
- Hendaus MA, Jomha FA, Alhammadi AH (2019) Vasopressin in the amelioration of social functioning in autism spectrum disorder. *J Clin Med* 8.
- Hering H, Sheng M (2001) Dendritic spines: structure, dynamics and regulation. *Nat Rev Neurosci* 2:880–888.
- Herringshaw AJ, Ammons CJ, DeRamus TP, Kana RK (2016) Hemispheric differences in language processing in autism spectrum disorders: A meta-analysis of neuroimaging studies. *Autism Res* 9:1046–1057.
- Hiday AC, Edler MC, Salek AB, Morris CW, Thang M, Rentz TJ, Rose KL, Jones LM, Baucum AJ (2017) Mechanisms and Consequences of Dopamine Depletion-Induced Attenuation of the Spinophilin/Neurofilament Medium Interaction. *Neural Plast* 2017.
- Hirata K, Nambara T, Kawatani K, Nawa N, Yoshimatsu H, Kusakabe H, Banno K, Nishimura K, Ohtaka M, Nakanishi M, Taniguchi H, Arahori H, Wada K, Ozono K,

- Kitabatake Y (2020) 4-Phenylbutyrate ameliorates apoptotic neural cell death in Down syndrome by reducing protein aggregates. *Sci Rep* *10*:1–14.
- Hockemeyer D, Soldner F, Cook EG, Gao Q, Mitalipova M, Jaenisch R (2008) A Drug-Inducible System for Direct Reprogramming of Human Somatic Cells to Pluripotency. *Cell Stem Cell* *3*:346–353.
- Hotulainen P, Hoogenraad CC (2010) Actin in dendritic spines: connecting dynamics to function. *J Cell Biol* *189*:619–629.
- Houge G et al. (2015) B56 δ -related protein phosphatase 2A dysfunction identified in patients with intellectual disability. *J Clin Invest* *125*:3051–3062.
- Howarth M, Takao K, Hayashi Y, Ting AY (2005) Targeting quantum dots to surface proteins in living cells with biotin ligase. *Proc Natl Acad Sci U S A* *102*:7583–7588.
- Howlin P, Goode S, Hutton J, Rutter M (2004) Adult outcome for children with autism. *J Child Psychol Psychiatry Allied Discip* *45*:212–229.
- Hu C, Feng P, Yang Q, Xiao L (2021) Clinical and Neurobiological Aspects of TAO Kinase Family in Neurodevelopmental Disorders. *Front Mol Neurosci* *14*:1–15.
- Hulme AJ, Maksour S, St-Clair Glover M, Mielle S, Dottori M (2022) Making neurons, made easy: The use of Neurogenin-2 in neuronal differentiation. *Stem Cell Reports* *17*:14–34.
- Hung V, Lam SS, Udeshi ND, Svinkina T, Guzman G, Mootha VK, Carr SA, Ting AY (2017) Proteomic mapping of cytosol-facing outer mitochondrial and ER membranes in living human cells by proximity biotinylation. *Elife* *6*:e24463.
- Hutchison M, Berman KS, Cobb MH (1998) Isolation of TAO1, a Protein Kinase That

- Activates MEKs in Stress-activated Protein Kinase Cascades. *J Biol Chem* 273:28625–28632.
- Hutsler JJ, Love T, Zhang H (2007) Histological and Magnetic Resonance Imaging Assessment of Cortical Layering and Thickness in Autism Spectrum Disorders. *Biol Psychiatry* 61:449–457.
- In D, Jensen SC, Noble KA, Kc B, Roux KH (2016) An improved smaller biotin ligase for BioID proximity labeling. *Mol Biol Cell*.
- Iossifov I et al. (2012) De Novo Gene Disruptions in Children on the Autistic Spectrum. *Neuron* 74:285–299.
- Iossifov I et al. (2014) The contribution of de novo coding mutations to autism spectrum disorder. *Nature* 515:216–221.
- Iwata R, Casimir P, Vanderhaeghen P (2020) Mitochondrial dynamics in postmitotic cells regulate neurogenesis. *Science* (80-) 369:858–862.
- Iyer J, Singh MD, Jensen M, Patel P, Pizzo L, Huber E, Koerselman H, Weiner AT, Lepanto P, Vadodaria K, Kubina A, Wang Q, Talbert A, Yennawar S, Badano J, Manak JR, Rolls MM, Krishnan A, Girirajan S (2018) Pervasive genetic interactions modulate neurodevelopmental defects of the autism-associated 16p11.2 deletion in *Drosophila melanogaster*. *Nat Commun* 9:2548.
- Jagtap S, Thanos JM, Fu T, Wang J, Lalonde J, Dial TO, Feiglin A, Chen J, Kohane I, Lee JT, Sheridan SD, Perlis RH (2019) Aberrant mitochondrial function in patient-derived neural cells from CDKL5 deficiency disorder and Rett syndrome. *Hum Mol Genet* 28:3625–3636.

- Jiang Y-H et al. (2013) Detection of Clinically Relevant Genetic Variants in Autism Spectrum Disorder by Whole-Genome Sequencing. *Ajhg* 93:249–263.
- Jin X, Simmons SK, Guo AX, Shetty AS, Ko M, Nguyen L, Jokhi V, Robinson E, Oyler P, Curry N, Deangeli G, Lodato S, Levin JZ, Regev A, Zhang F, Arlotta P (2020) In vivo Perturb-Seq reveals neuronal and glial abnormalities associated with autism risk genes. *Science* (80-) 370:eaaz6063.
- Johannessen CM et al. (2010) COT drives resistance to RAF inhibition through MAP kinase pathway reactivation. *Nature* 468:968–972.
- John GB, Shang Y, Li L, Renken C, Mannella CA, Selker JML, Rangell L, Bennett MJ, Zha J (2005) The Mitochondrial Inner Membrane Protein Mitofilin Controls Cristae Morphology. *Mol Biol Cell* 16:1543–1554.
- Journiac N et al. (2020) Cell Metabolic Alterations due to Mcph1 Mutation in Microcephaly. *Cell Rep* 31.
- Jovičić A, Mertens J, Boeynaems S, Bogaert E, Chai N, Yamada SB, Paul JW, Sun S, Herdy JR, Bieri G, Kramer NJ, Gage FH, Van Den Bosch L, Robberecht W, Gitler AD (2015) Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. *Nat Neurosci* 18:1226–1229.
- Kanellopoulos AK, Mariano V, Spinazzi M, Fiala A, Achsel T, Kanellopoulos AK, Mariano V, Spinazzi M, Woo YJ, Mclean C, Pech U (2020) Aralar Sequesters GABA into Hyperactive Article Aralar Sequesters GABA into Hyperactive Mitochondria , Causing Social Behavior Deficits. *Cell* 180:1178-1197.e20.
- Kapfhamer D, Taylor S, Zou ME, Lim JP, Kharazia V, Heberlein U (2013) *Taok2*

controls behavioral response to ethanol in mice. *Genes, Brain Behav* 12:87–97.

Kelley KW, Paşca SP (2022) Human brain organogenesis: Toward a cellular understanding of development and disease. *Cell* 185:42–61.

Khacho M, Harris R, Slack RS (2019) Mitochondria as central regulators of neural stem cell fate and cognitive function. *Nat Rev Neurosci* 20:34–48.

Khan TA et al. (2020) Neuronal defects in a human cellular model of 22q11.2 deletion syndrome.

Khattak S et al. (2015) Human induced pluripotent stem cell derived neurons as a model for Williams-Beuren syndrome. *Mol Brain* 8:77.

Kim BR, Coyaud E, Laurent EMN, St-germain J, Laar E Van De, Tsao M, Raught B, Moghal N (2017) Identification of the SOX2 Interactome by BioID Reveals EP300 as a Mediator of SOX2-dependent Squamous Differentiation and Lung Squamous Cell Carcinoma Growth * □. :1864–1888.

Kim E et al. (2016) Systematic functional interrogation of rare cancer variants identifies oncogenic alleles. *Cancer Discov* 6:714–726.

Kishino T, Lalande M, Wagstaff J (1997) UBE3A/E6-AP mutations cause Angelman syndrome. *Nat Genet* 15:70–73.

Koire A, Katsonis P, Kim YW, Buchovecky C, Wilson SJ, Lichtarge O (2021) A method to delineate de novo missense variants across pathways prioritizes genes linked to autism. *Sci Transl Med* 13:eabc1739.

Kosmicki JA, Samocha KE, Howrigan DP, Sanders SJ, Slowikowski K, Lek M, Karczewski KJ, Cutler DJ, Devlin B, Roeder K, Buxbaum JD, Neale BM,

- MacArthur DG, Wall DP, Robinson EB, Daly MJ (2017) Refining the role of de novo protein truncating variants in neurodevelopmental disorders using population reference samples. *Nat Genet* 49:052886.
- Krencik R, Seo K, van Asperen J V., Basu N, Cvetkovic C, Barlas S, Chen R, Ludwig C, Wang C, Ward ME, Gan L, Horner PJ, Rowitch DH, Ullian EM (2017) Systematic Three-Dimensional Coculture Rapidly Recapitulates Interactions between Human Neurons and Astrocytes. *Stem Cell Reports* 9:1745–1753.
- Krey JF, Paşca SP, Shcheglovitov A, Yazawa M, Schwemberger R, Rasmuson R, Dolmetsch RE (2013) Timothy syndrome is associated with activity-dependent dendritic retraction in rodent and human neurons. *Nat Neurosci* 16:201–209.
- Kucera M, Isserlin R, Arkhangorodsky A, Bader GD (2016) AutoAnnotate: A Cytoscape app for summarizing networks with semantic annotations [version 1; referees: 2 approved]. *F1000Research* 5:1–12.
- Kurochkin I, Khrameeva E, Tkachev A, Stepanova V, Vanyushkina A, Stekolshchikova E, Li Q, Zubkov D, Shichkova P, Halene T, Willmitzer L, Giavalisco P, Akbarian S, Khaitovich P (2019) Metabolome signature of autism in the human prefrontal cortex. *Commun Biol* 2:1–10.
- Kuzmanov U, Emili A (2013) Protein-protein interaction networks: Probing disease mechanisms using model systems. *Genome Med* 5:1–12.
- Kwan V, Meka DP, White SH, Hung CL, Holzapfel NT, Walker S, Murtaza N, Unda BK, Schwanke B, Yuen RKCC, Habing K, Milsom C, Hope KJ, Truant R, Scherer SW, Calderon de Anda F, Singh KK (2016) DIXDC1 Phosphorylation and Control of

- Dendritic Morphology Are Impaired by Rare Genetic Variants. *Cell Rep* 17:1892–1904.
- Kwan V, Rosa E, Xing S, Murtaza N, Singh K, Holzapfel NT, Berg T, Lu Y, Singh KK (2021) Proteomic Analysis Reveals Autism-Associated Gene DIXDC1 Regulates Proteins Associated with Mitochondrial Organization and Function. *J Proteome Res* 20:1052–1062.
- Lage K (2014) Protein-protein interactions and genetic diseases: The interactome. *Biochim Biophys Acta - Mol Basis Dis* 1842:1971–1980.
- Lai M-C, Lombardo M V, Baron-Cohen S (2014) Autism. *Lancet* 383:896–910.
- Lancaster MA, Renner M, Martin C, Wenzel D, Bicknell LS, Hurles ME, Homfray T, Penninger JM, Jackson AP, Knoblich JA (2013) Cerebral organoids model human brain development and microcephaly. *Nature* 501:373–379.
- Lauren A. Weiss et al. (2008) Association between microdeletion and microduplication at 16p11.2 and autism. *N Engl J Med* 358:667–675.
- Lautz JD, Tsegay KB, Zhu Z, Gniffke EP, Welsh JP, Smith SEP (2021) Synaptic protein interaction networks encode experience by assuming stimulus-specific and brain-region-specific states. *Cell Rep* 37:110076.
- Lek M et al. (2016) Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536:285–291.
- Lennox AL et al. (2020) Pathogenic DDX3X Mutations Impair RNA Metabolism and Neurogenesis during Fetal Cortical Development. *Neuron* 106:404-420.e8.
- Leterrier C (2018) The axon initial segment: An updated viewpoint. *J Neurosci* 38:2135–

2145.

Leussis MP, Berry-Scott EM, Saito M, Jhuang H, de Haan G, Alkan O, Luce CJ,

Madison JM, Sklar P, Serre T, Root DE, Petryshen TL (2013) The

ANK3 Bipolar Disorder Gene Regulates Psychiatric-Related Behaviors

That Are Modulated by Lithium and Stress. *Biol Psychiatry* 73:683–690.

Levy M, Faas GC, Saggau P, Craigen WJ, Sweatt JD (2003) Mitochondrial regulation of

synaptic plasticity in the hippocampus. *J Biol Chem* 278:17727–17734.

Lewis DA (2002) The human brain revisited: Opportunities and challenges in

postmortem studies of psychiatric disorders. *Neuropsychopharmacology* 26:143–

154.

Li J et al. (2015) Identification of Human Neuronal Protein Complexes Reveals

Biochemical Activities and Convergent Mechanisms of Action in Autism Spectrum

Disorders. *Cell Syst* 1:361–374.

Li J, Ryan SK, Deboer E, Cook K, Fitzgerald S, Lachman HM, Wallace DC, Goldberg

EM, Anderson SA (2019) Mitochondrial deficits in human iPSC-derived neurons

from patients with 22q11.2 deletion syndrome and schizophrenia. *Transl Psychiatry*

9.

Li P, Li J, Wang L, Di LJ (2017) Proximity Labeling of Interacting Proteins: Application

of BioID as a Discovery Tool. *Proteomics* 17:1–10.

Li Z, Okamoto KI, Hayashi Y, Sheng M (2004) The importance of dendritic

mitochondria in the morphogenesis and plasticity of spines and synapses. *Cell*

119:873–887.

Licznerski P et al. (2020) ATP Synthase c-Subunit Leak Causes Aberrant Cellular

Metabolism in Fragile X Syndrome. *Cell* 182:1170-1185.e9.

Lin GN, Corominas R, Lemmens I, Yang X, Tavernier J, Hill DE, Vidal M, Sebat J,

Iakoucheva LM (2015) Spatiotemporal 16p11.2 Protein Network Implicates Cortical

Late Mid-Fetal Brain Development and KCTD13-Cul3-RhoA Pathway in

Psychiatric Diseases. *Neuron* 85:742–754.

Lin YT et al. (2018) APOE4 Causes Widespread Molecular and Cellular Alterations

Associated with Alzheimer’s Disease Phenotypes in Human iPSC-Derived Brain

Cell Types. *Neuron* 98:1141-1154.e7.

Lindhurst MJ, Fiermonte G, Song S, Struys E, De Leonardis F, Schwartzberg PL, Chen

A, Castegna A, Verhoeven N, Mathews CK, Palmieri F, Biesecker LG (2006)

Knockout of Slc25a19 causes mitochondrial thiamine pyrophosphate depletion,

embryonic lethality, CNS malformations, and anemia. *Proc Natl Acad Sci U S A*

103:15927–15932.

Liu X, Salokas K, Tamene F, Jiu Y, Öhman T, Varjosalo M, Weldatsadik RG, Öhman T,

Varjosalo M (2018) An AP-MS- and BioID-compatible MAC-tag enables

comprehensive mapping of protein interactions and subcellular localizations. *Nat*

Commun 9.

Liu Y, Liu H, Sauvey C, Yao L, Zarnowska ED, Zhang SC (2013) Directed

differentiation of forebrain GABA interneurons from human pluripotent stem cells.

Nat Protoc 8:1670–1679.

Loh KH, Stawski PS, Draycott AS, Udeshi ND, Lehrman EK, Wilton DK, Svinkina T,

- Deerinck TJ, Ellisman MH, Stevens B, Carr SA, Ting AY (2016) Proteomic Analysis of Unbounded Cellular Compartments : Synaptic Clefts Resource. *Proteomic Analysis of Unbounded Cellular Compartments : Synaptic Clefts. Cell* 166:1295-1307.e21.
- Lord C, Brugha TS, Charman T, Cusack J, Dumas G, Frazier T, Jones EJH, Jones RM, Pickles A, State MW, Taylor JL, Veenstra-VanderWeele J (2020) Autism spectrum disorder. *Nat Rev Dis Prim* 6.
- Lord C, Elsabbagh M, Baird G, Veenstra-Vanderweele J (2018) Autism spectrum disorder. *Lancet* 392:508–520.
- Lowe M, Gonatas NK, Warren G (2000) The mitotic phosphorylation cycle of the cis-Golgi matrix protein GM130. *J Cell Biol* 149:341–356.
- Lu G, Middleton RE, Sun H, Naniong MV, Ott CJ, Mitsiades CS, Wong KK, Bradner JE, Kaelin WG (2014) The myeloma drug lenalidomide promotes the cereblon-dependent destruction of ikaros proteins. *Science (80-)* 343:305–309.
- Lu J, Zhong X, Liu H, Hao L, Huang CTL, Sherafat MA, Jones J, Ayala M, Li L, Zhang SC (2016) Generation of serotonin neurons from human pluripotent stem cells. *Nat Biotechnol* 34:89–94.
- Lutz AK et al. (2020) Autism-associated SHANK3 mutations impair maturation of neuromuscular junctions and striated muscles. *Sci Transl Med* 12.
- Madison JM et al. (2021) Regulation of purine metabolism connects KCTD13 to a metabolic disorder with autistic features. *iScience* 24:101935.
- Maenner MJ et al. (2021) Prevalence and Characteristics of Autism Spectrum Disorder

- Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2018. *MMWR Surveill Summ* 70:1–16.
- Malecki J, Ho AYY, Moen A, Dahl HA, Falnes P (2015) Human METTL20 is a mitochondrial lysine methyltransferase that targets the β subunit of electron transfer flavoprotein (ETF β) and modulates its activity. *J Biol Chem* 290:423–434.
- Mao P, Joshi K, Li J, Kim SH, Li P, Santana-Santos L, Luthra S, Chandran UR, Benos P V., Smith L, Wang M, Hu B, Cheng SY, Sobol RW, Nakano I (2013) Mesenchymal glioma stem cells are maintained by activated glycolytic metabolism involving aldehyde dehydrogenase 1A3. *Proc Natl Acad Sci U S A* 110:8644–8649.
- Marchetto MCN, Carromeu C, Acab A, Yu D, Yeo GW, Mu Y, Chen G, Gage FH, Muotri AR (2010) A Model for Neural Development and Treatment of Rett Syndrome Using Human Induced Pluripotent Stem Cells. *Cell* 143:527–539.
- Marcogliese PC et al. (2022) Drosophila functional screening of de novo variants in autism uncovers damaging variants and facilitates discovery of rare neurodevelopmental diseases. *Cell Rep* 38.
- Marini C, Scheffer IE, Nabbout R, Suls A, De Jonghe P, Zara F, Guerrini R (2011) The genetics of Dravet syndrome. *Epilepsia* 52:24–29.
- Markmiller S et al. (2018) Context-Dependent and Disease-Specific Diversity in Protein Interactions within Stress Granules. *Cell* 172:590–604.
- Matsumura K et al. (2020) Pathogenic POGZ mutation causes impaired cortical development and reversible autism-like phenotypes. *Nat Commun* 11:1–16.

- May DG, Scott KL, Campos AR, Roux KJ (2020) Comparative Application of BioID and TurboID for Protein-Proximity Biotinylation. *Cells* 9.
- Mefford HC et al. (2008) Recurrent Rearrangements of Chromosome 1q21.1 and Variable Pediatric Phenotypes. *N Engl J Med* 359:1685–1699.
- Meili F, Wei WJ, Sin WC, Meyers WM, Dascalu I, Callaghan DB, Rogic S, Pavlidis P, Haas K (2021) Multi-parametric analysis of 57 SYNGAP1 variants reveal impacts on GTPase signaling, localization, and protein stability. *Am J Hum Genet* 108:148–162.
- Menzies C, Naz S, Patten D, Lacoste B, Alquier T, Bennett BM (2021) Distinct basal metabolism in three mouse models of neurodevelopmental disorders. *eNeuro* 8:1–14.
- Merico D, Isserlin R, Stueker O, Emili A, Bader GD (2010) Enrichment map: A network-based method for gene-set enrichment visualization and interpretation. *PLoS One* 5:e13984.
- Mertens J et al. (2015) Differential responses to lithium in hyperexcitable neurons from patients with bipolar disorder. *Nature* 527:95–99.
- Meyer K, Feldman HM, Lu T, Drake D, Lim ET, Ling KH, Bishop NA, Pan Y, Seo J, Lin YT, Su SC, Church GM, Tsai LH, Yankner BA (2019) REST and Neural Gene Network Dysregulation in iPSC Models of Alzheimer’s Disease. *Cell Rep* 26:1112–1127.e9.
- Miles JH (2011) Autism spectrum disorders-A genetics review. *Genet Med* 13:278–294.
- Minde DP, Ramakrishna M, Lilley KS (2020) Biotin proximity tagging favours unfolded

proteins and enables the study of intrinsically disordered regions. *Commun Biol* 3:1–13.

Mitsopoulos C, Zihni C, Garg R, Ridley AJ, Morris JDH (2003) The Prostate-derived Sterile 20-like Kinase (PSK) Regulates Microtubule Organization and Stability. *J Biol Chem* 278:18085–18091.

Modabbernia A, Velthorst E, Reichenberg A (2017) Environmental risk factors for autism: an evidence-based review of systematic reviews and meta-analyses. *Mol Autism* 8:1–16.

Moore TM, Garg R, Johnson C, Coptcoat MJ, Ridley AJ, Morris JDH (2000) PSK, a Novel STE20-like Kinase Derived from Prostatic Carcinoma That Activates the c-Jun N-terminal Kinase Mitogen-activated Protein Kinase Pathway and Regulates Actin Cytoskeletal Organization. *J Biol Chem* 275:4311–4322.

Muffat J, Li Y, Yuan B, Mitalipova M, Omer A, Corcoran S, Bakiasi G, Tsai L-HH, Aubourg P, Ransohoff RM, Jaenisch R (2016) Efficient derivation of microglia-like cells from human pluripotent stem cells. *Nat Med* 22:1358–1367.

Murtaza N et al. (2022) Neuron-specific protein network mapping of autism risk genes identifies shared biological mechanisms and disease relevant pathologies. *BioRxiv*.

Murtaza N, Uy J, Singh KK (2020) Emerging proteomic approaches to identify the underlying pathophysiology of neurodevelopmental and neurodegenerative disorders. *Mol Autism* 11:27.

Nageshappa S et al. (2015) Altered neuronal network and rescue in a human MECP2 duplication model. *Mol Psychiatry* 21:1–11.

- Nalivaeva NN, Turner AJ, Zhuravin IA (2018) Role of prenatal hypoxia in brain development, cognitive functions, and neurodegeneration. *Front Neurosci* *12*:1–21.
- Namba T, Dóczy J, Pinson A, Xing L, Kalebic N, Wilsch-Bräuning M, Long KR, Vaid S, Lauer J, Bogdanova A, Borgonovo B, Shevchenko A, Keller P, Drechsel D, Kurzchalia T, Wimberger P, Chinopoulos C, Huttner WB (2020) Human-Specific ARHGAP11B Acts in Mitochondria to Expand Neocortical Progenitors by Glutaminolysis. *Neuron* *105*:867–881.e9.
- Namba T, Nardelli J, Gressens P, Huttner WB (2021) Perspective Metabolic Regulation of Neocortical Expansion in Development and Evolution. *Neuron* *109*:408–419.
- Nathanson JL, Yanagawa Y, Obata K, Callaway EM (2009) Preferential labeling of inhibitory and excitatory cortical neurons by endogenous tropism of adeno-associated virus and lentivirus vectors. *NSC* *161*:441–450.
- Neal EG, Chaffe H, Schwartz RH, Lawson MS, Edwards N, Fitzsimmons G, Whitney A, Cross JH (2008) The ketogenic diet for the treatment of childhood epilepsy: a randomised controlled trial. *Lancet Neurol* *7*:500–506.
- Neal EG, Chaffe H, Schwartz RH, Lawson MS, Edwards N, Fitzsimmons G, Whitney A, Cross JH (2009) A randomized trial of classical and medium-chain triglyceride ketogenic diets in the treatment of childhood epilepsy. *Epilepsia* *50*:1109–1117.
- Neale BM et al. (2012) Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* *485*:242–245.
- Nehme R et al. (2018) Combining NGN2 Programming with Developmental Patterning Generates Human Excitatory Neurons with NMDAR-Mediated Synaptic

- Transmission. *Cell Rep* 23:2509–2523.
- Nestler EJ, Hyman SE (2010) Animal models of neuropsychiatric disorders. *Nat Neurosci* 13:1161–1169.
- Ni Y, Hui C, Du Y, Yue Z, Li S, Yu H, He Y-YY, Wang D, Chen Y-N, Du H-Y, Shi Z-Y, He L, He Y-YY, Wang D (2018) Serum proteomic profiling for autism using magnetic bead - assisted matrix - assisted laser desorption ionization time - of - flight mass spectrometry : a pilot study. *World J Pediatr*:233–237.
- Niinae T, Imami K, Sugiyama N, Ishihama Y (2020) Proximity Labeling-assisted Identification of Endogenous Kinase Substrates. *BioRxiv*.
- Nimchinsky EA, Oberlander AM, Svoboda K (2001) Abnormal Development of Dendritic Spines in FMR1 Knock-Out Mice. *21*:5139–5146.
- Nishina H, Wada T, Katada T (2004) Physiological roles of SAPK/JNK signaling pathway. *J Biochem* 136:123–126.
- Nourbakhsh K, Ferreccio AA, Bernard MJ, Yadav S (2021) TAOK2 is an ER-localized kinase that catalyzes the dynamic tethering of ER to microtubules. *Dev Cell* 56:3321-3333.e5.
- Novarino G et al. (2012) Mutations in BCKD-kinase lead to a potentially treatable form of autism with epilepsy. *Science* (80-) 338:394–397.
- Nowakowski TJ, Bhaduri A, Pollen AA, Alvarado B, Mostajo-Radji MA, Di Lullo E, Haeussler M, Sandoval-Espinosa C, Liu SJ, Velmeshev D, Ounadjela JR, Shuga J, Wang X, Lim DA, West JA, Leyrat AA, Kent WJ, Kriegstein AR (2017) Spatiotemporal gene expression trajectories reveal developmental hierarchies of the

- human cortex. *Science* (80-) 358:1318–1323.
- O’Roak BJ et al. (2012a) Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science* (80-) 338:1619–1622.
- O’Roak BJ et al. (2012b) Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* 485:246–250.
- Of AR, National THE, Ofner M, Coles A, Decou L Lou, Do MT, Bienek A, Snider J, Ugnat A-M (2018) Autism Spectrum Disorder among Children and Youth in Canada 2018.
- Ohno H, Hirabayashi S, Iizuka T, Ohnishi H, Fujita T, Hata Y (2002) Localization of p0071-interacting proteins, plakophilin-related armadillo-repeat protein-interacting protein (PAPIN) and ERBIN, in epithelial cells. *Oncogene* 21:7042–7049.
- Okazaki K, Yamamuro K, Iida J, Kishimoto T (2019) Guanfacine monotherapy for ADHD/ASD comorbid with Tourette syndrome: A case report. *Ann Gen Psychiatry* 18:1–5.
- Pan YH, Wu N, Yuan XB (2019) Toward a better understanding of neuronal migration deficits in autism spectrum disorders. *Front Cell Dev Biol* 7:1–8.
- Papke CM, Smolen KA, Swingle MR, Cressey L, Heng RA, Toporsian M, Deng L, Hagen J, Shen Y, Chung WK, Kettenbach AN, Honkanen RE (2021) A disorder-related variant (E420K) of a PP2A-regulatory subunit (PPP2R5D) causes constitutively active AKT-mTOR signaling and uncoordinated cell growth. *J Biol Chem* 296:100313.
- Parikshak NN, Luo R, Zhang A, Won H, Lowe JK, Chandran V, Horvath S, Geschwind

- DH (2013) Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell* 155:1008.
- Paşca SP, Portmann T, Voineagu I, Yazawa M, Shcheglovitov A, Paşca AM, Cord B, Palmer TD, Chikahisa S, Nishino S, Bernstein JA, Hallmayer J, Geschwind DH, Dolmetsch RE (2011) Using iPSC-derived neurons to uncover cellular phenotypes associated with Timothy syndrome. *Nat Med* 17:1657–1662.
- Patzke C, Han Y, Covy J, Yi F, Maxeiner S, Wernig M, Södhof TC (2015) Analysis of conditional heterozygous STXBP1 mutations in human neurons. *J Clin Invest* 125:3560–3571.
- Peça J, Feliciano C, Ting JT, Wang W, Wells MF, Venkatraman TN, Lascola CD, Fu Z, Feng G (2011) Shank3 mutant mice display autistic-like behaviours and striatal dysfunction. *Nature* 472:437–442.
- Peñagarikano O, Abrahams BS, Herman EI, Winden KD, Gdalyahu A, Dong H, Sonnenblick LI, Gruver R, Almajano J, Bragin A, Golshani P, Trachtenberg JT, Peles E, Geschwind DH (2011) Absence of CNTNAP2 Leads to Epilepsy, Neuronal Migration Abnormalities, and Core Autism-Related Deficits. *Cell* 147:235–246.
- Pennemann FL, Mussabekova A, Urban C, Stukalov A, Andersen LL, Grass V, Lavacca TM, Holze C, Oubraham L, Benamrouche Y, Girardi E, Boulos RE, Hartmann R, Superti-Furga G, Habjan M, Imler JL, Meignin C, Pichlmair A (2021) Cross-species analysis of viral nucleic acid interacting proteins identifies TAOs as innate immune regulators. *Nat Commun* 12:1–22.
- Penzes P, Cahill ME, Jones KA, VanLeeuwen J-E, Woolfrey KM (2011) Dendritic spine

- pathology in neuropsychiatric disorders. *Nat Neurosci* *14*:285–293.
- Pick JE, Ziff EB (2018) Regulation of AMPA receptor trafficking and exit from the endoplasmic reticulum. *Mol Cell Neurosci* *91*:3–9.
- Pintacuda G, Hsu Y-HH, Tsafou K, Li KW, Martin JM, Risemann J, Gonzalez-Lozano MA, Egri SB, Jaffe J, Smit AB, Fornelos N, Eggan KC, Lage K (2021) Interaction studies of risk proteins in human induced neurons reveal convergent biology and novel mechanisms underlying autism spectrum disorders. medRxiv:<https://doi.org/10.1101/2021.10.07.21264575>.
- Pinto D et al. (2010) Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* *466*:368–372.
- Post KL et al. (2020) Multi-model functionalization of disease-associated PTEN missense mutations identifies multiple molecular mechanisms underlying protein dysfunction. *Nat Commun* *11*.
- Pucilowska J et al. (2018) Pharmacological Inhibition of ERK Signaling Rescues Pathophysiology and Behavioral Phenotype Associated with 16p11.2 Chromosomal Deletion in Mice. *J Neurosci* *38*:6640–6652.
- Qaiser F, Yin Y, Mervis CB, Morris CA, Klein-Tasman BP, Tam E, Osborne LR, Yuen RKC (2021) Rare and low frequency genomic variants impacting neuronal functions modify the Dup7q11.23 phenotype. *Orphanet J Rare Dis* *16*:1–11.
- Quesnel-vallières M, Weatheritt RJ, Cordes SP, Blencowe BJ (2019) Autism spectrum disorder: insights into convergent mechanisms from transcriptomics. *Nat Rev Genet* *20*:51–63.

- Ramaswami G, Won H, Gandal MJ, Haney J, Wang JC, Wong CCY, Sun W, Prabhakar S, Mill J, Geschwind DH (2020) Integrative genomics identifies a convergent molecular subtype that links epigenomic with transcriptomic differences in autism. *Nat Commun* *11*:1–14.
- Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F (2013) Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* *8*:2281–2308.
- Rangaraju V, Lauterbach M, Schuman EM (2019) Spatially Stable Mitochondrial Compartments Fuel Local Translation during Plasticity. *Cell*:1–12.
- Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, Vilo J (2019) G:Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res* *47*:W191–W198.
- Regulator K, Sensitivity PI, Narita T, Nussenzweig A, Lukas J, Choudhary C (2018) Article DNA Repair Network Analysis Reveals Shieldin as a Article DNA Repair Network Analysis Reveals Shieldin as a Key Regulator of NHEJ and PARP Inhibitor Sensitivity. *Cell* *173*:972-974.e23.
- Reim D, Distler U, Halbedl S, Verpelli C, Sala C, Bockmann J, Tenzer S, Boeckers TM, Schmeisser MJ (2017) Proteomic Analysis of Post-synaptic Density Fractions from Shank3 Mutant Mice Reveals Brain Region Specific Changes Relevant to Autism Spectrum Disorder. *Front Mol Neurosci* *10*:1–10.
- Reimand J, Isserlin R, Voisin V, Kucera M, Tannus-Lopes C, Rostamianfar A, Wadi L, Meyer M, Wong J, Xu C, Merico D, Bader GD (2019) Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and

- EnrichmentMap. *Nat Protoc* 22:924–934.
- Richards AL, Eckhardt M, Krogan NJ (2021) Mass spectrometry-based protein–protein interaction networks for the study of human diseases. *Mol Syst Biol* 17:1–18.
- Richardson CD, Ray GJ, DeWitt MA, Curie GL, Corn JE (2016) Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat Biotechnol* 34:339–344.
- Richter M et al. (2019) Altered TAOK2 activity causes autism-related neurodevelopmental and cognitive abnormalities through RhoA signaling. *Mol Psychiatry* 24:1329–1350.
- Risch N, Hoffmann TJ, Anderson M, Croen LA, Grether JK, Windham GC (2014) Familial recurrence of autism spectrum disorder: Evaluating genetic and environmental contributions. *Am J Psychiatry* 171:1206–1213.
- Robinson EB et al. (2016) Genetic risk for autism spectrum disorders and neuropsychiatric variation in the general population. *Nat Genet* 48:552–555.
- Robinson EB, Samocha KE, Kosmicki JA, McGrath L, Neale BM, Perlis RH, Daly MJ (2014) Autism spectrum disorder severity reflects the average contribution of de novo and familial influences. *Proc Natl Acad Sci U S A* 111:15161–15165.
- Ronemus M, Iossifov I, Levy D, Wigler M (2014) The role of de novo mutations in the genetics of autism spectrum disorders. *Nat Rev Genet* 15:133–141.
- Rose S, Bennuri SC, Davis JE, Wynne R, Slattery JC, Tippet M, Delhey L, Melnyk S, Kahler SG, MacFabe DF, Frye RE (2018) Butyrate enhances mitochondrial function during oxidative stress in cell lines from boys with autism. *Transl Psychiatry* 8.

- Rose S, Frye RE, Slattery J, Wynne R, Tippett M, Melnyk S, James SJ (2014) Oxidative stress induces mitochondrial dysfunction in a subset of autistic lymphoblastoid cell lines. *Transl Psychiatry* 4:1–8.
- Rossignol DA, Frye RE (2012) Mitochondrial dysfunction in autism spectrum disorders: A systematic review and meta-analysis. *Mol Psychiatry* 17:290–314.
- Roux KJ, Kim DI, Burke B (2013) BioID: A screen for protein-protein interactions. *Curr Protoc Protein Sci*.
- Roux KJ, Kim DI, Raida M, Burke B (2012) A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J Cell Biol* 196:801–810.
- Roux L, Buzsáki G (2015) Tasks for inhibitory interneurons in intact brain circuits. *Neuropharmacology* 88:10–23.
- Ruggiero A, Katsenelson M, Slutsky I (2021) Trends in Mitochondria : new players in homeostatic regulation of firing rate set points. *Trends Neurosci* 44:605–618.
- Russo GL, Sonsalla G, Natarajan P, Breunig CT, Bulli G, Merl-Pham J, Schmitt S, Giehl-Schwab J, Giesert F, Jastroch M, Zischka H, Wurst W, Stricker SH, Hauck SM, Masserdotti G, Götz M (2021) CRISPR-Mediated Induction of Neuron-Enriched Mitochondrial Proteins Boosts Direct Glia-to-Neuron Conversion. *Cell Stem Cell* 28:524-534.e7.
- Ruzzo EK, Pe L, Jung J, Prober DA, Geschwind DH, Wall DP, Jung J, Wang L, Kashfaghghi D (2019) Inherited and De Novo Genetic Risk for Autism Article Inherited and De Novo Genetic Risk for Autism Impacts Shared Networks. *Cell* 178:850–866.

- Rylaarsdam L, Guemez-Gamboa A (2019) Genetic Causes and Modifiers of Autism Spectrum Disorder. *Front Cell Neurosci* 13:1–15.
- Sakai Y et al. (2011) Protein interactome reveals converging molecular pathways among autism disorders. *Sci Transl Med* 3:86ra49.
- Samocha KE et al. (2014) A framework for the interpretation of de novo mutation in human disease. *Nat Genet* 46:944–950.
- Sanders SJ et al. (2012) De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* 485:1–6.
- Sanders SJ et al. (2015) Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci. *Neuron* 87:1215–1233.
- Sandin S, Lichtenstein P, Kuja-halkola R, Larsson H, Hultman CM, Reichenberg A (2014) The Familial Risk of Autism. *JAMA - J Am Med Assoc* 311:1770–1777.
- Satterstrom FK et al. (2020) Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell* 180:568-584.e23.
- Savardi A, Borgogno M, De Vivo M, Cancedda L (2021) Pharmacological tools to target NKCC1 in brain disorders. *Trends Pharmacol Sci* 42:1009–1034.
- Schmeisser MJ et al. (2012) Autistic-like behaviours and hyperactivity in mice lacking ProSAP1/Shank2. *Nature* 486:256–260.
- Schmidt O, Pfanner N, Meisinger C (2010) Mitochondrial protein import: From proteomics to functional mechanisms. *Nat Rev Mol Cell Biol* 11:655–667.
- Schmitz MHA, Held M, Janssens V, Hutchins JRA, Hudecz O, Ivanova E, Goris J,

- Trinkle-Mulcahy L, Lamond AI, Poser I, Hyman AA, Mechtler K, Peters JM, Gerlich DW (2010) Live-cell imaging RNAi screen identifies PP2A-B55 α and importin- β 21 as key mitotic exit regulators in human cells. *Nat Cell Biol* 12:886–893.
- Schwenk J, Boudkkazi S, Kocylowski MK, Brechet A, Zolles G, Bus T, Costa K, Kollwe A, Jordan J, Bank J, Bildl W, Sprengel R, Kulik A, Roeper J, Schulte U, Fakler B (2019) An ER Assembly Line of AMPA-Receptors Controls Excitatory Neurotransmission and Its Plasticity. *Neuron* 104:680-692.e9.
- Seeling JM, Miller JR, Gil R, Moon RT, White R, Virshup DM (1999) Regulation of β -catenin signaling by the B56 subunit of protein phosphatase 2A. *Science* (80-) 283:2089–2091.
- Semple BD, Blomgren K, Gimlin K, Ferriero DM, Noble-Haeusslein LJ (2013) Brain development in rodents and humans: Identifying benchmarks of maturation and vulnerability to injury across species. *Prog Neurobiol* 106–107:1–16.
- Shang L, Henderson LB, Cho MT, Petrey DS, Fong CT, Haude KM, Shur N, Lundberg J, Hauser N, Carmichael J, Innis J, Schuette J, Wu YW, Asaikar S, Pearson M, Folk L, Retterer K, Monaghan KG, Chung WK (2016) De novo missense variants in PPP2R5D are associated with intellectual disability, macrocephaly, hypotonia, and autism. *Neurogenetics* 17:43–49.
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: A Software Environment for Integrated Models. *Genome Res* 13:2498–2504.

- Sharma K, Schmitt S, Bergner CG, Tyanova S, Kannaiyan N, Manrique-Hoyos N, Kongi K, Cantuti L, Hanisch UK, Philips MA, Rossner MJ, Mann M, Simons M (2015) Cell type- and brain region-resolved mouse brain proteome. *Nat Neurosci* 18:1819–1831.
- Shcheglovitov A, Shcheglovitova O, Yazawa M, Portmann T, Shu R, Sebastiano V, Krawisz A, Froehlich W, Bernstein J a, Hallmayer JF, Dolmetsch RE (2013) SHANK3 and IGF1 restore synaptic deficits in neurons from 22q13 deletion syndrome patients. *Nature* 503:267–271.
- Shen L, Feng C, Zhang K, Chen Y, Gao Y (2019a) Proteomics Study of Peripheral Blood Mononuclear Cells (PBMCs) in Autistic Children. *Front Cell Neurosci* 13:1–16.
- Shen M, Wang F, Li M, Sah N, Stockton ME, Tidei JJ, Gao Y, Korabelnikov T, Kannan S, Vevea JD, Chapman ER, Bhattacharyya A, van Praag H, Zhao X (2019b) Reduced mitochondrial fusion and Huntingtin levels contribute to impaired dendritic maturation and behavioral deficits in Fmr1-mutant mice. *Nat Neurosci* 22:386–400.
- Sheridan SD, Theriault KM, Reis SA, Zhou F, Madison JM, Daheron L, Loring JF, Haggarty SJ (2011) Epigenetic characterization of the FMR1 gene and aberrant neurodevelopment in human induced pluripotent stem cell models of fragile X syndrome. *PLoS One* 6.
- Shi Y, Kirwan P, Livesey FJ (2012a) Directed differentiation of human pluripotent stem cells to cerebral cortex neurons and neural networks. *Nat Protoc* 7:1836–1846.
- Shi Y, Kirwan P, Smith J, Robinson HPCC, Livesey FJ (2012b) Human cerebral cortex development from pluripotent stem cells to functional excitatory synapses. *Nat*

Neurosci 15:477–486.

Shulyakova N, Andreatza AC, Mills LR, Eubanks JH (2017) Mitochondrial dysfunction in the pathogenesis of rett syndrome: Implications for mitochondria-targeted therapies. *Front Cell Neurosci* 11:1–9.

Silva JC, Gorenstein M V., Li GZ, Vissers JPC, Geromanos SJ (2006) Absolute quantification of proteins by LCMSE: A virtue of parallel MS acquisition. *Mol Cell Proteomics* 5:144–156.

Sit ST, Manser E (2011) Rho GTPases and their role in organizing the actin cytoskeleton. *J Cell Sci* 124:679–683.

Smith SEP, Zhou Y-D, Zhang G, Jin Z, Stoppel DC, Anderson MP (2011) Increased gene dosage of Ube3a results in autism traits and decreased glutamate synaptic transmission in mice. *Sci Transl Med* 3:103ra97.

Snel B, Lehmann G, Bork P, Huynen MA (2000) String: A web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. *Nucleic Acids Res* 28:3442–3444.

Solowska JM, Garbern JY, Baas PW (2010) Evaluation of loss of function as an explanation for SPG4-based hereditary spastic paraplegia. *Hum Mol Genet* 19:2767–2779.

Sowa ME, Bennett EJ, Gygi SP, Harper JW (2009) Defining the Human Deubiquitinating Enzyme Interaction Landscape. *Cell* 138:389–403.

Spence EF, Dube S, Uezu A, Locke M, Soderblom EJ, Soderling SH (2019) In vivo proximity proteomics of nascent synapses reveals a novel regulator of cytoskeleton-

mediated synaptic maturation. *Nat Commun* 10:1–16.

Spratt PWE, Alexander RPD, Ben-Shalom R, Sahagun A, Kyoung H, Keeshen CM, Sanders SJ, Bender KJ (2021) Paradoxical hyperexcitability from NaV1.2 sodium channel loss in neocortical pyramidal cells. *Cell Rep* 36:1–15.

Spratt PWE, Ben-Shalom R, Keeshen CM, Burke KJ, Clarkson RL, Sanders SJ, Bender KJ (2019) The Autism-Associated Gene *Scn2a* Contributes to Dendritic Excitability and Synaptic Function in the Prefrontal Cortex. *Neuron* 103:673–685.e5.

Stein LR, Imai SI (2012) The dynamic regulation of NAD metabolism in mitochondria. *Trends Endocrinol Metab* 23:420–428.

Stessman HAF et al. (2017) Targeted sequencing identifies 91 neurodevelopmental-disorder risk genes with autism and developmental-disability biases. *Nat Genet.*

Stogsdill JA, Eroglu C (2017) The interplay between neurons and glia in synapse development and plasticity. *Curr Opin Neurobiol* 42:1–8.

Su H, Fan W, Coskun PE, Vesa J, Gold JA, Jiang YH, Potluri P, Procaccio V, Acab A, Weiss JH, Wallace DC, Kimonis VE (2011) Mitochondrial dysfunction in CA1 hippocampal neurons of the *UBE3A* deficient mouse model for Angelman syndrome. *Neurosci Lett* 487:129–133.

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP (2005) Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102:15545–15550.

Subramanian L, Calcagnotto ME, Paredes MF (2020) Cortical Malformations: Lessons in

- Human Brain Development. *Front Cell Neurosci* 13:1–17.
- Suetterlin P et al. (2018) Altered neocortical gene expression, brain overgrowth and functional over-connectivity in *chd8* haploinsufficient mice. *Cereb Cortex* 28:2192–2206.
- Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S (2007) Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell* 131:861–872.
- Takano T, Wallace JT, Baldwin KT, Purkey AM, Uezu A, Courtland JL, Soderblom EJ, Shimogori T, Maness PF, Eroglu C, Soderling SH (2020) Chemico-genetic discovery of astrocytic control of inhibition in vivo. *Nature* 588:296–302.
- Takeda K, Ichijio H (2002) Neuronal p38 MAPK signalling: An emerging regulator of cell fate and function in the nervous system. *Genes to Cells* 7:1099–1111.
- Tang G, Gudsnuk K, Kuo SH, Cotrina ML, Rosoklija G, Sosunov A, Sonders MS, Kanter E, Castagna C, Yamamoto A, Yue Z, Arancio O, Peterson BS, Champagne F, Dwork AJ, Goldman J, Sulzer D (2014) Loss of mTOR-Dependent Macroautophagy Causes Autistic-like Synaptic Pruning Deficits. *Neuron* 83:1131–1143.
- Tang YG, Zucker RS (1997) Mitochondrial involvement in post-tetanic potentiation of synaptic transmission. *Neuron* 18:483–491.
- Taurines R, Dudley E, Conner AC, Grassl J, Jans T, Guderian F, Andreas CM, Gerlach M, Thome J (2010) Serum protein profiling and proteomics in autistic spectrum disorder using magnetic bead-assisted mass spectrometry. :249–255.
- Thompson A, Schäfer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, Neumann T, Hamon

- C (2003) Tandem mass tags: A novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal Chem* 75:1895–1904.
- Thomson JA (1998) Embryonic stem cell lines derived from human blastocysts. *Science* (80-) 282:1145–1147.
- Tian Y, Voineagu I, Paşca SP, Won H, Chandran V, Horvath S, Dolmetsch RE, Geschwind DH (2014) Alteration in basal and depolarization induced transcriptional network in iPSC derived neurons from Timothy syndrome. *Genome Med* 6:75.
- Tiberi L, Van Den Aemele J, Dimidschstein J, Piccirilli J, Gall D, Herpoel A, Bilheu A, Bonnefont J, Iacovino M, Kyba M, Bouschet T, Vanderhaeghen P (2012) BCL6 controls neurogenesis through Sirt1-dependent epigenetic repression of selective Notch targets. *Nat Neurosci* 15:1627–1635.
- Tonge BJ, Bull K, Brereton A, Wilson R (2014) A review of evidence-based early intervention for behavioural problems in children with autism spectrum disorder: the core components of effective programs, child-focused interventions and comprehensive treatment models. *Curr Opin Psychiatry* 27:158–165.
- Tracy TE et al. (2022) Tau interactome maps synaptic and mitochondrial processes associated with neurodegeneration. *Cell* 185:712-728.e14.
- Tracy TE, Madero-pérez J, Swaney D, Chang TS, Konrad C, Ward ME, Stevenson E, Hüttenhain R, Kauwe G, Mercedes M, Sweetland-martin L, Chen X, Mok S, Min S, Wang C, Sohn PD, Martin J (2021) Tau interactome mapping reveals dynamic processes in synapses and mitochondria associated with neurodegenerative disease. [bioRxiv:https://doi.org/10.1101/2021.06.17.448349](https://doi.org/10.1101/2021.06.17.448349).

- Tremblay R, Lee S, Rudy B (2016) GABAergic Interneurons in the Neocortex: From Cellular Properties to Circuits. *Neuron* 91:260–292.
- Tromans S, Chester V, Kiani R, Alexander R, Brugha T (2018) The Prevalence of Autism Spectrum Disorders in Adult Psychiatric Inpatients: A Systematic Review. *Clin Pract Epidemiol Ment Heal* 14:177–187.
- Trost B et al. (2020) Genome-wide detection of tandem DNA repeats that are expanded in autism. *Nature* 586:80–86.
- Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, Mann M, Cox J (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* 13:731–740.
- Uddin LQ, Dajani DR, Voorhies W, Bednarz H, Kana RK (2017a) Progress and roadblocks in the search for brain-based biomarkers of autism and attention-deficit/hyperactivity disorder. *Transl Psychiatry* 7:e1218.
- Uddin M, Woodbury-Smith M, Chen A, Brunga L, Lamoureux S, Pellecchia G, Yuen RKC, Faheem M, Stavropoulos DJ, Drake J, Hahn CD, Shlien A, Marshall CR, Turner LA, Minassian BA, Scherer SW (2017b) Germline and somatic mutations in STXBP1 with diverse neurodevelopmental phenotypes. *Neurol Genet* 3.
- Uezu A, Kanak DJ, Bradshaw TWA, Soderblom EJ, Catavero CM, Burette AC, Weinberg RJ, Soderling SH (2016) Identification of an elaborate complex mediating postsynaptic inhibition. *Science* (80-) 353:1123–1129.
- Ultanir SK, Yadav S, Hertz NT, Oses-Prieto JA, Claxton S, Burlingame AL, Shokat KM, Jan LY, Jan Y-N (2014) MST3 Kinase Phosphorylates TAO1/2 to Enable Myosin

Va Function in Promoting Spine Synapse Development. *Neuron* 84:968–982.

Urano J, Sato T, Matsuo T, Otsubo Y, Yamamoto M, Tamanoi F (2007) Point mutations in TOR confer Rheb-independent growth in fission yeast and nutrient-independent mammalian TOR signaling in mammalian cells. *Proc Natl Acad Sci U S A* 104:3514–3519.

Urresti J, Zhang P, Moran-Losada P, Yu NK, Negraes PD, Trujillo CA, Antaki D, Amar M, Chau K, Pramod AB, Diedrich J, Tejwani L, Romero S, Sebat J, Yates JR, Muotri AR, Iakoucheva LM (2021) Cortical organoids model early brain development disrupted by 16p11.2 copy number variants in autism. *Mol Psychiatry* 26:7560–7580.

Varghese M, Keshav N, Jacot-Descombes S, Warda T, Wicinski B, Dickstein DL, Harony-Nicolas H, De Rubeis S, Drapeau E, Buxbaum JD, Hof PR (2017) Autism spectrum disorder: neuropathology and animal models. *Acta Neuropathol* 134:537–566.

Velikonja T, Fett AK, Velthorst E (2019) Patterns of Nonsocial and Social Cognitive Functioning in Adults with Autism Spectrum Disorder: A Systematic Review and Meta-analysis. *JAMA Psychiatry* 76:135–151.

Velmeshev D, Schirmer L, Jung D, Haeussler M, Perez Y, Mayer S, Bhaduri A, Goyal N, Rowitch DH, Kriegstein AR (2019) Single-cell genomics identifies cell type-specific molecular changes in autism. *Science (80-)* 364:685–689.

Ventura P, de Giambattista C, Spagnoletta L, Trerotoli P, Cavone M, Di Gioia A, Margari L (2020) Methylphenidate in autism spectrum disorder: A long-term follow

up naturalistic study. *J Clin Med* 9:1–17.

Verstreken P, Ly C V., Venken KJT, Koh TW, Zhou Y, Bellen HJ (2005) Synaptic mitochondria are critical for mobilization of reserve pool vesicles at *Drosophila* neuromuscular junctions. *Neuron* 47:365–378.

Vismara L a, Rogers SJ (2010) Behavioral treatments in autism spectrum disorder: what do we know? *Annu Rev Clin Psychol* 6:447–468.

Voineagu I, Wang X, Johnston P, Lowe JK, Tian Y, Horvath S, Mill J, Cantor RM, Blencowe BJ, Geschwind DH (2011) Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature* 474:380–384.

Wagner A et al. (2021) Metabolic modeling of single Th17 cells reveals regulators of autoimmunity. *Cell* 184:4168-4185.e21.

Wang C, Ward ME, Chen R, Liu K, Tracy TE, Chen X, Xie M, Sohn PD, Ludwig C, Meyer-Franke A, Karch CM, Ding S, Gan L (2017a) Scalable Production of iPSC-Derived Human Neurons to Identify Tau-Lowering Compounds by High-Content Screening. *Stem Cell Reports* 9:1221–1233.

Wang K et al. (2009) Common genetic variants on 5p14 . 1 associate with autism spectrum disorders. *Nature* 459:528–533.

Wang P, Lin M, Pedrosa E, Hrabovsky A, Zhang Z, Guo W, Lachman HM, Zheng D (2015) CRISPR/Cas9-mediated heterozygous knockout of the autism gene CHD8 and characterization of its transcriptional networks in neurodevelopment. *Mol Autism* 6:55.

Wang P, Mokhtari R, Pedrosa E, Kirschenbaum M, Bayrak C, Zheng D, Lachman HM

- (2017b) CRISPR / Cas9-mediated heterozygous knockout of the autism gene CHD8 and characterization of its transcriptional networks in cerebral organoids derived from iPS cells. :1–17.
- Wang S, Tukachinsky H, Romano FB, Rapoport TA (2016a) Cooperation of the ER-shaping proteins atlastin, lunapark, and reticulons to generate a tubular membrane network. *Elife* 5:1–29.
- Wang Y, Picard M, Gu Z (2016b) Genetic Evidence for Elevated Pathogenicity of Mitochondrial DNA Heteroplasmy in Autism Spectrum Disorder. *PLoS Genet* 12:1–24.
- Warrier V et al. (2022) Genetic correlates of phenotypic heterogeneity in autism. *Nat Genet*.
- Weiner DJ et al. (2017) Polygenic transmission disequilibrium confirms that common and rare variation act additively to create risk for autism spectrum disorders. *Nat Genet* 49:978–985.
- White BH (2016) What genetic model organisms offer the study of behavior and neural circuits. *J Neurogenet* 30:54–61.
- Wilfert AB et al. (2021) Recent ultra-rare inherited variants implicate new autism candidate risk genes. *Nat Genet* 53:1125–1134.
- Willsey AJ et al. (2013) Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. *Cell* 155:997.
- Willsey HR, Exner CRT, Xu Y, Everitt A, Sun N, Wang B, Dea J, Schmunk G, Zaltsman Y, Teerikorpi N, Kim A, Anderson AS, Shin D, Seyler M, Nowakowski TJ, Harland

- RM, Willsey AJ, State MW (2021) Parallel in vivo analysis of large-effect autism genes implicates cortical neurogenesis and estrogen in risk and resilience. *Neuron* 109:788–804.
- Willsey HR, Willsey AJ, Wang B, State MW (2022) Genomics, convergent neuroscience and progress in understanding autism spectrum disorder. *Nat Rev Neurosci* 23:323–341.
- Wilton DK, Dissing-Olesen L, Stevens B (2019) Neuron-Glia Signaling in Synapse Elimination. *Annu Rev Neurosci* 42:107–127.
- Wolff JJ, Gu H, Gerig G, Elison JT, Styner M, Gouttard S, Botteron KN, Dager SR, Dawson G, Estes AM, Evans AC, Hazlett HC, Kostopoulos P, McKinstry RC, Paterson SJ, Schultz RT, Zwaigenbaum L, Piven J (2012) Differences in white matter fiber tract development present from 6 to 24 months in infants with autism. *Am J Psychiatry* 169:589–600.
- Wong WR, Brugman KI, Maher S, Oh JY, Howe K, Kato M, Sternberg PW (2019) Autism-Associated missense genetic variants impact locomotion and neurodevelopment in *Caenorhabditis elegans*. *Hum Mol Genet* 28:2271–2281.
- Wu Y, Li R, Sun S, Weile J, Roth FP (2021) Improved pathogenicity prediction for rare human missense variants. *Am J Hum Genet* 108:1891–1906.
- Xu J et al. (2018) Inhibition of STEP 61 ameliorates deficits in mouse and hiPSC-based schizophrenia models. *Mol Psychiatry* 23:271–281.
- Xu ZX, Kim GH, Tan JW, Riso AE, Sun Y, Xu EY, Liao GY, Xu H, Lee SH, Do NY, Lee CH, Clipperton-Allen AE, Kwon S, Page DT, Lee KJ, Xu B (2020) Elevated

protein synthesis in microglia causes autism-like synaptic and behavioral aberrations. *Nat Commun* 11.

Yacoub N al, Romanowska M, Haritonova N, Foerster J (2012) Optimized production and concentration of lentiviral vectors containing large inserts. *J Gene Med* 9:579–584.

Yadav S, Oses-Prieto JA, Peters CJ, Zhou J, Pleasure SJ, Burlingame AL, Jan LY, Jan Y-N, Samuel J, Burlingame AL, Jan LY, Jan Y-N (2017) TAOK2 Kinase Mediates PSD95 Stability and Dendritic Spine Maturation through Septin7 Phosphorylation. *Neuron* 93:1–15.

Yang R, Walder-Christensen KK, Kim N, Wu D, Lorenzo DN, Badea A, Jiang YH, Yin HH, Wetsel WC, Bennett V (2019) ANK2 autism mutation targeting giant ankyrin-B promotes axon branching and ectopic connectivity. *Proc Natl Acad Sci U S A* 116:15262–15271.

Yang W et al. (2021) Rheb mediates neuronal-activity-induced mitochondrial energetics through mTORC1-independent PDH activation. *Dev Cell* 56:811-825.e6.

Yardeni T, Cristancho AG, McCoy AJ, Schaefer PM, McManus MJ, Marsh ED, Wallace DC (2021) An mtDNA mutant mouse demonstrates that mitochondrial deficiency can result in autism endophenotypes. *Proc Natl Acad Sci U S A* 118:1–11.

Yasuda S, Tanaka H, Sugiura H, Okamura K, Sakaguchi T, Tran U, Takemiya T, Mizoguchi A, Yagita Y, Sakurai T, De Robertis EM, Yamagata K (2007) Activity-Induced Protocadherin Arcadlin Regulates Dendritic Spine Number by Triggering N-Cadherin Endocytosis via TAO2 β and p38 MAP Kinases. *Neuron* 56:456–471.

- Yi F, Danko T, Botelho SC, Patzke C, Pak C, Wernig M, Südhof TC (2016) Autism-associated SHANK3 haploinsufficiency causes I channelopathy in human neurons. *Science* (80-) 352:aaf2669.
- Yoon S, Munoz A, Yamrom B, Lee Y ha, Andrews P, Marks S, Wang Z, Reeves C, Winterkorn L, Krieger AM, Buja A, Pradhan K, Ronemus M, Baldwin KK, Levy D, Wigler M, Iossifov I (2021) Rates of contributory de novo mutation in high and low-risk autism families. *Commun Biol* 4.
- Youn JY, Dunham WH, Hong SJ, Knight JDR, Bashkurov M, Chen GI, Bagci H, Rathod B, MacLeod G, Eng SWM, Angers S, Morris Q, Fabian M, Côté JF, Gingras AC (2018) High-Density Proximity Mapping Reveals the Subcellular Organization of mRNA-Associated Granules and Bodies. *Mol Cell* 69:517-532.e11.
- Yuen RKC et al. (2017) Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. *Nat Neurosci* 20:13–17.
- Zang Y, Chaudhari K, Bashaw GJ (2021) Chapter Four - New insights into the molecular mechanisms of axon guidance receptor regulation and signaling. In: *Molecular Mechanisms of Neural Development and Insights into Disease* (Bashaw GJBT-CT in DB, ed), pp 147–196. Academic Press.
- Zehnder T, Petrelli F, Romanos J, Polleux F, Santello M, Bezzi P, Zehnder T, Petrelli F, Romanos J, Figueiredo ECDO, Lewis TL (2021) Article Mitochondrial biogenesis in developing astrocytes regulates astrocyte maturation and synapse formation
Graphical abstract II Mitochondrial biogenesis in developing astrocytes regulates astrocyte maturation and synapse formation.

- Zerbo O, Iosif AM, Walker C, Ozonoff S, Hansen RL, Hertz-Picciotto I (2013) Is Maternal Influenza or Fever during Pregnancy Associated with Autism or Developmental Delays? Results from the CHARGE (childhood Autism Risks from Genetics and Environment) Study. *J Autism Dev Disord* 43:25–33.
- Zhang Y, Pak CH, Han Y, Ahlenius H, Zhang Z, Chanda S, Marro S, Patzke C, Acuna C, Covy J, Xu W, Yang N, Danko T, Chen L, Wernig M, S?dhof TC (2013) Rapid single-step induction of functional neurons from human pluripotent stem cells. *Neuron* 78:785–798.
- Zhao X, Bhattacharyya A (2018) Human Models Are Needed for Studying Human Neurodevelopmental Disorders. *Am J Hum Genet* 103:829–857.
- Zhou T, Raman M, Gao Y, Earnest S, Chen Z, Machius M, Cobb MH, Goldsmith EJ (2004) Crystal Structure of the TAO2 Kinase Domain. *Structure* 12:1891–1900.
- Zihni C, Mitsopoulos C, Tavares IA, Baum B, Ridley AJ, Morris JDH, Blebbing M, Zihni C, Mitsopoulos C, Tavares IA, Baum B, Ridley AJ, Morris JDH (2006) Prostate-derived Sterile 20-like Kinase 1- Induces Apoptosis. *J Biol Chem* 282:6484–6493.