Functional genomic dissection of human brain development, degeneration, and dysfunction in pediatric neurogenetic disorders

Functional genomic dissection of human brain development, degeneration, and dysfunction in pediatric neurogenetic disorders

A DISSERTATION PRESENTED
BY
JENNY LAI
TO
THE DIVISION OF MEDICAL SCIENCES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN THE SUBJECT OF
NEUROSCIENCE

HARVARD UNIVERSITY
CAMBRIDGE, MASSACHUSETTS

AUGUST 2023

Thesis advisors: Timothy W. Yu and Eunjung Alice Lee Jenny Lai

Functional genomic dissection of human brain development, degeneration, and dysfunction in pediatric neurogenetic disorders

Abstract

Human genetics has implicated hundreds of genes in brain development, maintenance, and function. Experimentally tractable animal models and culture systems have contributed greatly to understanding the function of these genes, but have limitations, beginning with their inability to capture human-specific biology.

The Introduction chapter provides background on recent experimental approaches for studying and modeling human neurogenetic diseases. The combination of modern technological developments in functional genomics applied to human brain tissue and stem-cell based models has offered novel ways to molecularly dissect human neurological disease.

Chapter 1 focuses on applying this modern human-based genomic approach to probe the molecular mechanisms underlying a human neurogenetic disorder in which the genetic cause has been known for over 25 years. Ataxia-Telangiectasia (A-T) is an autosomal recessive multi-system disorder caused by mutations in the gene *ATM* that presents with progressive cerebellar neurodegeneration. Understanding disease pathogenesis has been hampered by the inability to generate animal models that recapitulate hallmark human features of disease. We address this gap by analyzing single-nucleus RNA-sequencing (snRNAseq) of postmortem brain tissue from individuals with A-T, generating mechanistic hypotheses that can then be tested using patient-derived induced pluripotent stem cells (iPSCs). We identified transcriptomic signatures of microglial inflammation in our snRNAseq analyses of postmortem A-T brains. To experimentally corroborate this finding,

we studied microglia and neurons generated from A-T patient versus control iPSCs. Transcriptomic profiling of A-T iPSC-derived microglia confirmed cell-intrinsic microglial activation of cytokine production and innate immune response pathways compared to controls. Furthermore, adding A-T microglia to co-cultures with either control or A-T iPSC-derived neurons was sufficient to induce cytotoxicity. Taken together, these studies reveal that cell-intrinsic microglial activation may play a critical role in the development and progression of neurodegeneration in Ataxia Telangiectasia.

Chapter 2 describes the discovery of a novel genetic cause of a neurodevelopmental disorder (NDD) that presents with Autism Spectrum Disorder (ASD), attention-deficit hyperactivity disorder (ADHD), and behavioral dysregulation. We assembled a cohort of 38 individuals with de novo loss-of-function (LoF) variants in the ciliogenic RFX family of transcription factors (*RFX3, RFX4*, and *RFX7*) and demonstrate that pathogenic variants in these genes are associated with an overlapping neurobehavioral phenotype. We show that *RFX3, RFX4*, and *RFX7* have enriched expression in the developing and adult human brain and their X-box binding motif is enriched in cis-regulatory regions of known ASD risk genes. These results implicate deleterious variation in RFX transcription factors in cases of monogenic NDD and highlight them as potential critical transcriptional regulators of neurobiological processes underlying NDD pathogenesis.

In Chapter 3 we dissect the molecular roles of *RFX3*, using human iPSC-derived neurons and forebrain organoids to understand how haploinsufficiency of *RFX3* alters brain development and neuronal function. Transcriptomic analyses of *RFX3* dosage demonstrated disruption of ciliary gene expression in RFX3-/- neurons (expected based on prior studies of *RFX3* in mice and worms) while analyses of *RFX3*+/- neurons revealed a previously unanticipated role for *RFX3* in synaptic gene expression. *RFX3* deficiency led to decreased synchronization of neural network activity and impaired induction of CREB targets in response to neuronal depolarization. Our results highlight a novel role of the ciliogenic transcription factor *RFX3* in shaping activity dependent responses and synaptic plasticity in human neurons that is disrupted by haploinsufficiency.

Chapter 4 highlights the potential for gene therapy development for neurogenetic conditions including *RFX3* haploinsufficiency. Antisense oligonucleotides (ASO) are short, synthetic, sequences that can modulate gene expression levels in a titratable manner through several different mechanisms. ASOs have the potential to restore functional protein levels in neurogenetic disease by targeting mRNA processing inefficiencies. We found that the wild-type allele of *RFX3* has a naturally occurring skipped exon that leads to nonsense-mediated decay (NMD). ASOs can promote inclusion of the exon, thereby rescuing the transcript from NMD and increasing *RFX3* levels in haploinsufficient human neurons. Our results suggest that the neurodevelopmental disorder caused by *RFX3* haploinsufficiency may be amenable to ASO-based therapy.

The Conclusion chapter synthesizes the insights from the studies presented and delineates potential future directions. Taken together, this thesis highlights how functional genomics applied to human brain tissue and iPSC-derived models of pediatric neurogenetic disorders can enable mechanistic and interventional insights.

Contents

Listing of figures

To my eternal partner, Jae, and our beloved parents: here's to the beautiful life we will live together.

Acknowledgments

Every day I think about how fortunate I am to live this life as a growing physician-scientist, one that has only been possible because of the supportive people in my life. As an undergraduate fascinated by the brain and introduced to stem cells and regenerative medicine around the time that induced pluripotent stem cell models, gene editing, and functional genomics were developed, I dreamed of one day performing experiments just like the ones in this thesis.

I first want to thank my incredible advisors, Tim Yu and Alice Lee, for making this dream possible. I first met Tim when he taught our genetics course in the fall of my first year of medical school. The way Tim taught us human genetics was so engaging and clinically relevant that it inspired me to consider studying human genetics for my PhD. I was excited to learn that Tim is also a physician-scientist and runs a lab at Harvard Medical School/Boston Children's Hospital. Tim is the physician-scientist I want to emulate. He has the patience, warmth, and kindness that brings healing and hope to the patients we work with. He also has the sharpness and creativity to make breakthroughs in science and medicine. I was introduced to Alice in my first year of graduate school when I was in

search for co-mentor with expertise in computational biology. As someone without previous bioinformatics experience, I am so thankful to Alice for welcoming me to her lab and providing me the opportunity to learn and grow my computational skills during my PhD. Alice has an incredible attention to detail, and I learned so much about computational biology, statistical analyses, and data visualization from her. She is patient, kind, and has a pure love for contributing to science. I am grateful for the supportive environment, mentorship, and guidance Tim and Alice have given me.

I also want to thank all the members of the Yu Lab and Lee Lab for being amazing colleagues. Their support, collaboration, feedback during lab meetings, and fun energy made lab a fun place to work for the past few years. I especially want to thank Didem Demirbas for all that she's taught me about iPSCs and differentiation into neurons and organoids. She is the reason the iPSC-based experiments were possible. I'll always have fond memories of being in the TC room with Didem for hours, on weekends, and late into the evening as we took care of our cells and chatted about life. I also want to thank Boxun Zhao for all the bioinformatics and library prep he taught me early in my PhD. Bo was generous with his time and taught me how to prepare sequencing libraries from start to finish - an experiment that took multiple long days on the weekend. I want to thank Junho Kim for being my computational mentor and helping me get started on single-cell RNA-sequencing analysis in my first few months of graduate school. His encouragement and support made the A-T project possible. I also want to thank Luca Fusar Bassini, Katie Phillips, and Harrison Wallace for being great teammates and all their help on the RFX3 project. I am also thankful to wonderful collaborators, including Mike Lodato and Eunha Kim, for the opportunities they provided me and all their support.

In addition to my PhD advisors, I am thankful to my MD-PhD advisor Chinfei Chen for her support and guidance from even before the MD-PhD journey started. She was always available to meet and provide comforting advice, especially during critical decisions and times of need in both medical school and graduate school. I would also like to thank the MD-PhD office, especially Loren

Walensky, Amy Cohen, and Liz Barks for their support throughout the MD-PhD journey. A special thank you to Loren who made my dreams of becoming a physician-scientist come true.

This journey all started 10 years ago when I sat in premed classes with my best friends and college roommates Dr. Eileen Feng and Dr. Tiffany Yu. I was so lucky to have almost all of my classes with Eileen, a fellow HDRB concentrator, where we learned about stem cells and research together. She is there for all the ups and downs of lab, from failed western blots and qPCRs to paper submissions. Tiff is my gageenang, someone who has always been my cheerleader and told me that I was "born to defend". I am incredibly thankful to them for their friendship throughout college, medical school, and graduate school, and for the life milestones we have been able to celebrate together. They are inspirational physicians who bring healing to those around them, and will always be great role models for the rest of my medical career.

The longest days in the hospital or lab were always made more bearable because of the support of my love (and at this time, fiance!) Jae. On our first date, I was just a junior in college with a dream to apply to MD-PhD programs. His constant support turned this dream into a reality. From always greeting me at the airport when I came back from MD-PhD interviews, to cooking delicious meals every day while I work or study, to discussing my scientific problems and brainstorming data/statistical approaches together, Jae is the best partner in life I could have dreamed of. I also want to thank Jae's family, Abeonim, Eomeonim, Jihye, Jae Cheon, and Sujin for their support and encouragement along this journey. They always make me feel at home when we visit them and I am thankful for our wonderful memories in Boston, Philly, and Korea together.

Most importantly, I want to thank my family. I want to thank my older sister Julie, for always being a role model and inspiration throughout our childhood. Julie sparked my interest in biology and medicine – we volunteered at a local hospital together, had research internships at the UofM together, and cared for our family together. As first-generation students, she was a trailblazer who showed me the way through high school and college. I want to thank my younger siblings, Brandon

and Brianna for their encouragement and support. They always cheered me on and made Boston a fun place during several summers along this journey. I want to thank my Ama for taking care of us while we were young and teaching us many life skills. Above all, I am immensely grateful to my parents Ma and Ba for their unconditional love and support. They are the most selfless, generous, thoughtful, and resilient people I know. Their hard work, courage, and sacrifices made this life possible. These degrees are truly earned by them.

0

Introduction

0.1 Benefits and limitations of traditional approaches to studying pediatric neurogenetic disease

There are an estimated 7,000 rare and ultra-rare diseases caused by deleterious variants in single genes. Remarkably, 90% of rare childhood diseases are neurological disorders or have major neurological effects[195]. Pediatric neurogenetic disorders often present in early childhood with brain

malformations, neurodevelopmental delay, neurodegeneration, or multi-systemic symptoms[93]. Human genetics has elucidated pathogenic variants in hundreds of genes that cause highly penetrant monogenic neurological disorders, but elucidating the mechanisms by which these mutations lead to disease has been challenging. Tractable and disease-recapitulating genetic models have the power to illuminate the core functions of genes, understand disease pathogenesis, and enable development of targeted therapies.

Animal models have been instrumental to advancing the understanding of both fundamental neurobiology and disease. First, animal models can have features that make them particularly accessible for studying aspects of nervous system development and function. One example is the large, externally developing nature of Xenopus embryos, which was instrumental to studies that elucidated the process of neural induction[90,91,92]. Another is the rapid life cycles and well-characterized genomes of Drosophila and C. elegans that make them excellent tools for genetic screens. Animal models are also amenable to complex genetic manipulation, such as conditional knockout with spatiotemporal specificity, to understand the function of a gene in a specific cell type or developmental time point. Second, animal models that are simpler with defined anatomy, such as C. elegans where all 302 neurons and 7,000 synapses are identified with clear lineages, anatomical positions, and morphology, provide elegant and robust systems for understanding how specific genes specify neuronal fate, wiring, and function[282]. Third, animal models serve as genetically tractable systems for studying behavior and their underlying circuits. Lastly, animal models of neurological disease can contribute to our understanding of disease pathogenesis by allowing for *in vivo* disease modeling under physiological conditions[264,117].

However, inherent differences between animal models and humans impose constraints with respect to extrapolating findings to human neurobiology and disease. While rodents are the most common animal model in neuroscience, [121] emerging insights on human-specific characteristics of brain development reveal several limitations. The rodent brain is simply significantly less com-

plex than the human brain; it has orders of magnitude less neurons, lacks gyri and sulci, and has decreased cellular diversity[17]. Moreover, while many neuronal and non-neuronal cell types are conserved between rodents and humans, species-specific cell types exist. For example, outer radial glia give rise to the majority of cortical neurons in humans and are critical for the expansion of the human cerebral cortex, but are relatively rare in the mouse cortex[210]. Rodents and humans also diverge on epigenomic, genomic, and transcriptomic levels. Cell type specific chromatin accessibility state and DNA methylation signatures differed between humans and mice[17]. The human genome also contains sequences with high rates of human-specific substitutions (defined as human accelerated regions)[209] and human-specific genomic rearrangements such as deletions and segmental duplications[147]. Transcriptomically, species-specific patterns of alternative splicing have been observed[19,181]. In addition, neuronal transcriptional response to external stimuli (activity-dependent gene expression programs) also differ between humans and mice[213,11]. These genetic differences may affect the regulation of gene expression and therefore cellular phenotypes in rodents compared to humans. Lastly, the symptoms for establishing neuropsychiatric diagnoses in humans often do not have clear corresponding symptoms in rodents[194], and animal models sometimes do not recapitulate the severity of human disease phenotypes. For example, Ataxia-Telangiectasia (A-T) is a progressive neurodegenerative condition caused by biallelic loss-of-function mutations in *ATM*, and is characterized by severe ataxia due to atrophy of the cerebellum[234]. Yet, rodent models of complete *ATM* loss fail to recapitulate the progressive ataxia and cerebellar degeneration characteristic of the human disease for reasons that are still unclear[144]. Overall, these limitations hinder our ability to understand the molecular mechanisms by which mutations in a single gene contribute to disease pathogenesis.

0.2 Approaches to studying human-specific biology

Powerful new experimental techniques have greatly enhanced our ability to understand unique aspects of human biology, ranging from the ability to understand postmortem human disease pathology at single-cell resolution, to the ability to recapitulate human developmental processes responsible for generating specific cell types and tissues with induced pluripotent stem cells (iPSCs). Human tissue and iPSC-derived models are complementary approaches that together provide an opportunity to discover unique insights into human neurobiology and disease while overcoming certain limitations of animal models.

Human postmortem brain tissue provides insights on patient-specific brain pathology. While traditionally post-mortem tissue could only be studied histologically, recent advances in molecular biology and functional genomic technologies permit molecular dissection of post-mortem tissue. Single-nucleus RNA-sequencing (snRNA-seq) of normal human postmortem brain tissue deepened our understanding of the cellular diversity of the human brain [138,137]. snRNA-seq has also revealed cell type specific molecular alterations in several human neurological disorders from neurodevelopment to neurodegeneration, including autism spectrum disorder (ASD), epilepsy, major depressive disorder, Alzheimer's disease (AD), and multiple sclerosis [275,207,191,175,118]. These studies illustrate how snRNA-seq of postmortem tissue can provide insight into the transcriptional signatures of disease, implicate particular cell types, and propose pathophysiological mechanisms.

There are several important factors to consider when studying human tissues. While postmortem tissue allows insights into patient pathology, it often represents the disease end point, and freshly isolated brain tissue from desired stages of disease in living patients (especially pediatric patients) is difficult to obtain. Moreover, individuals have covariates such as postmortem interval (PMI), age, sex, and comorbid conditions that need to be controlled through closely matched cohorts of patients and controls [302]. In addition, postmortem tissue is not dynamic and cannot be manip-

ulated for mechanistic studies, making it difficult to distinguish between primary and secondary consequences of disease. Therefore, combining the study of human tissues with an experimentally manipulatable system permits the testing of pathophysiological mechanisms.

Human induced pluripotent stem cells (iPSCs) have revolutionized our ability to generate experimentally manipulatable human-based models of disease (1). iPSCs are cells that are capable of self-renewal and differentiation into any somatic cell type[259]. In a landmark study, human dermal fibroblasts were reprogrammed into iPSCs using four transcription factors, OCT_4, SOX_2, KLF_4, and C-MYC[259]. Since then, numerous protocols for differentiating iPSCs into cell types of the brain have emerged. Early protocols for generating cortical glutamatergic neurons involved small-molecule inhibition of BMP and TGFb pathways ("dual SMAD inhibition")[153,239]. Then, it was demonstrated that overexpression of a single transcription factor, *NGN2*, led to efficient generation of a relatively homogeneous population of cortical L2/3-like excitatory neurons in under 2 weeks[299]. GABAergic neurons can also be generated from iPSCs using small molecule inhibitors of WNT and SMAD signaling coupled with SHH to induce ventral telencephalic fate[160], or transcription factor reprogramming with overexpression of *ASCL* and *DLX2*[290]. Protocols for differentiation of more specialized neurons, such as dopaminergic and motor neurons have also been developed[135,100]. In addition to iPSC-derived neuronal models, it is also possible to differentiate glial cells from iPSCs. Overexpression of *SOX9* and *NFIB* differentiates iPSCs into astrocyte-like cells, while overexpression of *SOX10, OLIG2*, and *NKX6.2* can induce formation of oligodendrocytes[38,63]. iPSCs can also be differentiated into brain-resident cells from non-neural lineages, such as microglia, through overexpression of *SPI1* and *CEBPA*[40]. Importantly, each of these iPSC-derived cell types molecularly and functionally resemble the corresponding cell type in the human brain. They had similar transcriptome profiles and demonstrated similar key functional properties, such as neuronal action potential firing[299], astrocytic glutamate uptake[38], oligodendrocytic myelin formation[63], and microglial phagocytosis[40]. In Chapter 1, we illustrate how single-cell analysis of pa-

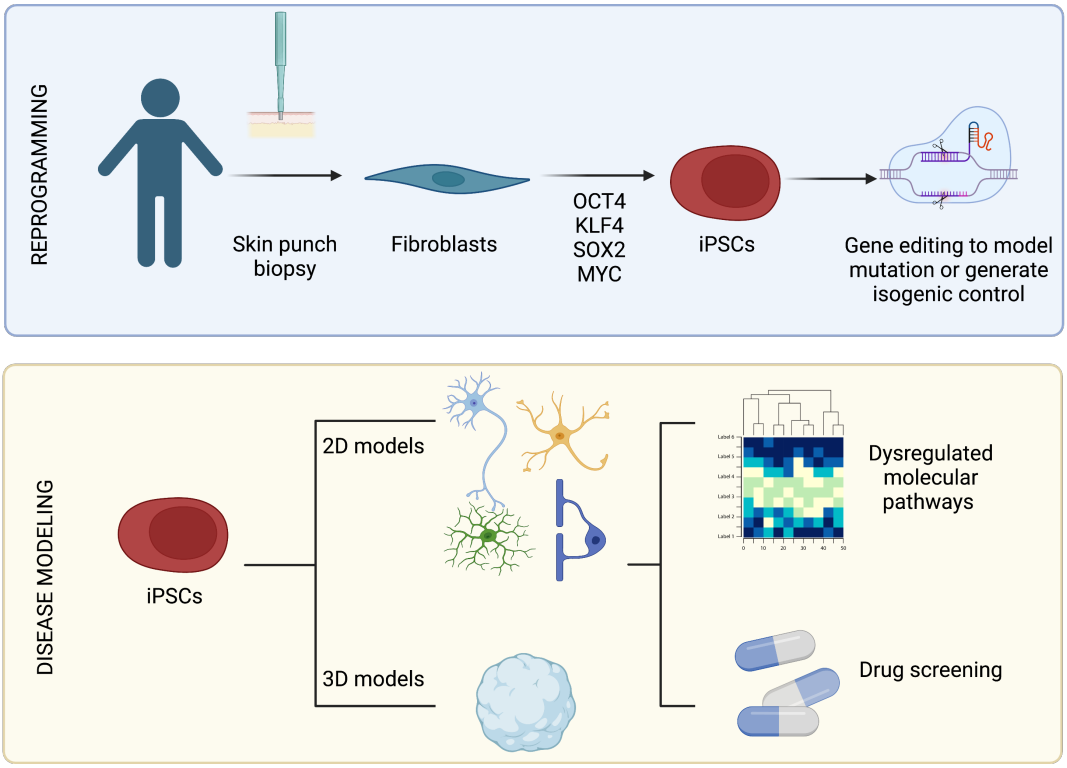


Figure 1: Generation of human iPSC-derived models of the brain. Human fibroblasts can be reprogrammed into iPSCs and then differentiated into 2D or 3D models of the brain to enable discovery of dysregulated molecular pathways or drug screening.

tient postmortem tissue can generate mechanistic hypotheses about cell type involvement in disease, which we then experimentally test using iPSC-derived models.

iPSC-derived models of the brain can provide insight on the molecular cascades that give rise to neurodevelopmental disorders by enabling the study of early development. iPSCs can be generated from patients to directly model the disease and link known genotypes to phenotypes. Advances in gene editing have further improved the utility of iPSC models by facilitating the creation of isogenic lines. Isogenic lines allow dissection of the effects of a genetic variant without potentially confounding differences in genetic background[9]. For example, CRISPR-Cas9 can be used to correct a pathogenic variant in a patient iPSC line and determine which phenotypes are rescued, and

therefore are likely caused by the variant. Alternatively, CRISPR-Cas9 can also be used to introduce pathogenic mutations into control iPSC lines. These isogenic systems enable well-controlled experiments to be designed to probe biological and pathological mechanisms and establish causality between genotype and phenotypes.

One of the first iPSC disease models was derived from a patient with spinal muscular atrophy (SMA)[60]. SMA is an autosomal recessive neurogenetic disorder caused by disruptions of the gene *SMN1* that leads to selective degeneration of lower motor neurons. This study demonstrated that iPSCs could be generated from patient dermal fibroblasts and differentiated into a motor neuron-like fate (marked by expression of choline acetyltransferase, *ChAT*). Compared to iPSC-derived neurons from the unaffected mother, the SMA neurons had a significant decline in the number and size of motor neurons by 6 weeks in culture, indicating that SMA iPSC-neurons recapitulate the disease phenotype of motor neuron degeneration[60]. Over the past decade, patient-derived iPSCs have been used to model many neurogenetic conditions including neurodevelopmental, neuropsychiatric, and neurodegenerative diseases. Fragile X syndrome (FXS) is one of most common inherited neurodevelopmental disorders caused by an expanded trinucleotide repeat (CGG) in the gene *FMR1*. iPSCs derived from multiple patients with FXS demonstrated impaired neuronal differentiation (decreased Tuj1 positive cells and neurite length) that correlated with CGG repeat length and epigenetic silencing of *FMR1*[238]. Rett syndrome is another common neurodevelopmental disorder in females caused by disruption of the X-linked gene *MECP2*[8]. Rett syndrome iPSCs also demonstrated defects in neuronal differentiation compared to controls, with decreased Tuj1-positive cells and decreased expression of neuronal maturation markers *PAX6* and *SCN1A*[170,129]. Patient-derived iPSCs also enable modeling of genetically complex or idiopathic neuropsychiatric conditions such as schizophrenia (SCZ) and idiopathic autism (ASD). iPSC-derived neurons from SCZ patients demonstrated decreased neuronal connectivity and number of neurites that could be reversed by treatment with the antipsychotic loxapine[30]. Idiopathic ASD iPSC-derived neural pro-

genitors demonstrated increased cell proliferation consistent with the macrocephalic phenotype in the ASD individuals, and their neurons had reduced synaptogenesis resulting in decreased neuronal network activity [169]. Finally, iPSC models have also been used to understand neurodegenerative diseases, including Alzheimer's disease [110], amyotrophic lateral sclerosis [230], and frontotemporal dementia [5]. These studies demonstrate that patient-derived iPSC models can illuminate molecular and cellular mechanisms of monogenic and genetically complex neurological disorders.

0.3 The rise of 3D organoid models

While iPSC-derived neurons and glial cells allow the study of a relatively homogeneous cell type, brain organoids are self-organized, 3D culture systems that can recapitulate additional features of the human brain, including diversity of cell types, structural organization, and developmental trajectories [140,204,274]. scRNA-seq of dorsal forebrain organoids revealed major cell types found in the developing human cortex, including progenitors, glutamatergic neurons, and astroglia [274]. Importantly, human iPSC-derived brain organoids can generate human-specific cell types, such as outer radial glia within progenitor zones [140]. The transcriptomes of cerebral organoids are also highly correlated with that of endogenous human fetal tissue and most closely resembles mid-gestation human cortex [274,24]. Organoids also temporally recapitulate developmental stages of neurogenesis, with the production of progenitors, followed by neurons resembling deep layer neurons, upper layer neurons, and finally astroglia [274,24,269,218]. Overall, organoids enable the study of multiple cell types in a disease, as well as non-cell autonomous disease phenotypes and processes.

In the past decade, several different methods for generating cerebral organoids have emerged [140,274,24,204]. Brain organoids can be broadly classified into "unguided" versus "guided". Unguided cerebral organoids spontaneously generate a high diversity of neural cell types that may reflect multiple anatomical regions such as retina, cortex, and cerebellum with variability between individual organoids [140,205].

Guided organoids are exposed to additional small molecules or factors that instruct the organoid to develop into a particular region of the nervous system, such as dorsal forebrain, ventral forebrain, hypothalamus, retina, or striatum [15,24,106,277,185]. Guided organoids tend to be more reproducible than unguided organoids in that individual organoids develop similar cell types at similar proportions [274]. Multiple organoids can also be combined to generate an "assembloid". Assembloids containing different regionalized organoids allow the modeling of processes such as circuit assembly and cell migration. For example, fused cortical and thalamic organoids developed reciprocal axon projections, resembling the human thalamus and cortex [287]. In other studies, assembly of dorsal and ventral forebrain organoids modeled the migration of interneurons produced in ventral forebrain into dorsal forebrain [15,24]. Assembloids can also be generated from organoids from different individuals or different species to study individual variation or species-specific phenotypes [205].

Remarkably, organoids can be transplanted *in vivo* into animals to model disease in a physiological context [167,218,205]. One of the first studies to transplant a human brain organoid into adult mouse brain demonstrated that the grafted organoid developed axon projections throughout the host brain, became vascularized by the host and infiltrated by host microglia, and functionally integrated into the host synaptic circuits [167]. Organoids have also been transplanted into newborn rat cortex when neuronal circuits are still under development [218]. The grafted organoids integrated into circuits with native rat neurons and optogenetic stimulation of human neurons could modulate the rat behavioral response in a reward-training task [218]. Transplanted organoid models have the advantages of modeling human neurological disease under physiological conditions to study circuit level phenotypes such as behavior.

0.4 Advantages and limitations of iPSC models

Each iPSC model system has unique characteristics that may serve as advantages or limitations depending on the scientific question of interest (Table 1). While methods for developing several types of neurons, glia, and region-specific cerebral organoids have been established, there are still cell types that have been relatively challenging to differentiate from iPSCs. For example, cerebellar organoids have taken longer to develop, although a very recent study was able to generate organoids containing cells resembling mature Purkinje and granule neurons[41]. In addition, differentiated cell types tend to be relatively immature and are more transcriptionally similar to human fetal cell types than adult cell types. iPSCs derived from patients and individuals are also genetically heterogeneous which may introduce variability in observations and hinder reproducibility[9]. Best practices recommended in the field[9] include the use of isogenic controls when possible, multiple iPSC-lines and clones. In order to compare diseases with controls, the same techniques for reprogramming and differentiation should be applied, and disease and control lines should be closely matched in terms of donor characteristics such as age and sex.

Another limitation of current organoid methods is the lack of intrinsic formation of non-neural lineage cell types such as microglia and endothelial cells. Microglia are the key brain-resident immune cells and are important for synaptic pruning, homeostatic maintenance, and response to injury[115]. While organoids do not naturally develop microglia, iPSC-derived microglia can invade and integrate organoids when co-cultured[212]. Vascularization is important for the delivery of oxygen and nutrients throughout the organoid and would enable longer-term culture that is currently limited by formation of necrotic centers due to the lack of oxygen and nutrient penetration[37]. One study demonstrated that induced expression of ETV2 could promote formation of a vascular-like network in human cortical organoids[37], and development of similar approaches for nutrient delivery and improved organoid viability are on-going.

Table 1: Advantages and limitations of iPSC-models

	2D iPSC-models	3D organoids	Transplanted organoids
Cell type diversity	Relatively homogeneous	Recapitulates diversity of human brain but lacks non-neural lineages	Infiltration of microglia and vascularization
Cell-cell interactions	Requires co-culture to study interaction between different cell types	Interactions in self-organized 3D environment	Can study interactions between graft and host brain
Cytoarchitecture	None	Resembles human brain (forms oSVZ)	Improved maturation and cytoarchitecture
Environment	*in vitro*	*in vitro*	*in vivo*
Cost	Relatively inexpensive	More expensive	Most expensive

0.5 Insights from functional genomics of human and iPSC models of neurogenetic disorders

Functional genomics involves genome-wide studies to understand gene functions and genotype-phenotype relationships[95]. Advances in multi-omic technologies can be applied to human brain tissue and iPSC-derived models to enable unbiased characterization of chromatin state, epigenetic signatures, DNA variants, and transcriptomic profiles of neurological diseases. For example, ATAC-seq is a method for mapping genome-wide chromatin accessibility, which can highlight genomic regions containing permissible or active regulatory elements[34]. ChIP-seq and CUT&RUN-seq are approaches for profiling genome-wide epigenetic marks and protein-DNA interactions, such as transcription factor binding sites[116,20,243]. Whole genome sequencing enables detection of genetic variants, including single nucleotide variants, insertions and deletions, and copy number variations[22]. RNA-seq characterizes the transcriptome of a biological specimen, which can vary for different tissues, developmental stages, or physiological conditions[279]. The recent development of single-cell and spatial technologies provides higher resolution characterization of cell type diversity, cell-cell interactions, and cell type specific alterations in disease[226,273]. In Chapter 1 and Chapter 3, we combine each of these functional genomic approaches ("multi-omic analysis") to understand neurodevelopment and degeneration in monogenic disorders.

The application of these functional genomic approaches to human tissue and model systems has begun to provide mechanistic insights on disease pathogenesis for neurodevelopmental disorders. Characterization of genome-wide histone acetylation in brains from individuals with ASD compared to controls revealed regions with abnormal acetylation near genes involved in synaptic transmission [256]. Transcription factor binding motifs within these differentially acetylated regions were enriched for RFX, PAR bZIP, AP-1, and MEF2 motifs, indicating that these transcription factors may either mediate or be impacted by aberrant histone acetylation in ASD [256]. RNA-seq analysis of postmortem brain from over 100 individuals with ASD and controls identified broad transcriptomic alterations across multiple cortical regions in ASD that converged on downregulation of genes involved in synaptic signaling [72]. Dysregulation of specific cell types in ASD has also been illuminated by snRNA-seq of cortical tissue from individuals with ASD compared to controls [275]. Characterization of cell type specific transcriptomic changes revealed that upper layer excitatory neurons were preferentially affected in ASD, and gene expression dysregulation in these neurons correlated with clinical severity [275]. Taken together, these studies of brain tissue from individuals with ASD implicate dysfunction of synaptic transmission and upper layer cortical neurons in ASD pathogenesis.

Functional genomic characterization of iPSC-derived neurons and organoid models of ASD offers the advantage of testing mechanistic hypotheses over postmortem tissue. One of the first studies to generate organoids from iPSCs from individuals with severe idiopathic ASD used transcriptome analyses to identify upregulation of genes involved in cell proliferation that was associated with accelerated cell cycle and overproduction of inhibitory neurons [171]. It was demonstrated that this phenotype was dependent on upregulation of *FOXG1*, and the degree of *FOXG1* upregulation also correlated with symptom severity [171]. Another study used gene editing to generate cortical organoids with heterozygous loss-of-function variants in three highly penetrant ASD risk genes (*SUV420H1*, *ARID1B*, and *CHD8*). scRNA-seq analysis of these three models of monogenic ASD

revealed convergent phenotypes involving accelerated development of inhibitory neurons[203]. These studies demonstrate the power of functional genomic characterization of iPSC-based models for investigating neurodevelopmental alterations in ASD. In Chapter 3, we use this approach to perform a functional genomic dissection of iPSC-derived neurons and organoid models to study the basic function of another highly penetrant ASD risk gene, *RFX3*, in human neurodevelopment.

0.6 Human based disease modeling for drug discovery

iPSC-based models serve as a platform that can be scaled for high throughput drug screening. Identifying compounds that modulate disease phenotypes in iPSC models may highlight specific molecular pathways as novel targets for disease-modifying therapies. In one study, compounds were screened in iPSC-derived neurons from a patient with a neurodevelopmental disorder caused by *SHANK3* haploinsufficiency, to identify candidates that could increase *SHANK3* expression and reverse disease phenotypes[48]. This led to the identification of lithium and valproic acid as candidate molecules with disease modifying potential. iPSCs can also be used to stratify patients into clinical cohorts based on their response to certain treatments. In a study of iPSC-derived neurons from patients with bipolar disorder, neuronal phenotypes segregated in a manner that correlated with clinical response to lithium[250]. In addition to screens for therapeutic efficacy, iPSC models can be leveraged to monitor and predict drug toxicity[104]. One study differentiated neurons from iPSCs from a population of individuals and identified 22 hit compounds with neurotoxicity *in vitro*, two of which were confirmed *in vivo* in mouse models[104]. Therefore, iPSCs can facilitate high throughput screening for drug efficacy and toxicity as a cost-effective step prior to preclinical animal studies.

iPSC models can also enable development of gene-based therapies. Antisense oligonucleotides (ASOs) are short, single-stranded nucleotides that modulate gene expression by complementary binding to mRNA transcripts[219]. One major class of ASOs are splice modulating ASOs. ASOs can

modulate splicing by blocking splicing enhancer or silencer cis-regulatory elements [102]. This makes them a potential therapeutic approach for correcting pathogenic splicing variants or regulating gene expression through exon inclusion or skipping. However, alternative splicing patterns can vary between cell types and species [200,198], making it difficult to understand the impact of a patient mutation on splicing patterns if using more accessible cell types such as patient fibroblasts or animal models. Differentiating patient iPSCs into the most relevant cell type allows us to identify the splicing patterns that can be targeted by an ASO, enabling ASO design and screening in models of relatively inaccessible tissues such as the brain. In Chapter 4, we demonstrate how ASOs can be designed and tested in iPSC-derived neurons to restore the expression of genes disrupted in neurodevelopmental disorders.

0.7 Concluding remarks

In conclusion, patient tissues and gene-edited iPSC models of the brain are complementary systems that together have the ability to capture molecular, cellular, and functional phenotypes of neurogenetic disorders. This thesis illustrates how these models, combined with functional genomics, present unique opportunities to investigate underlying mechanisms and enable development of novel therapies for human disease.

"Alone we are rare. Together we are strong."

A-T Children's Project

1

ATM-deficiency induced microglial activation promotes neurodegeneration in Ataxia-Telangiectasia

This section is based on the following manuscript in which I made significant contributions during my PhD studies:

Lai J, Demirbas D, Kim J, Jeffries AM, Tolles A, Park J, Chittenden TW, Buckley PG, Yu TW, Lodato MA, Lee EA. *ATM*-deficiency induced microglial activation promotes neurodegeneration in Ataxia-Telangiectasia. bioRxiv. 2023. doi: https://doi.org/10.1101/2021.09.09.459619. **Supplementary Material:** All supplementary material can be found in the supplement of Lai *et al*., bioRxiv 2023.

1.1 Introduction

Selective degeneration of the cerebellum occurs in genetic ataxias such as Ataxia Telangiectasia (A-T). However, the mechanisms underlying cerebellar degeneration remain poorly understood. A-T is an autosomal recessive multi-system disorder caused by mutations in the gene *ATM*[234]. It affects as many as 1 in 40,000 live births and typically presents in early childhood with a median age of diagnosis of 6 years)[206]. Clinical features of A-T include progressive impairment of gait and coordination due to cerebellar neurodegeneration, immunodeficiency, and increased predisposition for

cancers. There are currently no therapies to slow neurodegeneration in A-T[206].

ATM is a multifunctional kinase known for its role in DNA damage response to double strand breaks (DSB)[234,143]. DNA damage leads to *ATM* activation, which phosphorylates downstream regulators of cell cycle arrest, DNA repair, and apoptosis[280]. Additionally, *ATM* plays a role in cellular metabolism. *ATM* can form a dimer that mediates mitochondrial redox sensing and regulates antioxidant capacity[298]. Moreover, *ATM* deficiency leads to upregulation of autophagy and perinuclear accumulation of lysosomes, implicating *ATM* in lysosomal trafficking[42].

While *ATM* has roles in diverse cellular functions, it remains unclear how *ATM* loss-of-function leads to selective and progressive degeneration of Purkinje and granule neurons in the cerebellum. *Atm*-null mice do not recapitulate the cerebellar degeneration found in human A-T patients[144], but bulk transcriptomic analyses of human brain tissue and neuronal models have begun to offer some insights into dysregulated pathways in A-T. Expression of *ITPR1*, a calcium channel that is highly expressed in Purkinje cells and is associated with spinocerebellar ataxias, is significantly altered in A-T cerebellum[113]. Transcriptome analyses of induced pluripotent stem cell-derived cerebellar-like neurons from individuals with A-T and controls have also revealed alterations in pathways related to synaptic vesicle cycling, oxidative stress, and insulin secretion[193]. However, it remains unresolved whether bulk transcriptomic changes reflect loss of certain cell types in A-T, or perturbation of cellular functions in specific cell types.

Emerging evidence implicates dysfunctional microglia in the pathogenesis of A-T. Human cellular models of microglia with *ATM* deficiency reveal dysfunctional phagocytosis of neuronal processes[27]. However, whether microglia activation is present in A-T patient brains, whether the microglia are reactive to cell-intrinsic or extrinsic signals, and their role in neurodegeneration remain unknown. Characterization of microglia and their transcriptional signatures in human A-T brain may provide further insight on mechanisms that underlie cerebellar degeneration in A-T.

Single-cell technology has provided insight on the cell-type-specific effects of neurological dis-

eases including Alzheimer's disease (AD), Huntington's disease, Multiple Sclerosis, Autism, and Major Depressive Disorder [175,235,275,145,190]. However, data from human adult cerebellum in health and degeneration has not yet been explored. Here, we present the largest single-nucleus transcriptomic atlas of adult human cerebellar vermis and the first atlas of cerebellar degeneration to-date. We profiled 126,356 nuclei from the postmortem human cerebellar vermis of six individuals with genetically confirmed A-T and seven matched control individuals. In addition, we profiled 86,354 nuclei from postmortem human prefrontal cortex (PFC) from two A-T and two matched control individuals. We annotated major cell types in the human cerebellum and PFC, and identified cell type proportion changes between A-T and controls. We also demonstrated cell-type-specific expression of cerebellar ataxia-associated disease genes, some of which were significantly dysregulated in A-T Purkinje neurons, supporting a convergence in disease pathophysiology underlying several hereditary ataxias. We analyzed the cell-type-specific molecular pathways perturbed in A-T across these brain regions, revealing prominent and widespread activation of pro-inflammatory pathways in A-T microglia. Pseudotime analysis revealed that activation of A-T microglia preceded upregulation of apoptosis related genes in granule and Purkinje neurons, while ligand-receptor analysis suggested that microglia have increased neurotoxic cytokine signaling to granule and Purkinje neurons in A-T. We experimentally interrogated the role of microglia in neurodegeneration using A-T patient and control iPSC-derived microglia (iMGL) and neurons (iN). Transcriptomic profiling of A-T patient iMGL revealed cell-intrinsic microglial activation of pro-inflammatory pathways, and A-T patient iMGL were sufficient to induce elevated cytotoxicity of co-cultures with control and A-T iNs. Overall, our data suggests that activated microglia are central to A-T pathophysiology and cerebellar degeneration.

1.2 Results

1.2.1 snRNA-seq profiling of A-T human cerebellum and prefrontal cortex

A-T presents with prominent loss of motor coordination associated with early selective atrophy of the cerebellum. To investigate cell type and region-specific perturbations in A-T, we performed single-nucleus RNA-sequencing (snRNA-seq) of human postmortem tissues from the cerebellar vermis (the region with the most prominent atrophy in A-T[261]) and prefrontal cortex (PFC) (Figure 1.1A). We profiled six cerebella from neuropathologically and genetically confirmed A-T cases and seven control cerebella, as well as matched A-T and control PFC from two cases each (Table S1.1-S1.4). The ages of the A-T cases ranged from 19 to 50 years old (mean age 28 years). Biallelic *ATM* variants (with evidence of pathogenicity in ClinVar[141]) were confirmed in all six A-T cases by whole genome sequencing (WGS) (Table S1.5, Figure A.2). All cases were balanced for age, sex, RNA integrity number (RIN), and postmortem interval (PMI) (Tables S1.1-S1.2). After quality control and doublet filtering[253,178,70], there were 126,356 cerebellar nuclei (51,297 A-T; 75,059 control) and 86,354 PFC nuclei (16,979 A-T; 69,375 control) for downstream analysis (Figure A.2).

1.2.2 Single-nucleus RNA-seq recapitulates neuropathological hallmarks and reveals increased glial populations in A-T cerebellum

To distinguish cell types in the cerebellum, we used Seurat[253] to perform graph-based clustering with the Leiden algorithm and annotated clusters based on the expression of canonical marker genes and existing single-cell atlases of the cerebellum[139,134]. Ten major cell types were identified, including the cerebellum specific granule cells (*RIMS1, GRM4*), Purkinje cells (*ITPR1, CALB1*), and Bergmann-glia (*TUBB2B, AQP4*), as well as interneurons (*GAD1, PVALB*), astrocytes (*TTN, AQP1*), oligodendrocytes (*PLP1, MBP*), oligodendrocyte precursor cells (OPC) (*PDGFRA, OLIG1*),

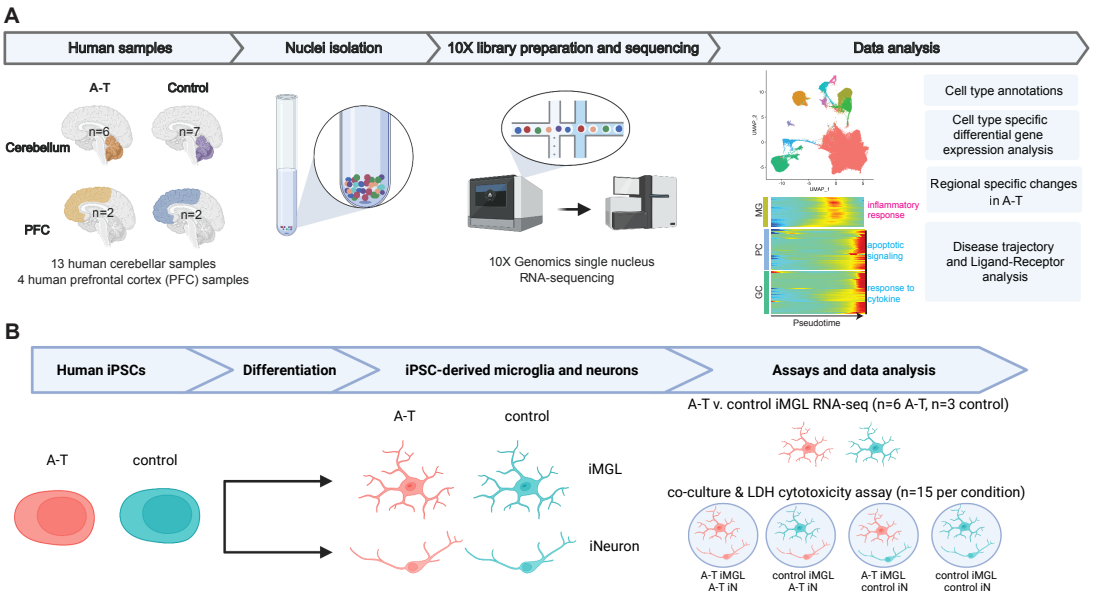


Figure 1.1: Dissecting cell type specific contributions to neurodegeneration in Ataxia Telangiectasia using single-nucleus RNA-sequencing of postmortem brain and patient iPSC-derived cultures. A. Schematic overview of snRNA-seq experimental design and computational analyses. B. Schematic overview of A-T patient and control human iPSC-derived experimental design. N represents the number of independent culture wells. MG, microglia. PC, Purkinje cells. GC, granule cells. iMGL, iPSC-derived microglia. iN, iPSC-derived neurons.

microglia (*CD74, CSF1R*), endothelial cells (*CLDN5, VWF*), and fibroblasts (*DCN, APOD*) (Figure 1.2A-B). We annotated PFC cell types by reference-based mapping with SingleR, using an existing human PFC snRNA-seq dataset as the reference[275,10]. There were thirteen cell types identified, including astrocytes, microglia, endothelial cells, oligodendrocytes, OPCs, and multiple subtypes of excitatory and inhibitory neurons (Figure A.3A).

Next, to understand the patterns of cerebellar cell loss in A-T, we examined the cell type proportions in A-T compared to controls. Due to the negative covariance structure of compositional data—as one feature increases, the other features must decrease—we implemented Dirichlet multinomial modeling using Hamiltonian Monte Carlo (DMM-HMC) to estimate the cell type proportions in A-T and control[86]. We generated multinomial distributions of cell type proportions for each sample based on the nuclei counts (Table S1.3), and used these to construct a Dirichlet distribution of cell type proportions in A-T and controls. To determine the relative shift in abundance of each cell type in A-T versus control, we subtracted the control posterior probability distribution of proportions from the A-T distribution (Figure 1.2D). Granule cells were significantly decreased in A-T (Figure 1.2C-D, 89% credibility interval). Purkinje cells were reduced in abundance, although not significantly, possibly due to individual sample variability and decreased power for detecting changes in rare cell types. In addition, astrocytes, Bergmann-glia, microglia, and oligodendrocytes showed significant increases in A-T compared to control (Figure 1.2D, 89% credibility interval). In contrast, permutation testing of disease labels (A-T versus control) showed no significant differences for any cell type, as expected (permutation n=500, Figure A.2E). Unlike A-T cerebellum, there was no evidence of neuronal loss in A-T PFC (Figure A.3B). However, oligodendrocytes were significantly decreased in A-T PFC compared to control (Figure A.3B), consistent with reports of cortical white matter degeneration in A-T[224,267]. Taken together, these results are consistent with the neuropathology autopsy reports from individuals with A-T, which describe granule and Purkinje cell atrophy and gliosis in the cerebellum while the PFC is grossly unremarkable (Table S1.4).

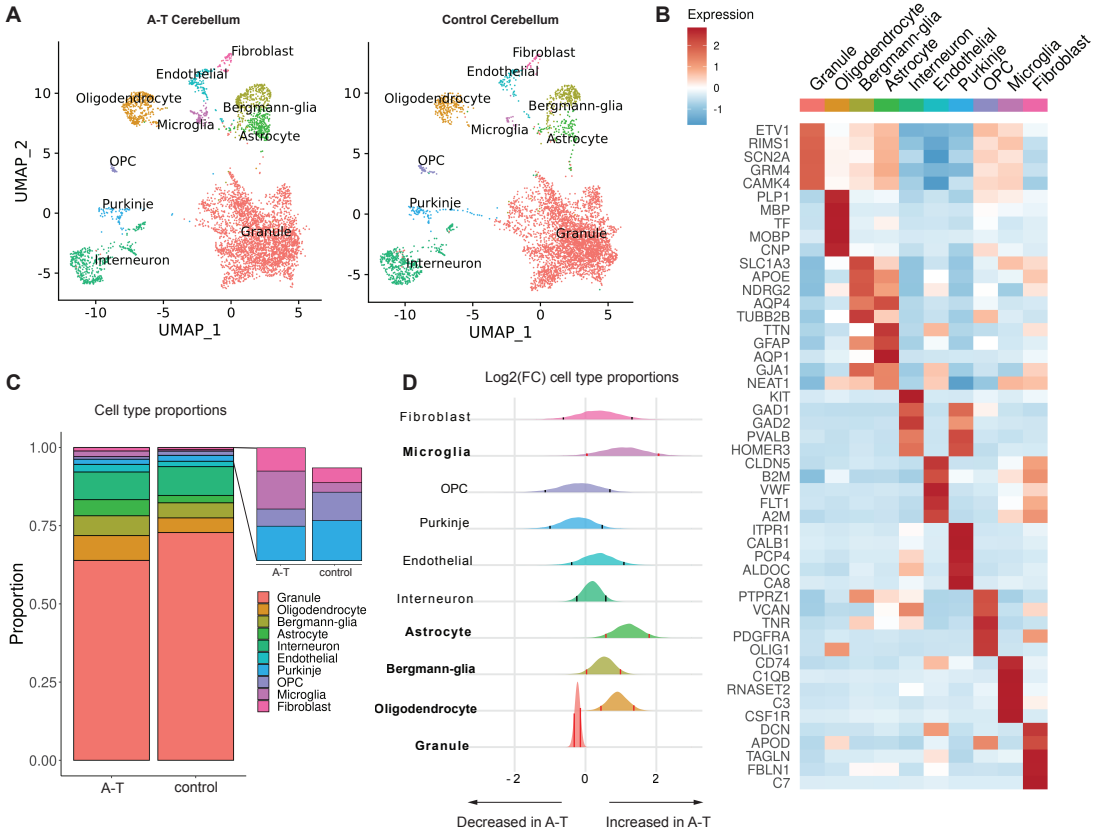


Figure 1.2: Cell types resolved in A-T and control human cerebellum by snRNA-seq. A. UMAP dimensionality reduction plot of major cell types identified in A-T and control human cerebellum, downsampled to 10,000 cells for each condition. B. Top five marker genes for each major cell type in control cerebellum. Heatmap depicts centered and scaled log-normalized expression values. C. Cell type proportions in A-T and control cerebellum. D. Relative abundance of cell types in AT versus control cerebellum, shown as the posterior distribution of Log2(proportion in AT/proportion in control) with 89% credible interval. Red bars highlight credible intervals that do not overlap 0. Bolded cell type labels indicate a significant difference in relative abundance.

1.2.3 Monogenic cerebellar disease genes are expressed in specific cell types

Next, we characterized the expression of *ATM* across cell types in the cerebellum and PFC. *ATM* had the highest expression in microglia in adult control human cerebellum and PFC (Figure A.4A). To investigate human developmental expression patterns, we profiled the expression of *ATM* in a snRNA-seq dataset of human control fetal cerebellum27, which also revealed the highest expression of *ATM* in microglia (Figure A.4B). Taken together, these data suggest an important role for *ATM* in microglia during normal cerebellar development.

Given that *ATM* is enriched in microglia, we asked whether additional hereditary ataxia genes are enriched in specific cerebellar cell types. We curated a list of genes implicated in human hereditary ataxias using the Online Mendelian Inheritance in Man (OMIM) database (Table S1.6) and profiled their expression in control cells. We found that like *ATM*, most hereditary ataxia disease genes demonstrate cell-type-specific patterns of expression (Figure A.4C; Bonferroni adjusted p-value < 0.05, t-test). Cerebellar disease gene expression was most commonly enriched in Purkinje cells, including the potassium and calcium ion related genes *KCND3, PRKCG, CACNA1G, ITPR1, TRPC3,* and *KCNC3*. Microglia also showed enriched expression of several ataxia disease genes (*ATM, TPP1, SETX,* and *VPS13D*) (Figure A.4C). These gene expression patterns implicate Purkinje cells and microglia in the pathogenesis of cerebellar ataxias.

1.2.4 Transcriptional dysregulation across cell types in A-T

To investigate the cell types and molecular pathways that are most perturbed in A-T cerebellum, we performed differential gene expression analysis between A-T and control for each cell type (Table S1.7). Astrocytes, microglia, and oligodendrocytes had remarkably high numbers of DEGs (n=1,631, 1,577, 1,563 respectively) while granule cells had the least (n=265). These data suggest that while granule cells degenerate in A-T, their gene expression program is less dysregulated than

other cerebellar cell types such as glia.

Next, we performed Gene Ontology (GO) analysis on the DEGs found in each cerebellar cell type to gain insight into the biological processes dysregulated in A-T. Consistent with the loss of Purkinje and granule cells in A-T, both cell types had significant upregulation of genes involved in apoptotic signaling (Figure 1.3A-B, Figure A.5A, Table S1.8). A-T Purkinje and granule cells also upregulated genes involved in ribosome assembly and biogenesis, which has been observed in several studies of aging brains[296,288]. A-T microglia specifically demonstrated upregulation of immune response related genes involved in microglial cell activation, phagocytosis, and cytokine production (Figure 1.3C).

Activated microglia associate with sites of injury and pathological changes across various neurodegenerative diseases[245] and can induce local neurodegeneration by phagocytosis of synapses and secretion of neurotoxic cytokines[192]. We thus sought to characterize the spatial distribution and number of microglia in A-T and control cerebellum by immunostaining for CD11B, a microglia marker that is upregulated in activated microglia[222], in four A-T cerebellum and three matched control cerebellum samples (Table S1.9). In A-T cerebellum, there were increased CD11B+ cells throughout the white matter, granule layer, and molecular layer compared to control cerebellum (Figure 1.3D). Across all layers, the number of CD11B+ microglia per area was significantly higher in A-T compared to control (Bonferroni adjusted p-values< 1e-04, Mann-Whitney U test) (Figure 1.3D). These data confirm that A-T cerebellum has a widespread increased presence of activated microglia.

1.2.5 Dysregulated calcium signaling in Purkinje neurons in A-T

GO analysis of genes downregulated in A-T revealed significant enrichment of calcium related pathways (calcium ion transport, regulation of cytosolic calcium ion concentration, cellular calcium ion homeostasis) in A-T Purkinje cells (Figure 1.3A, Figure A.5, Table S1.10). Among these genes was

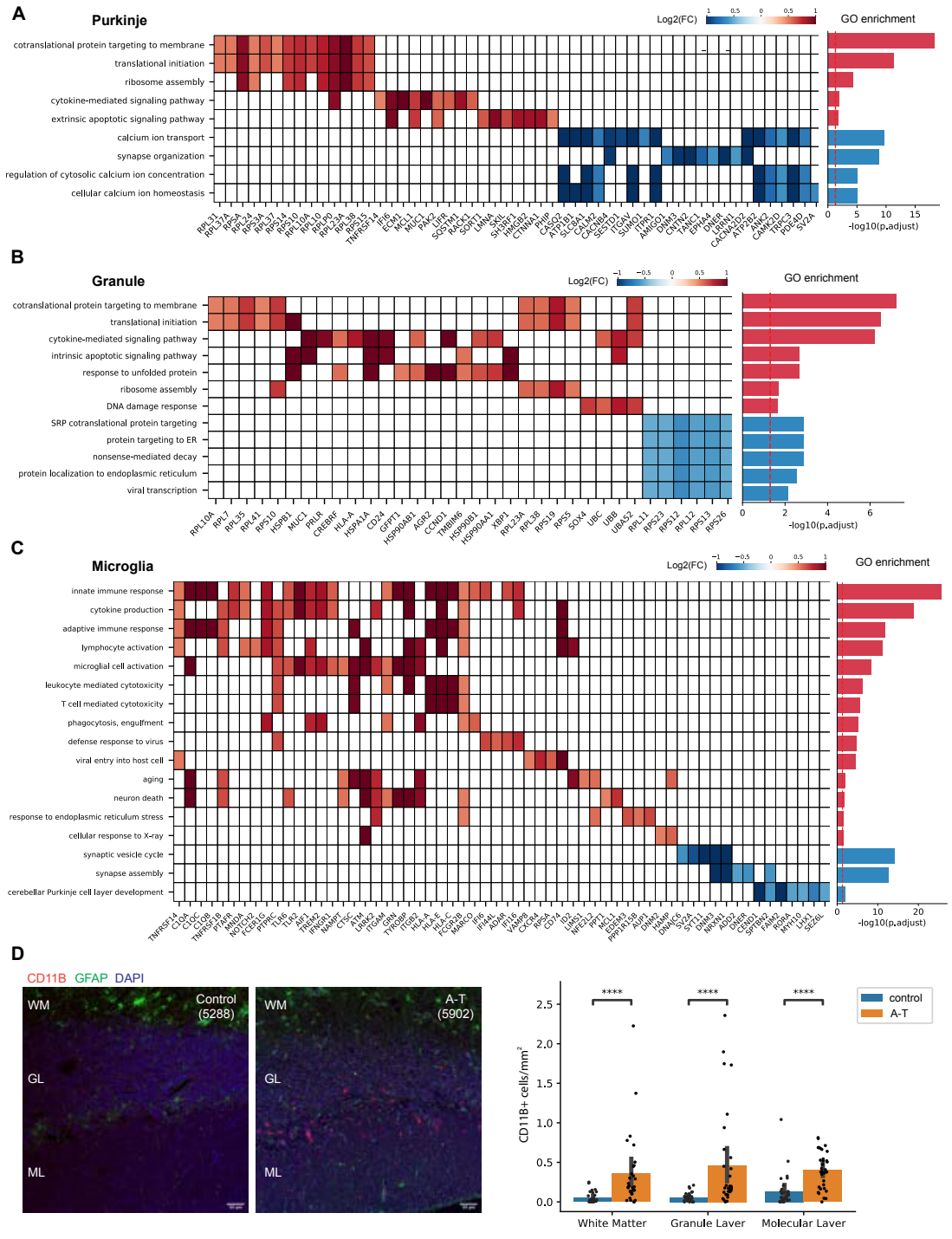


Figure 1.3: Gene Ontology (GO) analysis of differentially expressed genes in A-T cerebellum.

Figure 1.3 (Continued). A-C. Select enriched GO biological processes among DEGs in A) A-T Purkinje neurons, B) granule neurons, C) microglia. Select genes in each pathway and their log2 fold-changes (A-T/control) shown in the heatmap. Barplot shows significance of pathway enrichment among upregulated genes (red) and downregulated genes (blue). D. CD11B immunostaining of postmortem cerebellar cortex from A-T and control. WM: white matter. GL: granule layer. ML: molecular layer. Quantification of CD11B-positive cells across cerebellar cortex layers in AT and control. Error bar represents 95% confidence interval. ****Bonferroni-adjusted p-value<1e-04, Mann-Whitney U test.

ITPR1, an endoplasmic reticulum (ER) calcium channel that causes spinocerebellar ataxias in an autosomal dominant manner[272,105] (Figure 1.3A). We asked whether additional cerebellar disease genes are differentially expressed in A-T. Hereditary ataxia genes with enriched expression in Purkinje cells (i.e., *KCNC3, PRKCG, ITPR1, KCND3, CACNA1G*) were significantly downregulated in A-T Purkinje cells (Figure A.7A). These disease genes were also enriched for calcium processes (calcium ion transport and regulation of cytosolic calcium ion concentration) (Figure A.7B). These results suggest that ATM may be an upstream activator of calcium signaling and points to converging underlying mechanisms of a subset of hereditary ataxias in Purkinje cells. *HDAC4* is a known negative regulator of calcium ion genes (*ITPR1, DAB1*)[229], and *ATM* is an inhibitor of *HDAC4*. Loss of *ATM* leads to nuclear accumulation of *HDAC4*, which could lead to dysregulated expression of *HDAC4* target genes[149]. In support of this, there was significant overlap between *HDAC4* targets and A-T downregulated genes (p-value 6.08e-22, Fisher's exact test, Figure A.7C). Taken together, these findings support that abnormal calcium ion homeostasis may be a consequence of *ATM* loss of function in Purkinje neurons, which may lead to aberrant firing patterns and ataxic phenotypes[87,96].

1.2.6 Activated microglia in A-T cerebellum share transcriptomic signatures with aging and neurodegeneration-associated microglia

Microglia in A-T cerebellum express activation markers, including strong upregulation of the complement genes, *C1QA, C1QB, C1QC,* and *C3* (Figure 1.4A). Increased expression of complement

is associated with microglial activation in neurodegenerative diseases such as Alzheimer's disease (AD)[285], where they have been shown to mediate synaptic loss and dysfunction[112,98]. To assess whether A-T cerebellar microglia share additional features associated with microglia in aging and other neurodegenerative disorders, we compared the upregulated genes in A-T cerebellar microglia with three gene sets: markers of aged human microglia, AD associated microglia markers from a human snRNA-seq study of AD, and markers of disease associated microglia (DAM) from a murine model of AD[175,199,123]. We found significant overlaps between A-T cerebellar microglia upregulated genes and these three gene sets with the highest overlap with aging microglia (Fisher's exact test p-values < 1e-13; Figure 1.4B-C). *TREM2, TSPO, MS4A6A*, complement components, and MHC class II genes were among the overlapping genes, which have been associated with neuroinflammation and microglia activation[59] (Figure 1.4C). GO enrichment analysis of the overlapping genes and A-T microglia specific upregulated genes revealed common enrichment in immune response and RNA processing related pathways, and A-T specific enrichment of cytoskeletal organization, ER stress, and regulation of neuron projection development (Figure 1.4D). Overall, these data show that A-T cerebellar microglia share transcriptomic signatures found in aging and other neurodegenerative diseases.

1.2.7 Differential activation of microglia in A-T cerebellum versus PFC

To systematically compare gene expression changes in A-T cerebellum with A-T PFC, we identified genes and pathways that are more significantly changed in A-T cerebellum than A-T PFC in cell types common to the two brain regions (microglia, OPC, oligodendrocytes, astrocytes, endothelial cells) (Figure A.8, Tables S1.11-1.16). While both cerebellar and PFC microglia showed upregulation of genes involved in immune response in A-T, microglia were more strongly activated in the cerebellum compared to the PFC. Cerebellar A-T microglia had significantly higher upregulation of genes related to inflammatory processes (adaptive immune response, leukocyte proliferation, glial

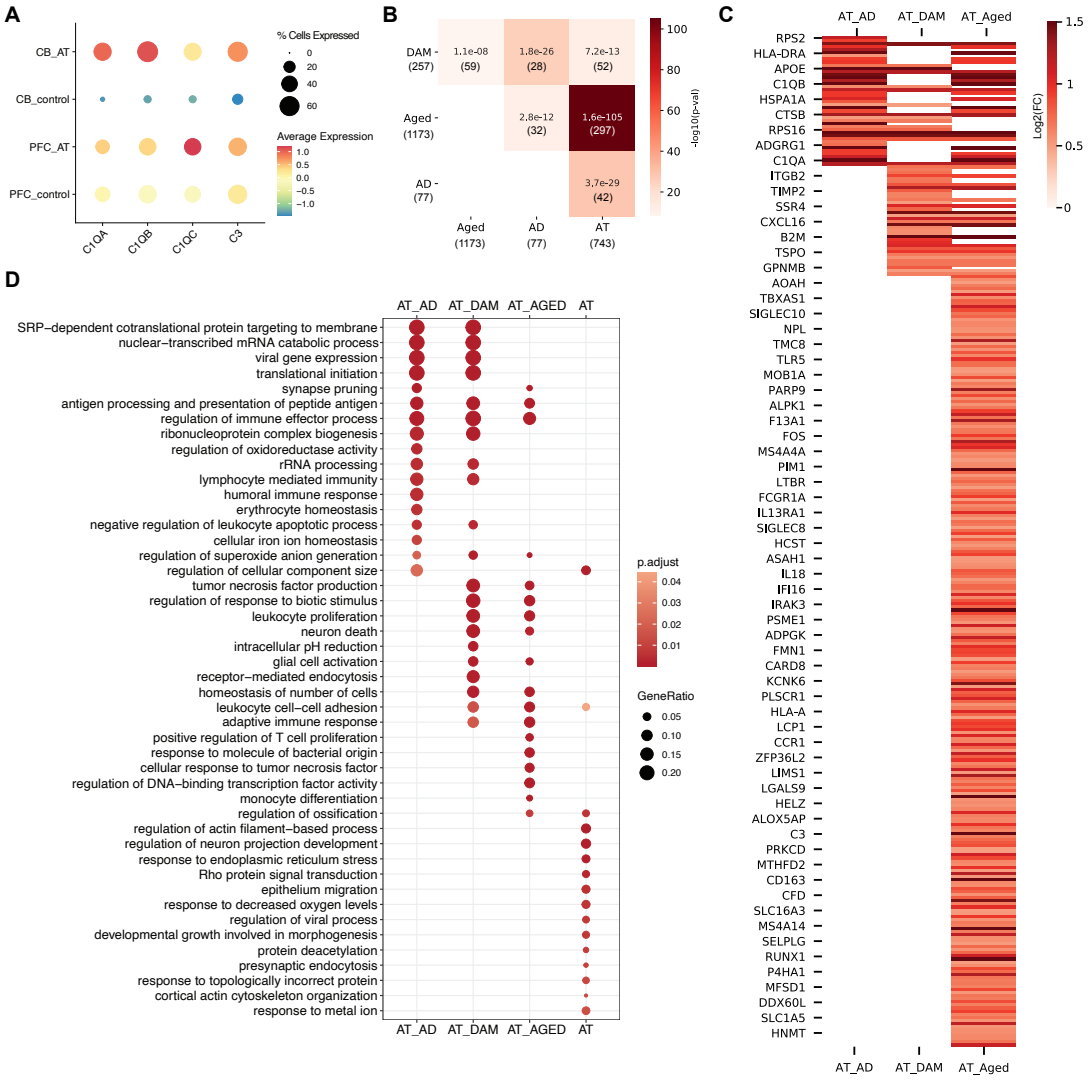


Figure 1.4: A-T cerebellar microglia share transcriptomic signatures with aging and neurodegenerative microglia. A. Average scaled expression of complement components C1QA, C1QB, C1QC, C3 in microglia from A-T and control cerebellum (CB) and prefrontal cortex (PFC). B. Overlap between A-T microglia upregulated genes and human Alzheimer's disease microglia markers (AD, Mathys et al., 2019), Disease Associated Microglia markers (DAM, Keren-Shaul et al., 2017), and human aging microglia markers (Aged, Olah et al., 2018). Overlap p-values from Fisher's exact test shown in each cell. Color represents -log10(overlap p-value). Number of genes in each set and intersection shown in parentheses. C. Heatmap of A-T microglia log2 fold-changes (FDR<0.05) for overlapping microglia markers. AT_AD: A-T and Alzheimer's disease microglia overlapping genes. AT_DAM: A-T and DAM overlapping genes. AT_AGED: A-T and human aging microglia overlapping genes. D. GO biological process enrichment of overlapping microglia markers. AT: A-T microglia only upregulated genes. Pathways with FDR<0.05 shown.

cell activation, tumor necrosis factor superfamily cytokine production) and neuronal death than A-T PFC microglia (Figure 1.5A, Figure A.7).

Microglia can be activated by neuronal death, a mechanism common to multiple neurodegenerative conditions[173]. However, since PFC neurons do not degenerate in A-T, we hypothesized that A-T microglia in the PFC and cerebellum are cell-intrinsically activated. One potential mechanism of intrinsic activation is the accumulation of DNA damage due to *ATM* deficiency. Deficits in DNA repair caused by *ATM* deficiency are associated with increased cytosolic DNA[109]. Cytosolic DNA is recognized by the CGAS-STING pathway, which triggers an innate immune response by inducing type I interferons and inflammatory cytokines46. We found that the CGAS-STING pathway, including *CGAS* and activators of CGAS signaling—*TRIM14, TRIM38,* and TRIM56[39,236,101]—were upregulated specifically in A-T microglia from the cerebellum and PFC, suggesting the presence of cytosolic DNA (Figure 1.5B-C). Moreover, the CGAS-STING pathway showed significant enrichment among genes upregulated in A-T cerebellum versus A-T PFC, suggesting that CGAS-STING activation was more prominent in cerebellar microglia (Figure 1.5C). CGAS-STING pathway genes (*TRIM14, TRIM38, IFI16, IRF7*) were also upregulated in aged human microglia but not AD or DAM gene sets, implicating DNA damage and cytoplasmic DNA as a common underlying cause of microglial activation in A-T and aging microglia, while AD microglia may be activated by another mechanism such as extrinsic stimuli (Table S1.17). Taken together, activation of the CGAS-STING pathway suggests that cytosolic DNA contributes to microglial activation that is more pronounced in AT cerebellum than PFC.

Self-DNA enters the cytosol in a mitosis dependent manner, through the formation and rupture of DNA-damage induced micronuclei[83,164]. Thus, we hypothesized that underlying differences in proliferation may explain the differential activation of CGAS-STING in cerebellar microglia compared to PFC microglia. We tested whether microglia in the cerebellum and PFC differ in proliferation rate by inferring the cell cycle status of each cell by the expression of canonical markers

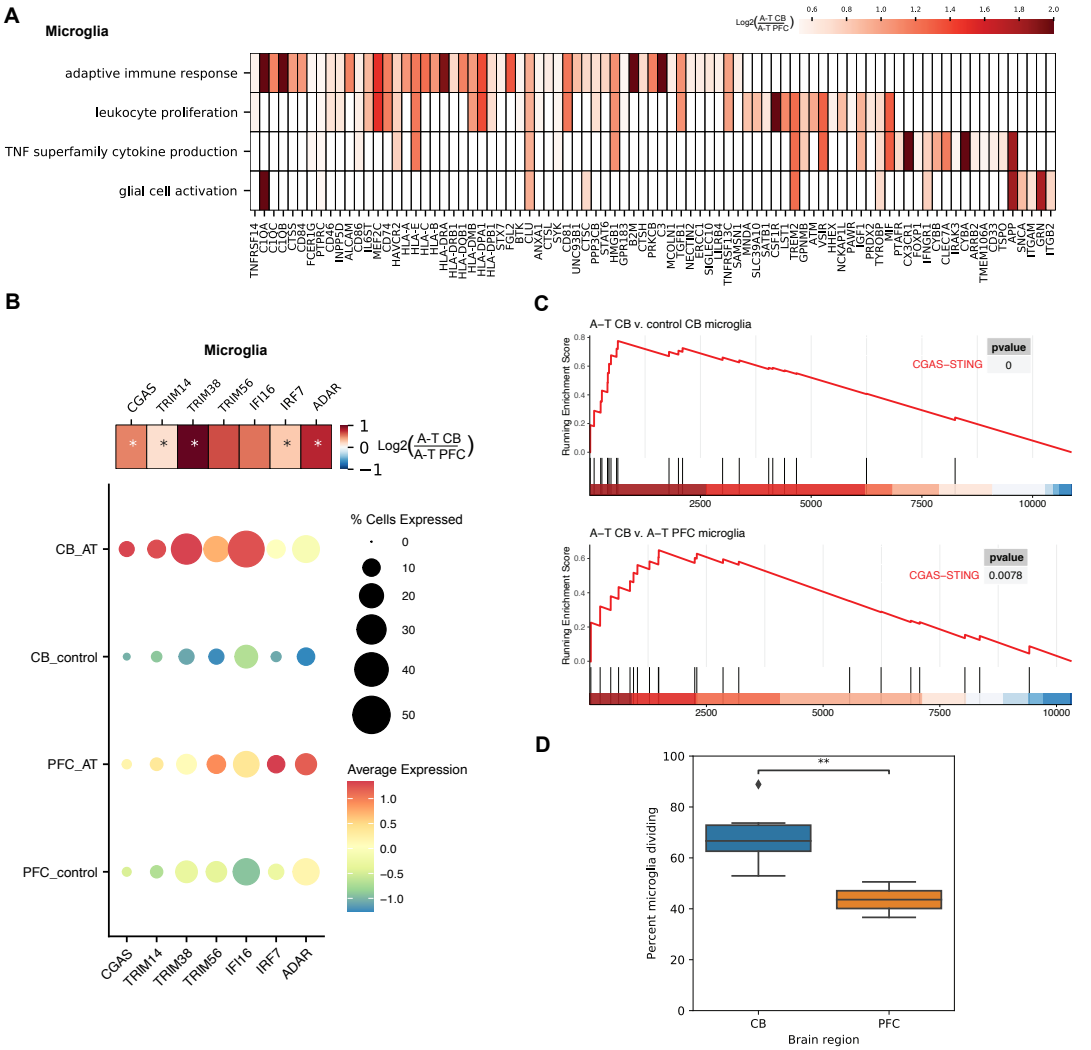


Figure 1.5: Stronger activation of microglia in A-T cerebellum compared to PFC. A. Heatmap showing enriched GO biological processes and log2 fold-changes of significant DEGs in each pathway with greater upregulation in A-T CB microglia than A-T PFC microglia. B. Dotplot of average scaled expression of CGAS-STING pathway genes in microglia of the cerebellum (CB) and PFC in A-T and control. Heatmap shows log2 fold-change of CGAS-STING pathway genes in A-T cerebellar microglia versus AT PFC microglia. *FDR<0.05. C. Gene Set Enrichment Analysis (GSEA) plot for the CGAS-STING pathway in A-T CB microglia versus control microglia (top), and A-T CB versus A-T PFC microglia (bottom). D. Percentage of microglia putatively in replicating phase in control cerebellum and PFC. **p-value <0.01, t-test.

of G1, S, and G2 phases. We found that cerebellar microglia have an increased percentage of cells in replicating phases compared to cortical microglia (p-value < 0.01, t-test) (Figure 1.5D). This is consistent with previous reports that cerebellar microglia have a higher turnover rate than cortical microglia, which may make them more vulnerable to accumulation of DNA damage due to *ATM* deficiency[252,262]. The differences in proliferation may underlie our observation that cerebellar microglia have a stronger CGAS-STING mediated immune response in A-T than cortical microglia.

1.2.8 A-T patient iPSC-derived microglia demonstrate cell-intrinsic activation of NF-kB and type I interferon signaling

To understand whether microglia might contribute to neurodegeneration in A-T, we sought to determine when cerebellar microglia become activated in disease progression relative to Purkinje and granule cell death. We employed pseudotime analysis to understand the cell-type-specific cascades of molecular events across disease progression in A-T[214]. For each cell type, cells were ordered along a continuous trajectory that progressed from the control to diseased state (Figure A.8A). We performed differential analysis to identify genes that change as a function of pseudotime and clustered these genes by their pseudotime expression patterns. We then applied GO analysis to determine the enriched biological functions of each gene cluster. For microglia, the beginning of pseudotime (healthy state) had high expression of genes involved in cellular respiration, oxidative phosphorylation, and synaptic vesicle exocytosis, consistent with a high metabolism gene expression profile characteristic of cerebellar microglia[77] (Figure A.8B). Over disease progression, upregulation of CGAS-STING pathway genes (i.e., *CGAS, TRIM14, TRIM38, TRIM56*) was followed by leukocyte activation, cytokine production, and innate immune response (Figure A.8B-C). In Purkinje neurons, disease progression was characterized first by a decline in genes related to calcium ion homeostasis (i.e., *ITPR1*), followed by late of a cluster of genes related to immune response and response to cytokine (i.e., *HSPA1A, B2M, ISG15*), and apoptotic signaling pathway (i.e., *XBP1, BCL2,*

CASP2) (Figure A.8D). Granule neurons first demonstrated a decrease in chromatin organization (i.e., *HDAC4, HDAC2, CHD5*) and glutamatergic synaptic signaling genes (i.e., *NRXN1, RELN, SYT1*), followed by upregulation genes related to oxidative phosphorylation (i.e., *PINK1, NDUFS5, UQCRH*), then upregulation of genes related to response to cytokine (i.e., *SP100, STAT6, ISG15*), response to unfolded protein (i.e., *HSPA1A, HSPE1, DNAJB1*), and programmed cell death (i.e., *IFI6, BCL2, XBP1*) at the end of pseudotime, suggesting that synaptic and metabolic dysfunction precedes granule neuron death in A-T (Figure A.8E).

We then compared molecular events between microglia, Purkinje, and granule cells along disease progression by aligning their pseudotime trajectories using dynamic time warping[36]. This revealed that the upregulation of inflammatory response and cytokine production genes in microglia (i.e., *C1QA, IL18, IFI16, CGAS, LTBR, TNFSF10*) occurs earlier in disease progression pseudotime than the onset of response to cytokine and apoptosis genes (*XBP1, MUC1, TFAP2A*) in Purkinje and granule cells (Figure 1.6A-B). These findings suggest that microglial activation and inflammation precedes Purkinje and granule cell death in A-T and support a model of cell-intrinsic microglial activation in A-T.

To experimentally characterize microglial activation in A-T, we developed iPSC-based experimental models using A-T patients and controls (Figure 1.1B, Figure A.9A). We generated human iPSCs from fibroblasts of an individual with A-T with compound heterozygous splice-altering variants in ATM (hg38 chr11:g.108332838C>T; chr11:g.108347266_108353792del27) and used the well-characterized PGP1 iPSC line as a control. We then differentiated microglia-like cells (iMGL) from the A-T patient and control iPSCs through a hematopoietic progenitor intermediate[1] (Figure A.9A). The resulting iMGLs were a highly homogeneous population with >85% of cells expressing microglia markers CD45 and CD11B assessed by flow cytometry (Figure A.9B, Figure A.10A). Profiling of the iMGLs by RNA-sequencing showed that they expressed canonical markers of microglia (*AIF1, CX3CR1, CSF1R, SPI1, TREM2*) but did not express myeloid lineage (*MPO, KLF2*), neu-

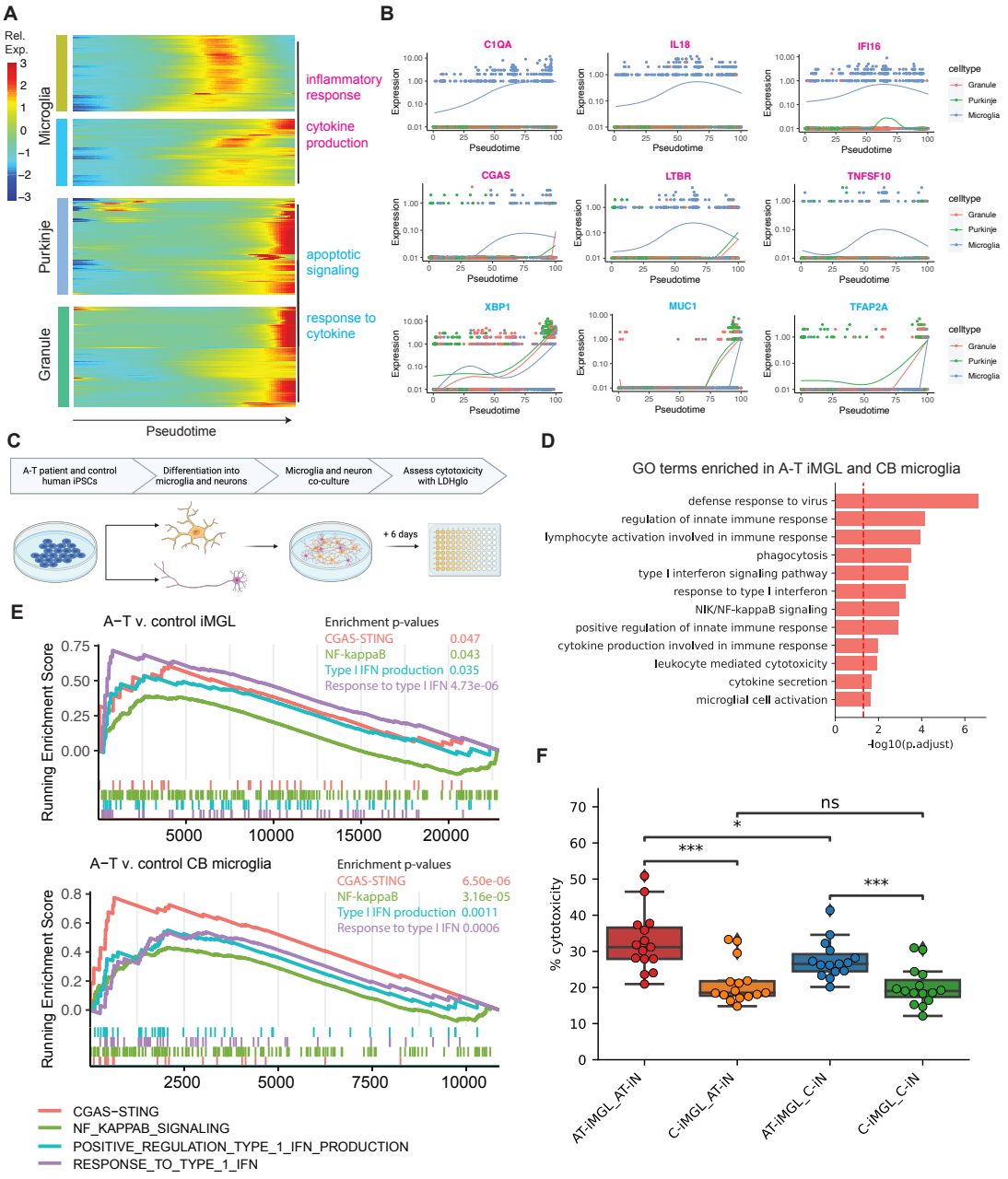


Figure 1.6: A-T patient iPSC-derived microglia reveal cell-intrinsic activation and increased cytotoxicity in neuronal co-cultures. A. Heatmap of genes that significantly change over aligned pseudotime in microglia, Purkinje, and granule neurons clustered by pseudotemporal expression patterns. Each cluster is annotated with enriched GO terms (FDR<0.05). B. Expression of select inflammatory genes (magenta) and apoptotic signaling and response to cytokine genes (cyan) over aligned pseudotime in microglia, Purkinje and granule neurons.

Figure 1.6 (Continued) C. Schematic of A-T and control iPSC-microglia (iMGL) and iPSC-neuron (iN) co-culture experiment. D. Gene Ontology (GO) terms enriched among genes upregulated in A-T iMGL and A-T cerebellar microglia versus control. E. Gene Set Enrichment Analysis (GSEA) plots for CGAS-STING, NF-kappaB, type I IFN production, and response to type I IFN pathways in A-T versus control iMGL (top) and A-T versus control CB microglia (bottom). F. Percent cytotoxicity of iMGL and iN co-cultures based on LDH-based cytotoxicity assay (bottom). C, PGP1 control. *p<0.05, ***p<0.001, ns, not significant, Mann-Whitney-U test.

ronal (*MAP2, SOX2*), and pluripotency (POU5F1) marker genes (Figure A.9C). Principal component analysis with transcriptomic profiles from A-T and control iMGLs along with our snRNA-seq data from human adult cerebellum and published snRNA-seq data from human fetal cerebellum[4] confirmed that iMGLs cluster more closely with microglia from adult and fetal human cerebellum than with other brain cell types (Figure A.9D, Figure A.10B). Our clustering analysis also confirmed that iMGLs were most correlated with microglia in adult and fetal human cerebellum (Figure A.10C).

We performed differential gene expression analysis of A-T versus control iMGL, revealing upregulation of genes involved in cytokine secretion, phagocytosis, NF-kB signaling, and type I interferon signaling (Figure A.9E, Table S1.18-1.20). To compare A-T iMGL and A-T cerebellar microglia, we obtained genes upregulated in both and identified enrichment of immune related GO terms, including microglial cell activation, cytokine secretion, NF-kB, and type I interferon signaling pathway (Figure 1.6D, Table S1.21). These results confirmed that A-T microglia grown in isolation take on an activated phenotype, demonstrating that this is a cell intrinsic property of microglia in A-T.

We hypothesized that cell-intrinsic activation of NF-kB and type I interferon signaling in A-T microglia could be due to cytoplasmic DNA triggering CGAS-STING, since a major consequence of CGAS-STING activation is initiation of NF-kB signaling and production of type I interferon related cytokines[97,257]. In support of cytoplasmic DNA as a potential trigger for microglia activation in A-T, A-T iMGL also showed positive enrichment of a CGAS-STING gene set (Figure 1.6E, Table S1.22). The similar enrichment of CGAS-STING, NF-kB, and type I interferon pathways

among genes upregulated in A-T iMGL and A-T postmortem microglia supports that A-T microglia are activated by a cell-intrinsic mechanism that may involve cytoplasmic DNA and CGAS-STING (Figure 1.6E, Table S1.22).

1.2.9 In-silico and in-vitro models of microglia-neuron interactions reveal that A-T microglia increase cytotoxicity

A-T Purkinje and granule neurons showed enrichment of response to cytokine genes late in disease progression, raising the possibility that cytokines produced by microglia induce molecular changes in Purkinje and granule neurons in A-T (Figure 1.6A). Consistently, ligand-receptor analysis revealed increased cytokine signaling (*IL1B, IL7, IL4, IL15*), type I interferon signaling (*IFNA2*), CD40 signaling (*CD40LG*), and complement signaling (*C3, C4*) from microglia to granule and Purkinje neurons in A-T versus control (Figure A.9F, Figure A.11). Notably, several of these pro-inflammatory cytokines, including *IL1B* and *IFNA2*, have been shown to induce neurotoxicity[292,168,124]. This suggests that activated microglia in A-T have increased secretion of pro-inflammatory cytokines that may have neurotoxic consequences.

To test this hypothesis experimentally, we differentiated A-T patient and control iPSCs into neurons using overexpression of Ngn263 and co-cultured them with A-T patient and control iMGLs (Figure 1.6C). iPSC-derived neurons (iNs) transcriptionally correlated with human midgestational cerebellum, making them a relevant model for studying microglia-neuron interactions in A-T (Figure A.9D). We co-cultured iMGLs and iNs for six days and assessed cytotoxicity by lactate dehydrogenase (LDH) release. We found that co-culturing control iNs with A-T iMGLs was sufficient for increased cytotoxicity compared to co-culturing with control iMGLs (Figure 1.6C, p-value<0.001). Moreover, A-T iMGLs co-cultured with A-T iNs demonstrated increased cytotoxicity compared to control iMGLs co-cultured with A-T iNs (Figure 1.6F, p-value<0.001). A-T iMGLs co-cultured with A-T iNs had the highest levels of cytotoxicity, suggesting that A-T iNs

are more vulnerable to A-T iMGLs. Taken together, these results indicate that A-T microglia have increased pro-inflammatory interactions with neurons and lead to increased cytotoxicity.

1.3 Discussion

In this study, we present an atlas of human cerebellar transcriptomes in health and degeneration at single-cell resolution. We identified major neuronal and glial cell types in the cerebellum and recovered known and novel transcriptomic signatures in A-T. We resolved cell-type-specific transcriptional perturbations in A-T, including the activation of pro-inflammatory pathways in microglia that was more pronounced in the cerebellum versus PFC. The activation of microglia in A-T PFC despite lack of cortical atrophy or neuronal loss suggests that A-T microglia are activated by a cell-intrinsic mechanism rather than in response to neurodegeneration, with brain region specific factors leading to stronger activation in the cerebellum. We found that genes in the CGAS-STING pathway, including CGAS itself, were significantly upregulated specifically in A-T microglia, with preferential activation in the cerebellum. Moreover, A-T patient iPSC-derived microglia cultured in the absence of neurons also demonstrated upregulation of genes involved in CGAS-STING, NF-kB, and type I interferon signaling, further suggesting a cell-intrinsic mechanism of activation. The CGAS-STING pathway is activated in response to cytosolic DNA, suggesting the presence of cytosolic DNA in A-T microglia. These results are consistent with recent reports of cytosolic DNA and CGAS-STING activation in microglia from rodent models of A-T[215,246]. Potential sources of cytosolic DNA include release of nuclear self-DNA[109,59], mitochondrial DNA released by defective mitochondria[145,54,249], reverse-transcribed transposable elements as observed in aging-related disorders[50], and phagocytosis of nucleic acids from cellular debris. It will be of great interest for future work to identify the source of cytosolic DNA in A-T microglia as a potential therapeutic avenue for modulating inflammation in A-T. Reducing inflammation in rodent models of A-T with ibuprofen

led to decreased microglial activation and neuronal apoptosis, further highlighting microglia as a therapeutic target in A-T[107].

Activated microglia have been intensely studied in the context of neurodegenerative disorders including Alzheimer's disease, Parkinson's disease, and ALS[94]. Our study highlights the involvement of microglia in human cerebellar degeneration as well. Through pseudotime analyses, we found that the rise in pro-inflammatory genes in A-T microglia precedes the onset of cytokine response, and apoptosis related genes in A-T Purkinje and granule neurons, and ligand-receptor analysis indicated increased pro-inflammatory signaling from microglia to neurons in A-T. These findings support a model of pathogenesis whereby cell-intrinsically activated microglia release neurotoxic cytokines and factors that trigger Purkinje and granule cell death. Recent work in human microglial cell lines demonstrated that persistent DNA damage due to loss of ATM leads to microglial activation and excessive phagocytosis of neurites, providing further validation for a cell-intrinsic mechanism of microglia activation that contributes to neuronal damage10. Rodent models of A-T also show evidence of neuroinflammation and microglial activation that leads to secretion of neurotoxic cytokines[215,246,148], but why these models do not fully recapitulate cerebellar degeneration remains unknown. Potential reasons may be that the rodent cerebellum has an increased numerical density (number/mm^3) of Purkinje cells compared to human cerebellum[132], or Purkinje cell transcriptomic differences between rodent and human may underlie the different phenotypes[32,289]. Here, our data from human A-T patients and patient-derived iPSC models provides critical support that microglial activation, previously demonstrated in vitro and in rodent models, is also found in human A-T patient brains and may contribute to A-T disease mechanisms. The pathological nature of pro-inflammatory signaling from activated microglia is also supported by clinical observations that treatment with glucocorticoids reduced ataxia symptoms in A-T[179]. The presence of activated microglia in A-T PFC without cortical neuron loss suggests that cerebellar Purkinje and granule neurons are more vulnerable to activated microglia. Moreover, the increased cytotoxicity of A-T

iNs co-cultured with A-T iMGLs suggest that A-T iNs have increased vulnerability to microglia. It is highly possible that cell autonomous and non-cell autonomous effects of ATM loss in cerebellar neurons and microglia interact to cause A-T pathology. Given that microglia and other brain-resident macrophages share several transcriptomic markers, it is possible that the microglia cluster may include a smaller percentage of other brain-resident macrophages. As markers of subtypes of brain-resident macrophages become elucidated, future studies will be able to better distinguish these cell types in the brain. Moreover, it is possible that some of the transcriptional alterations identified in certain cell types reflect processes other than degeneration, such as compensatory changes in surviving cells. It will be important for future work to further dissect the developmental onset, contribution of cell type interactions, and cellular responses to cerebellar degeneration in A-T.

This study identifies cell-type-specific changes in A-T across multiple brain regions and serves as a resource as an adult human cerebellum vermis dataset in health and degeneration. While granule and Purkinje neurons have been the most studied cell types in A-T due to their selective degeneration, this study highlights the central role of microglia in A-T cerebellar degenerative pathogenesis.

1.4 Acknowledgements

We acknowledge the NIH NeuroBioBank for providing human tissues used in this study. We would like to thank Brad Margus from the AT Children's Project for helpful scientific discussion. We also acknowledge Allison R. McLean and Jeffery R. Gulcher from Genuity Sciences for their assistance with snRNA-seq data coordination, Ronald Mathieu from HSCI-BCH Flow Cytometry Research Facility for his assistance with flow cytometry, and HSCI iPS Core Facility for their assistance with iPSC derivation. Illustrations were created with BioRender.com.

This work was supported by the National Institute of Health grants: K01 AG051791 (E.A.L), DP2 AG072437 (E.A.L), R00 AG054748 (M.A.L), AT Children's Project (E.A.L, M.A.L.), Suh

Kyungbae Foundation (E.A.L), Paul G. Allen Family Foundation (E.A.L), The American Federation for Aging Research (M.A.L. and A.M.J.), and Charles Hood Child Health Foundation (E.A.L., M.A.L). J.L. was supported by award T32GM007753 from the National Institute of General Medical Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health.

1.5 Author Contributions

E.A.L., M.A.L., J.K., and J.L. conceptualized the study. M.A.L., A.J., T.W.C., and P.G.B. performed single-nucleus RNA-sequencing experiments. J.L and J.K analyzed the single-nucleus RNA-sequencing data. J.L and D.C performed iPSC-derived experiments. A.T. performed the immunohistochemistry analyses. J.L and J.P performed WGS analysis. E.AL., T.W.Y, and M.A.L. supervised the study. J.L. wrote the manuscript. All authors reviewed and revised the manuscript.

1.6 Methods

1.6.1 Isolation of single nuclei from fresh frozen brain tissue samples

All human postmortem brain tissue samples were obtained from the NIH NeuroBioBank. Samples were matched as closely as possible for age, sex, RIN, and PMI. All available neuropathology reports and medical records for each sample were obtained. Detailed sample information is in Table S1.1. All cerebellar vermis samples are from the same anatomical section (section 1) and include the entire vermis, and all cortex samples are from BA9/BA46, which are adjacent Brodmann areas. Sectioning was performed as described by the University of Maryland Brain and Tissue Bank Sectioning Protocol: "LEFT HEMISPHERE: CEREBRUM A cut is made just posterior to the cerebral peduncle and the midbrain/pons/cerebellum are removed as a unit from the left hemisphere (Cut #3).

The remaining cerebrum is sectioned coronally, at approximate 1 cm intervals beginning from the frontal pole apex and proceeding caudally. As each section is isolated, it is gently rinsed with water, blotted dry, assigned a sequential numeric identifier (odd numbers only!), and placed in the freezing bath. The handling of sections is best aided by the use of a plastic spatula. Each frozen section is placed into individual plastic bags appropriately labeled and sealed. All bags are then stored in a -80 degree Centigrade freezer prior to shipping. Frozen sections of the cerebrum are identified as sections 1,3,5,7,9. LEFT HEMISPHERE: CEREBELLUM The remaining cerebellum is placed in a vertical plane (its normal anatomic position) and sectioned at 0.5 to 0.6 cm intervals beginning from the medial surface (vermis) and moving laterally. Each resulting section is assigned a sequential identifier (odd numbers only!). Frozen sections of the cerebellum are identified as sections 1,3,5,7,9" (https://medschool.umaryland.edu/btbankold/Brain-Protocol-Methods/Brain-Sectioning—Protocol-Method-2/).

The nuclei isolation protocol was adapted from two previous publications [174,248]. Specifically, the protocol was performed as follows: all procedures were performed on ice or at 4C. Approximately 100 mg of fresh frozen samples were processed using a Dounce homogenizer in 5 ml of tissue lysis buffer (0.32M sucrose, 5mM CaCl2, 3mM MgAc2, 0.1mM EDTA, 10mM Tris-HCl (pH 8), 0.1% Triton X-100, 1mM DTT). The homogenized solution was loaded on top of a sucrose cushion (1.8M sucrose, 3mM MgAc2, 10mM Tris-HCl (pH 8), 1mM DTT) and spun in an ultracentrifuge in an SW28 rotor (13,300 RPM, 2hrs, 4c) to separate nuclei. After spinning, supernatant was removed and nuclei were resuspended (1% BSA in PBS plus 25ul 40U/ul RNAse inhibitor) then filtered through a 40um cell strainer. After filtration, nuclei were counted and diluted to a concentration of 1000 cells/ul.

1.6.2 Droplet-based snRNA-seq

Droplet-based libraries were generated using the Chromium Single Cell 3' v3 reagent kits (10x Genomics) according to manufacturer's instructions. Resulting libraries were sequenced on an Illumina Novaseq at 150 paired-end reads with target depth 20,000 reads/nucleus.

1.6.3 snRNA-seq analysis

snRNA-seq data for each individual was preprocessed and aligned to the GRCh38-3.0.0 reference genome using CellRanger count. After alignment, ambient/background RNA was removed from the count matrix using CellBender remove-background (v0.2.0)[70]. The resulting count matrices were used to create Seurat objects per individual using Seurat v4[253]. Further quality control removed cells with less than 200 genes per cell, greater than 5000 genes per cell, and greater than 30% mitochondrial content. DoubletFinder[178] was used to detect and remove doublets with an expected multiplet rate of 7.6% for 16,000 cells loaded per sample based on the 10x user guide.

SCTransform was used to perform normalization of expression values per sample[80]. After normalization, samples were integrated to align common clusters across individual datasets using Seurat's integration method[253]. Dimensionality reduction was performed using RunPCA and RunUMAP. A shared nearest neighbor graph was constructed using FindNeighbors with dims 1:30, k.param 100, and cosine metric. Clusters were then identified using the Leiden algorithm at a resolution of 1.5.

Marker genes for each cluster were obtained using FindAllMarkers with the default minimum log2 fold-change of 0.25. Cluster markers were used to assign cerebellar cell type identities based on known literature cell type markers. For annotating clusters in the prefrontal cortex, SingleR was used for reference-based annotation since reference datasets from the human prefrontal cortex are available[10]. Velmeshev et al., 2019 was used as the prefrontal cortex reference[275].

To identify changes in cell type proportions in AT compared to controls, we used a Bayesian statistics approach to compositional data analysis previously described in the context of microbial ecology[86]. Compositional data has a negative covariance structure that is accounted for by the multinomial and Dirichlet probability distributions. We implemented Dirichlet multinomial modeling (DMM) using Hamiltonian Monte Carlo (HMC) using the R packages rstan and bayestestR (Stan Development Team (2020). "RStan: the R interface to Stan." R package version 2.21.2, http://mc-stan.org/.)[166]. The input to the model was a matrix of counts where the rows correspond to replicates (individual samples) and the columns correspond to cell types. We then modeled the replicates using the multinomial distribution for the probability of each cell type per individual. The Dirichlet distribution was used to model the multinomial parameters. The prior on the Dirichlet parameters was another Dirichlet distribution with a fixed parameter alpha 10^{-7}, which gives uniform cell type proportions in expectation. The resulting posterior distributions were Dirichlet distributions of the cell type proportions in AT and control. We then subtracted the posterior probability distribution (89% credible interval) of control from AT to see whether there are significant differences in relative cell type composition. A cell type shift in log2 abundance ratio (AT/control) was considered significant if the 89% credible did not include 0.

1.6.4 Differential gene expression and Gene Ontology analysis

Differential gene expression analysis was performed on each cell type using edgeR[176]. The integrated Seurat object was subset to obtain Seurat objects for each cell type. The read counts were modeled and normalized using a negative binomial distribution with the trimmed mean of M-values (TMM) normalization method. To ensure robust signals, only genes that were expressed in at least 5% of one cell type were included in the analysis. The design matrix formula was disease.status + sex + age + cellular detection rate. Differentially expressed genes between AT and control were identified using the likelihood ratio test (glmLRT). Genes with FDR < 0.05 and a log2 fold-change

greater than 0.50 were used for downstream analysis.

Pseudobulk differential gene expression analysis was performed using DESeq2 [162]. To generate pseudobulk data, we took the sum of raw counts for each gene over all cells in each sample, resulting in a gene by sample counts matrix. The design matrix was specified as sex + age + disease.status. Differentially expressed genes between AT and controls were identified after running lfcShrink. Genes with FDR < 0.05 were considered statistically significant.

Gene Ontology (GO) enrichment analysis was performed on the upregulated and downregulated differentially expressed genes in each cell type using clusterProfiler [294]. GO biological processes with FDR < 0.05 were considered significant.

Gene Set Enrichment Analysis was performed on differentially expressed genes ordered by log2 fold-change (descending order) with the function GSEA() with default settings from clusterProfiler. Human gene sets of interest were obtained from the Molecular Signatures Database (MSigDB) [254,157] and are listed in Supplementary Table 1.21.

CellChat [114] was used to perform ligand-receptor analysis in order to identify differential cell-cell communication in AT compared to control. Differentially active signaling pathways with FDR<0.05 were considered significant.

1.6.5 Cell Cycle Scores

AddModuleScore calculates the average expression of a gene set in each cell, subtracted by the average expression of a control gene set, as previously described [265]. A positive score means the module of genes is expressed more highly in a particular cell than expected by chance. Cell cycle phase scores were calculated using this method based on canonical cell cycle markers using the Seurat function CellCycleScoring, which then outputs the predicted classification (G1, S, or G2M phase) for each cell.

1.6.6 Disease gene analysis

Known monogenic cerebellar disease genes were obtained from the Online Mendelian Inheritance in Man (OMIM) database. The mouse cerebellum data was obtained from Saunders et al., 2018[233]. To determine the correlation of cerebellar disease gene expression between humans and mice, cosine similarity was calculated between the human and mouse expression vectors of the genes for each cell type.

1.6.7 Pseudotime analysis

Pseudotime trajectory analysis was performed using Monocle2[214]. First, Seurat objects for each cell type (Microglia, Purkinje, and granule cells) were converted into the Monocle2 compatible cell dataset form using as.CellDataSet. Cells were clustered in an unsupervised manner using clusterCells. The top 1000 differentially expressed genes between AT and control in each cell type were used as the ordering genes to order cells along the pseudotime trajectory. The root state of the trajectory was defined as the cluster with the highest proportion of control cells. Genes that change as a function of pseudotime were identified using differentialGeneTest with the model formula specified as sm.ns(Pseudotime), which fits a natural spline to model gene expression as a smooth nonlinear function of pseudotime. Genes that change over pseudotime with an FDR<0.05 were considered significant. These genes were then clustered according to their pseudotime expression profiles. GO enrichment analysis was performed on each cluster using clusterProfiler to identify the biological processes that co-vary across pseudotime.

To align the pseudotime trajectories between cell types, the method dynamic time warping (DTW) was applied as previously described[36]. DTW aligns similar temporal processes that differ in length by calculating the optimal matching between points in each sequence. For each cell type, ordered cells in pseudotime were binned into 100 "pseudocells" and the average expression of each

gene was calculated within each pseudocell. The resulting gene x pseudocell matrices were the input for the dtw package in R. The Purkinje cell matrix was set as the reference and the other cell type was the query. The cost matrix was calculated as 1-corr(M1, M2), where corr(M1, M2) is the cosine similarity matrix between cell type 1 matrix (M1, reference) and cell type 2 matrix (M2, query). The query cell type pseudotime values were then warped into the reference to obtain aligned pseudotimes.

1.6.8 Whole Genome Sequencing and Variant Calling

DNA was isolated using the Qiagen DNA Mini kit (Qiagen cat. no. 51304) according to the protocol for tissues. Approximately 25mg of fresh frozen tissue was minced into small pieces. Tissue was transferred to a 1.5ml microcentrifuge tube and 180µl of Buffer ATL as added. 20ul of proteinase K was added prior to 4 hours of agitation at 56 degrees centigrade on a thermomixer (1400 RPM). DNA isolation proceeded as written in the protocol with the inclusion of the optional RNase A step. Libraries were prepared with the TruSeq DNA PCR-free library kit (Illumina). Libraries were sequenced at 30X coverage with Illumina NovaSeq 6000 S4, 2 x 150 bp. WGS data for each individual was processed using GATK491. Specifically, we used the DRAGEN-GATK whole genome germline pipeline for variant discovery with the WGS_Maximum_Quality mode on Terra (terra.bio). Reads were aligned using hg38 as the reference genome. Resulting VCF files were annotated using ANNOVAR92. Sample information and variants are listed in Table S1.5.

1.6.9 Immunostaining and Imaging

Postmortem human formalin fixed vermis samples from four subjects diagnosed with Ataxia Telangiectasia and three age matched controls were obtained from the University of Maryland Brain and Tissue Bank (Table S1.9). The samples were embedded in paraffin by the UMCCTS Biospecimen,

Tissue and Tumor Bank, and sectioned on a LEICA RM2125 microtome. Sections were cut at 20 μm thickness, placed on uncharged untreated Fisherfinest premium microscope slides (product number 12-544-7), and dried overnight at 37°C. Eight sections (four control and four with AT pathology) were selected at random, deparaffinized with xylene and rehydrated in a graded ethanol wash series. This was followed by antigen retrieval, done by heating citrate buffer (pH 6.2) to 90°C and leaving the slides to soak for 45 minutes, then allowing the solution and slides to cool to room temperature. The samples were then subjected to brief PBS washes before application of a primary antibody solution consisting of blocking solution, 10% triton in 10x PBS for permeabilization, and primary antibody. Primary antibodies used included CD11B/Integrin Alpha M Monoclonal Antibody (Proteintech, product number 66519-1-Ig, concentration 1:100), and Anti-GFAP antibody produced in rabbit (Millipore Sigma product number HPA056030, concentration 1:250). The samples were incubated at 4°C overnight. The following day the samples were washed for an hour and a half in cold 1x PBS before application of secondary antibody solution consisting of blocking solution, 10% triton, and Donkey anti-Mouse Cross-Adsorbed Secondary Antibody, Alexa Fluor plus 555 (ThermoFisher Scientific, Invitrogen, product number A32773, concentration 1:125) and Donkey anti-Rabbit Cross-Adsorbed Secondary Antibody Alexa Fluor plus 488 (ThermoFisher Scientific, Invitrogen, product number A32790, concentration 1:125). After incubating overnight in the dark at 4°C the samples were washed in 1x PBS for an hour and a half in the dark. The tissue was stained for DAPI (ThermoFisher Scientific, Product Number 62248, concentration 1:10,000 in 1x PBS), washed and then treated with TrueBlack lipofuscin autofluorescence quencher (Biotium, product number 23007) according to the manufacturer's protocol. The samples were subjected to 30 minutes of washes in Barnstead water before being allowed to air dry. The tissue was mounted in Everbrite (Biotium, product number 23003) with Platinum Line cover glass (FisherScientific, Product Number 15-183-89) and left to cure overnight. Imaging was performed on a Zeiss LSM700 inverted microscope using ZEN Black 2012 Software.

1.6.10 Maintenance and culture of iPSCs

Fibroblasts from an individual with A-T (2 years old, female) were reprogrammed into iPSCs using non-integrative Sendai virus. The patient had the following compound-heterozygous variants in ATM: hg38 chr11:g.108332838C>T; chr11:g.108347266_108353792del27. Control iPSCs from Personal Genome Project participant 1 (PGP1) were used as the control line. Both iPSC lines were mycoplasma negative, karyotypically normal, and expressed pluripotency markers OCT4, SOX2, NANOG, SSEA4, TRA-1-60. iPSCs were maintained in 6-well or 10 cm plates (Corning) coated with LDEV-free hESC-qualified Matrigel (Corning Catalog #354277) in feeder-free conditions with complete mTeSR-plus medium (STEMCELL Technologies) in a humidified incubator (5% CO_2, 37°C). iPSCs were fed fresh media daily and passaged every 3-4 days.

1.6.11 Differentiation of iPSCs to Hematopoietic Progenitor Cells

Human iPSC-derived hematopoietic progenitors (HPC) were generated using the STEMdiff Hematopoietic Kit (STEMCELL Technologies) according to the manufacturer instructions based on a protocol by Abud et al., 2017 [1]. Briefly, iPSCs were dissociated into aggregates using Gentle Cell Dissociation Reagent (STEMCELL Technologies) and 40-80 aggregates were plated per well on Matrigel coated 12-well plates. Cells were maintained in Medium A to induce a mesoderm-like state until day 3. From day 3 to day 12, cells were maintained in Medium B with half-media changes every 2 days to promote further differentiation into hematopoietic progenitor cells. To confirm the differentiation from iPSC to HPC, we performed flow cytometry on day 12 at the end of HPC differentiation. Cells were harvested and incubated with CD45-APC (1 µl per 50,000 cells in 100 µL, STEMCELL Technologies Catalog #60018AZ) and CD34-PE (1 µl per 50,000 cells in 100 µL, STEMCELL Technologies Catalog #60119PE) antibodies for 15 minutes in the dark at 4C. In addition to double-stained samples, we included unstained samples, Fluorescence Minus One (FMO)

controls, and compensation single-stain controls using OneComp eBeads (Thermo Scientific Catalog #01111142). Samples were washed twice with FACS buffer (Phosphate Buffered Saline (PBS) with 2% FBS) and transferred to 5 ml Round Bottom Polystyrene FACS tubes. Flow cytometry was performed at the HSCI-BCH Flow Cytometry Research Facility using the BD FACSAria II. Analysis was performed using FlowJo software.

1.6.12 Differentiation of Hematopoietic Progenitor Cells to Microglia

Human iPSC-derived microglia (iMGL) were generated from human iPSC-derived hematopoietic progenitors using the STEMdiff Microglia Differentiation Kit and Microglia Maintenance Kit (STEMCELL Technologies) according to manufacturer instructions. Briefly, hematopoietic progenitor cells from day 12 of the STEMdiff Hematopoietic Kit protocol were collected and transferred to a Matrigel coated 6-well plate (200,000 cells per well) with STEMdiff Microglia Differentiation Medium. From day 0 to 24, 1 mL Microglia Differentiation Medium was added to each well every other day. After day 24, cells were maintained in STEMdiff Microglia Maturation Medium with half-medium volume added every other day. To confirm the differentiation from HPC to iMGL, we performed flow cytometry on day 24 at the end of microglia differentiation. Cells were harvested and blocked with CD32 antibody (2 µl per 50,000 cells in 100 µL, STEMCELL Technologies Catalog #60012) for 15 minutes at 4C. Samples were then incubated with CD45-APC (1 µl per 50,000 cells in 100 µL, STEMCELL Technologies Catalog #60018AZ) and CD11B-FITC (1 µl per 50,000 cells in 100 µL, STEMCELL Technologies Catalog #60040FI) antibodies for 15 minutes in the dark at 4C. In addition to double-stained samples, we included unstained samples, FMO controls, and compensation single-stain controls using OneComp eBeads (Thermo Scientific Catalog #01111142). Samples were washed twice with FACS buffer (Phosphate Buffered Saline (PBS) with 2% FBS) and transferred to 5 ml Round Bottom Polystyrene FACS tubes. Flow cytometry was performed at the HSCI-BCH Flow Cytometry Research Facility using the BD FACSAria II.

Analysis was performed using FlowJo software.

1.6.13 RNA-sequencing of A-T and control iMGL

RNA was harvested from day 29 iMGL samples using the PureLink RNA Mini kit (Life Technologies Catalog # 12183018A). Libraries were prepared with the Illumina Stranded mRNA prep. Sequencing was performed on Illumina HiSeq, 2x150bp configuration at approximately 20X depth per sample. Reads were aligned to hg38 using STAR v2.7.9a followed by processing with featureCounts to obtain a gene-level counts matrix. Differential gene expression analysis was performed using DESeq2 with control set as the reference condition and results coefficient set as 'condition_AT_vs_control'.

1.6.14 Differentiation of iPSCs to neurons

Human iPSC-derived neurons (iNs) were generated from iPSCs transduced with the lentiviruses Tet-O-Ngn2-Puro and FUW-M2rtTA based on a previously published protocol[299]. On day -1 of differentiation, iPSC colonies were dissociated into single-cells using Accutase (STEMCELL Technologies) and 4-8 million cells were plated on Matrigel coated 10 cm dishes with complete mTeSR plus medium supplemented with Y-27632 (10 uM). On day 0, cells were fed N2 media (DMEM/F-12 media, 1X N2, 1X Nonessential Amino Acids) supplemented with doxycycline (2 μg/mL), BDNF (10 ng/mL), NT3 (10 ng/mL), and laminin (0.2 μg/mL). On day 1, media was replaced with N2 media with puromycin (1 μg/mL) in addition to the supplements listed above. On day 2, cells were fed B27 media (Neurobasal media, 1X B27, 1X Glutamax) supplemented with puromycin (1 μg/mL), doxycycline (2 μg/mL), Ara-C (2 uM), BDNF (10 ng/mL), NT3 (10 ng/mL), and laminin (0.2 μg/mL). On day 3, cells were dissociated into single cells using Accutase and replated onto polyethylenimine/laminin coated 96-well plates (10,000 cells/well) with B27 media supple-

mented with Y-27632 (10 μM), doxycycline (2 μg/mL), Ara-C (2 μM), BDNF (10 ng/mL), NT3 (10 ng/mL), and laminin (0.2 μg/mL). On day 5, cells were fed with Conditioned Sudhof Neuronal Growth Medium (1:1 ratio of Astrocyte Conditioned Media and Neurobasal Media, 1X B27, Glutamax, $NaHCO_3$, and transferrin) supplemented with BDNF (10 ng/mL), NT3 (10 ng/mL), and laminin (0.2 μg/mL).

1.6.15 Co-culture of iMGL and iNs

Day 29 iMGL were added to Day 6 iNs plated in a 96-well plate at a 1:5 ratio (2,000 iMGL:10,000 neurons per well) with 15 independent wells per co-culture condition (A-T iMGL/A-T iN; A-T iMGL/control iN; control iMGL/A-T iN; control iMGL/control iN). The co-cultures were maintained for 6 days.

1.6.16 Cytotoxicity assay

Cytotoxicity of co-cultures were assessed with the LDH-Glo Cytotoxicity Assay (Promega Catalog #J2380) according to the manufacturer instructions. 5 μl of medium from each co-culture well was diluted in 95 μl LDH Storage Buffer (200 mM Tris-HCl pH 7.3, 10% Glycerol, 1% BSA). LDH activity was measured by combining 10 μl LDH Detection Reagent with 10 μl diluted sample in a 384-well white opaque-walled assay plate (Greiner Bio-One Catalog # 784080) after incubating for 45 minutes at room temperature. LDH standards were prepared according to manufacturer instructions. Fresh culture medium was used as a no-cell control. To generate a maximum LDH release control, 2 μl of 10% Triton-X 100 was added per 100 μl to 96-wells containing iNs for 10 minutes followed by sample collection in LDH storage buffer as described above. Plate was read using the Spectramax iD5 plate reader with the luminescence read mode. Percent cytotoxicity was calculated according to the kit's protocol as % Cytotoxicity = 100 × (Experimental LDH Release –

Medium Background) / (Maximum LDH Release Control – Medium Background).

1.6.17 Statistics and reproducibility

No statistical methods were used to predetermine sample sizes. All statistical tests were performed in R (v4.0.1) or Python (v3.8) with multiple hypothesis correction using the Benjamini-Hochberg procedure (FDR<0.05) unless otherwise specified. Statistical tests are noted in figure legends, and include Fisher's exact test, Student's t-test and Mann-Whitney U-test.

1.7 Supplemental Data

Supplementary figures are included in the Thesis Appendix. Tables are included in the supplementary data of the manuscript (Lai et al., 2023, doi: https://doi.org/10.1101/2021.09.09.459619) upon which this chapter is based.

"Genetics is the key to understanding and treating many diseases, but it is also the key to understanding what makes us human."

Francis Collins

2

Disruption of RFX family transcription factors causes autism, attention-deficit/hyperactivity disorder, intellectual disability, and dysregulated behavior

This section is based on the following manuscript in which I made significant contributions during my PhD studies:

Harris HK*, Nakayama T*, **Lai J***, et al. Disruption of RFX family transcription factors causes autism, attention-deficit/hyperactivity disorder, intellectual disability, and dysregulated behavior. Genet Med. 2021 Mar 3; PubMed ID: 33658631. **These authors contributed equally.* **Supplementary Material:** All supplementary material can be found in the supplement of Harris *et al*., Genet Med. 2021.

2.1 Introduction

Autism spectrum disorder (ASD), marked by deficits in social communication and the presence of restricted interests and repetitive behavior, is highly heritable and genetically heterogeneous, with *de*

novo loss-of-function variants as known contributors to ASD risk[232]. ASD is often comorbid with other neurodevelopmental diagnoses, including attention-deficit/hyperactivity disorder (ADHD). Emerging evidence also points to a role of *de novo* loss-of-function variants in ADHD[18].

RFX3 is a member of the regulatory factor X (RFX) gene family which encodes transcription factors with a highly-conserved DNA binding domain. *RFX3* is expressed in several tissues including developing and adult brain, and other RFX family members (RFX1, 4, 5, and 7) are also highly expressed in brain tissue, with expression patterns of RFX1, 3, 4 and 7 clustering tightly[255]

We report a series of 38 individuals from 33 families with deleterious, mostly *de novo* variants in three brain-expressed members of the RFX family: *RFX3*, *RFX4*, or *RFX7*. *RFX3* was among 102 genes recently identified as statistically enriched for *de novo* variants in a large-scale analysis of trio exome data from individuals with ASD,[232] but to date *RFX4* and *RFX7* have not been previously associated with human disease. Analysis of case clinical data reveals common features including intellectual disability (ID), ASD, and/or ADHD, delineating a novel neurobehavioral phenotype associated with RFX haploinsufficiency.

2.2 Results

2.2.1 Case series of individuals with *de novo* or inherited *RFX3* variants

We identified and obtained clinical information from 18 individuals bearing loss-of-function variants in *RFX3* via GeneMatcher[244]. Genotypic information is provided in Table 2.1, clinical phenotypes are summarized in Tables 2.2 and S1, and predicted variant impacts are summarized in Supplemental Table S2.2. A total of 15 distinct variants were identified: two frameshift variants, two canonical splice donor variants, eight missense variants, one in-frame deletion, one 42 kb deletion removing the last two exons of *RFX3*, and one 227 kb deletion involving only *RFX3*. In one family, an affected parent transmitted a frameshift variant to three affected children; all other variants were

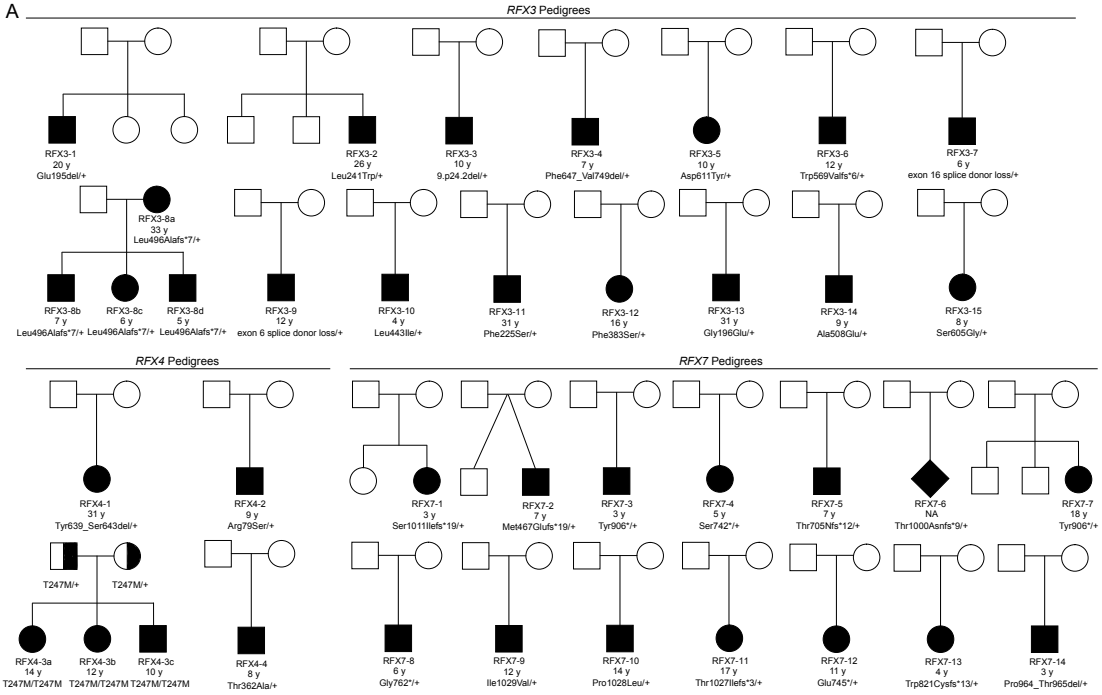


Figure 2.1: Pedigrees of reported individuals with *RFX3*, *RFX4*, and *RFX7* variants. A. *RFX3*, *RFX4*, and *RFX7* case pedigrees. All pedigrees show *de novo* origin of variants except for *RFX3*-8a-d: a 33 year-old affected mother carrying the variant p.(Leu496Alafs*7) with transmission to three children, and pedigree *RFX4*-3a-c: three affected children homozygous for p.(Thr247Met).

de novo and novel (Figure 2.1A).

There were thirteen males and five females, with no sex-based differences in severity of phenotype. All individuals had neurodevelopmental delays, with formally recorded clinical diagnoses of ASD (72%) and ID of varying severity (borderline to moderate) or global developmental delay in young children (78%) and ADHD (56%) (Table S2.1). Many showed a distinct behavioral pattern marked by easy excitability/overstimulation, hypersensitivity to sensory (particularly auditory) stimuli, anxiety, emotional dysregulation and/or aggression (13/15 [87%] with specific behavioral information provided). Three individuals were reported to have seizures (17%). Some individuals had sleep difficulties (44%) including limited total duration of sleep, frequent awakenings, or early morning awakenings. Subtle non-specific and non-recurrent dysmorphisms were commonly re-

ported (61%), including broad nasal bridge, high, arched palate, and hand and foot abnormalities (tapered fingers, widely spaced toes), but no consistent recognizable features were shared by all individuals. Both macrocephaly (six individuals) and microcephaly (two individuals) were reported (8/11 individuals [73%] with a head circumference measurement or percentile provided). Magnetic resonance imaging (MRI) of the brain was available for eight individuals, with reports of nonspecific findings in four, including white matter changes, uncal asymmetry, partially empty sella, or prominent ventricles. One individual had mild thinning of the corpus callosum. Five of seven individuals (71%) who were past the onset of puberty (ages 12-30 years) had reports of behavioral and/or cognitive worsening at the time of puberty/adolescence. Three had increased aggression specifically noted. Three were described as having manic and/or psychotic symptoms, specifically two described as having hallucinations (one requiring psychiatric hospitalization) and another described as having conversations with imaginary friends. Three were reported to have had decline in cognition, one in adolescence and another around 28 years of age.

2.2.2 Variants in additional RFX family genes are associated with similar neurodevelopmental phenotypes

Additional individuals were ascertained who harbored loss-of-function variants in other closely related genes of the RFX family. Fourteen individuals bearing *de novo* loss-of-function variants in *RFX7* were identified (Tables 1 and 2), including four frameshift variants, five stop gain variants, one in-frame deletion, and two missense variants (Table 2.1, Table S2.2). Slightly more males were identified than females (eight males, six females) without differences in phenotype based on sex. All individuals had language delay, and most had ID/global developmental delay (93%) (Table S2.1). While formal diagnoses of ASD (36%) and/or ADHD (29%) were less consistent, autistic features and/or significant behavioral challenges akin to those seen in *RFX3* individuals were reported in the majority of cases, including excitability/overstimulation, sensitivity to sensory (particularly audi-

Table 2.1: Molecular findings in individuals with ASD, ADHD, and/or ID and variants in *RFX3*, *RFX4*, or *RFX7*

Gene	Individual	Inheritance	*gDNA (GRCh38)*	cDNA	Protein	Category	Domain
RFX3	*RFX3*-1	*de novo*	*chr9:g.3293222_3293224del*	c.584_586del	p.(Glu195del)	Inframe deletion	DBD
	RFX3-2	*de novo*	*chr9:g.3293086A>C*	c.722T>G	p.(Leu241Trp)	Missense	DBD
	RFX3-3	*de novo*	*chr9:g.9p24.2del*	NA	(gene deletion)	Deletion	(all)
	RFX3-4	*de novo*	*chr9:g.3197716_3239762del*	c.1968+8270_*27326del	(exon 17 and exon 18 deleted) p.Phe647_Val749del	Deletion	DD
	RFX3-5	*de novo*	*chr9:g.3248169C>A*	c.1831G>T	p.(Asp611Tyr)	Missense	DD
	RFX3-6	*de novo*	*chr9:g.3257101dupG*	c.1704dup	p.(Trp569Valfs*6)	Frameshift	DD
	RFX3-7	*de novo*	*chr9:g.3248031C>T*	c.1968+1G>A	exon 16 splice donor loss	Splicing	DD
	RFX3-8a-d	inherited	*chr9:g.3263053_3263054del*	c.1486_1487del	p.(Leu496Alafs*7)	Frameshift	DD
	RFX3-9	*de novo*	*chr9:g.3301541C>T*	c.549+5G>A	exon 6 splice donor loss	Splicing	DBD
	RFX3-10	*de novo*	*chr9:g.3270401G>T*	c.1327C>A	p.(Leu443Ile)	Missense	EDD
	RFX3-11	*de novo*	*chr9:g.3293134A>G*	c.674T>C	p.(Phe225Ser)	Missense	DBD
	RFX3-12	de novo	chr9:g.3271057A>G	c.1148T>C	p.(Phe383Ser)	Missense	EDD
	RFX3-13	de novo	chr9:g.3293221C>T	c.587G>A	p.(Gly196Glu)	Missense	DBD
	RFX3-14	de novo	chr9:g.3263017G>T	c.1523C>A	p.(Ala508Glu)	Missense	DD
	RFX3-15	de novo	chr9:g.3256992T>C	c.1813A>G	p.(Ser605Gly)	Missense	DD
RFX4	*RFX4*-1	*de novo*	chr12:g.106750773_106750787del	c.1915_1929del	p.(Tyr639_Ser643del)	Inframe deletion	NA
	RFX4-2	*de novo*	chr12:g.106654271C>A	c.235C>A	p.(Arg79Ser)	Missense	DBD
	RFX4-3a-c	recessive (homozygous)	chr12:g.106696353C>T	c.740C>T	p.(Thr247Met)	Missense	NA
	RFX4-4	*de novo*	chr12:g.106715490A>G	c.1084A>G	p.(Thr362Ala)	Missense	DD
RFX7	*RFX7*-1	*de novo*	chr15:g.56094696del	c.3032del	p.(Ser1011Ilefs*19)	Frameshift	NA
	RFX7-2	unknown (adopted)	chr15:g.56096328_56096329del	c.1399_1400del	p.(Met467Glufs*19)	Frameshift	NA
	RFX7-3	*de novo*	chr15:g.56095010G>T	c.2718C>A	p.(Tyr906*)	Stop Gain	NA
	RFX7-4	*de novo*	chr15:g.56095503G>C	c.2225C>G	p.(Ser742*)	Stop Gain	NA
	RFX7-5	*de novo*	chr15:g.56095615dupT	c.2113dup	p.(Thr705Asnfs*12)	Frameshift	NA
	RFX7-6	*de novo*	chr15:g.56094730dupT	c.2998dup	p.(Thr1000Asnfs*9)	Frameshift	NA
	RFX7-7	*de novo*	chr15:g.56095010G>C	c.2718C>G	p.(Tyr906*)	Stop Gain	NA
	RFX7-8	*de novo*	chr15:g.56095444C>A	c.2284G>T	p.(Gly762*)	Stop Gain	NA
	RFX7-9	*de novo*	chr15:g.56094643T>C	c.3085A>G	p.(Ile1029Val)	Missense	NA
	RFX7-10	*de novo*	chr15:g.56094645G>A	c.3083C>T	p.(Pro1028Leu)	Missense	NA
	RFX7-11	*de novo*	chr15:g.56094648del	c.3080del	p.(Thr1027Ilefs*3)	Frameshift	NA
	RFX7-12	*de novo*	chr15:g.56095495C>A	c.2233G>T	p.(Glu745*)	Stop Gain	NA
	RFX7-13	*de novo*	chr15:g.56095266_56095269dup	c.2459_2462dup	p.(Trp821Cysfs*13)	Frameshift	NA
	RFX7-14	*de novo*	*chr15:g.56094864_56094869del*	*c.2859_2864del*	*p.(Pro964_Thr965del)*	*Inframe deletion*	*NA*

tory) stimuli, a high pain threshold, emotional dysregulation, aggression, and anxiety (8/8, 100% of those with specific behavioral information provided). Abnormal head size (five individuals with microcephaly and three with macrocephaly) was noted in 7/11 (64%) that provided head circumference measurements. In 5/11 patients (45%) who had neuroimaging, MRI abnormalities were observed (Dandy-Walker malformation, cerebellar tonsillar herniation, an abnormality of the basal ganglia, and a fourth case with limited information but an "abnormal brain MRI" noted). Subtle clinical dysmorphisms were reported in 86% including abnormalities of the hands and feet such as widely spaced toes, syndactyly, or long tapered fingers (50%). Again, no consistent dysmorphisms were evident across individuals.

Six individuals with probable loss-of-function *RFX4* variants were also identified (Tables 2.1). Three were individuals who harbored *de novo RFX4* variants, including an in-frame deletion (*RFX4* p.(Tyr639_Ser643del)), and two predicted damaging missense variants (*RFX4* p.(Arg79Ser) and p.(Thr362Ala)) (Table 2.1, Table S2.2). We also report a pedigree in which three additional related

individuals (siblings) were homozygous for a missense variant in *RFX4* (p.Thr247Met) altering a well-conserved threonine residue. Parents of these siblings were first cousins, each heterozygous for the mutation and without any known neurobehavioral phenotype, and a heterozygous sibling was similarly reported as neurotypical; this pedigree therefore raises the possibility that the *RFX4* phenotype may be associated with both monoallelic and biallelic inheritance as has been described for several other genetic conditions[84]. Of these six individuals, three were female and three were male. All were noted to have ID or global developmental delay (100%) and most had documented ASD (83%). Four individuals were normocephalic, and one was microcephalic. Neuroimaging was performed in two and demonstrated asymmetric volume loss in one individual and absent pituitary gland in another individual with hypopituitarism. The latter individual also presented with cleft lip and palate. Seizures were described in two individuals (33%). No consistent dysmorphisms were evident.

2.2.3 *RFX3*, *RFX4*, and *RFX7* variant analyses

In total, 33 distinct variants in *RFX* family members (15 *RFX3*, 4 *RFX4*, and 14 *RFX7* were identified (Table 2.1). Excluding related individuals, each case involved a novel variant (e.g., there were no recurrent variants). *RFX3*, *RFX4*, and *RFX7* each exhibit intolerance to loss-of-function variation in human population databases (gnomAD, pLI scores = 1.00). All variants were absent from gnomAD except for *RFX7* p.Pro964_Thr965del, which is detected at a very low frequency in gnomAD v2.1.1 (AF 0.00007677) leading us to formally classify it as a variant of uncertain significance (VUS) (see Supplemental *RFX7* Case Descriptions Individual 14 for further details). The fact that the majority of variants identified are predicted to cause outright protein truncation or gene deletion (20 out of 33) strongly supports a loss-of-function / haploinsufficiency model. Of the thirteen missense variants, eleven were predicted to be damaging by at least four of six algorithms (NsynD score >=4) and two missense variants were predicted damaging by at least 2 algorithms (*RFX4* p.(Thr247Met)

and p.(Thr362Ala)) (Table S2.2). All missense variants affect highly conserved amino acids (PhastCons vertebrate, mammalian, and primate scores ranging from 0.99-1.00) (Table S2.2).

RFX transcription factors are defined by a conserved, specialized winged-helix type DNA binding domain (DBD) that recognize the X-box motif. In addition to the DBD, *RFX3* and *RFX4* have three known domains that are associated with dimerization (DD)[255]. *RFX4* and *RFX7* variants did not exhibit clustering to specific functional domains, but all of the non-truncating (missense or in-frame deletion) variants identified in *RFX3* were found to be located in the DBD or one of the dimerization domains (Figure 2.2A, Table 2.1). We engineered five of the non-truncating variants – p.(Glu195del), p.(Leu241Trp), p.(Phe383Ser), p.(Leu443Ile), and p.(Asp611Tyr) – into a V5-*RFX3* heterologous expression vector for protein stability analyses in HeLa cells (Figure S1). The majority of these variants resulted in significant decreases in detectable *RFX3* levels, consistent with a destabilizing impact on protein expression. Two missense variants, p.(Leu241Trp) and p.(Leu443Ile) (residing in the DNA binding domain or the second extended protein dimerization domain, respectively) did not appear to impact protein stability, raising the possibility that they might disrupt more specific functional interactions of *RFX3* to be investigated more thoroughly in the future.

We examined 35 additional reported variants in *RFX3*, *RFX4*, and *RFX7* from prior studies of *de novo* or inherited variants in ASD and neuropsychiatric conditions (Table S2.3, Figure A.13)[232,225,278,150,136]. Missense variants from the literature tended to be of milder predicted deleteriousness than those reported here (Figure 2.2B). Sixteen were in *RFX3*, including five *de novo* variants (four protein truncating and one missense variant predicted damaging by all six algorithms, NsynD6, supportive of likely deleteriousness), seven inherited variants (four CNVs, one frameshift variant (p.(Pro408fs)) and three missense variants (p.(Thr151Ala), p.(Ala101Thr), p.(Arg615His); NsynD scores 3-6), and four copy number variants (all microdeletions) were reported for which parental inheritance was not established (Table S2.3). Among previously reported *RFX7* variants, one was a *de novo* frameshift

and one was an inherited frameshift variant. There were also six reported inherited missense variants (6/6 with NsynD >4), and two *de novo* missense variants that are likely benign. Finally, there were nine previously reported *RFX4* variants, only one of which was *de novo* (a missense variant lacking strong evidence of pathogenicity), and eight inherited missense variants of varying predicted deleteriousness (NsynD scores 3-6).

2.2.4 *RFX* expression is enriched in human brain

RFX3, *RFX4* and *RFX7* have been reported to have relatively high expression in human fetal cortex[119]. To determine whether specific cell types are affected by RFX haploinsufficiency, we examined single-cell transcriptomes from developing and adult human cortex (Figure 2.3A-F, Figure A.14A-B)[275,208]. In developing human cortex, *RFX3* and *RFX7* exhibited the strongest brain expression, with *RFX3* most highly expressed in maturing excitatory upper enriched neurons, *RFX4* most highly expressed in outer radial glia, and *RFX7* most highly expressed in interneurons from the medial ganglionic eminence (Figure 2.3A-C). We also examined RFX expression patterns in the adult human cortex (Figure 2.3D-F, Figure A.14A-B)[275,208]. Again, *RFX3* and *RFX7* exhibited the highest expression. *RFX3* was most highly expressed in glutamatergic layer 2/3 neurons, followed by astrocytes. *RFX7* was expressed in both inhibitory and excitatory neurons. *RFX4* expression was much lower overall, but highest in astrocytes (Figure 2.3F). These expression profiles suggest that RFX deleterious variants may lead to our observed neurodevelopmental phenotypes by altering early developmental cell fates or by impacting the function of upper-layer cortical neurons, astrocytes, and interneurons.

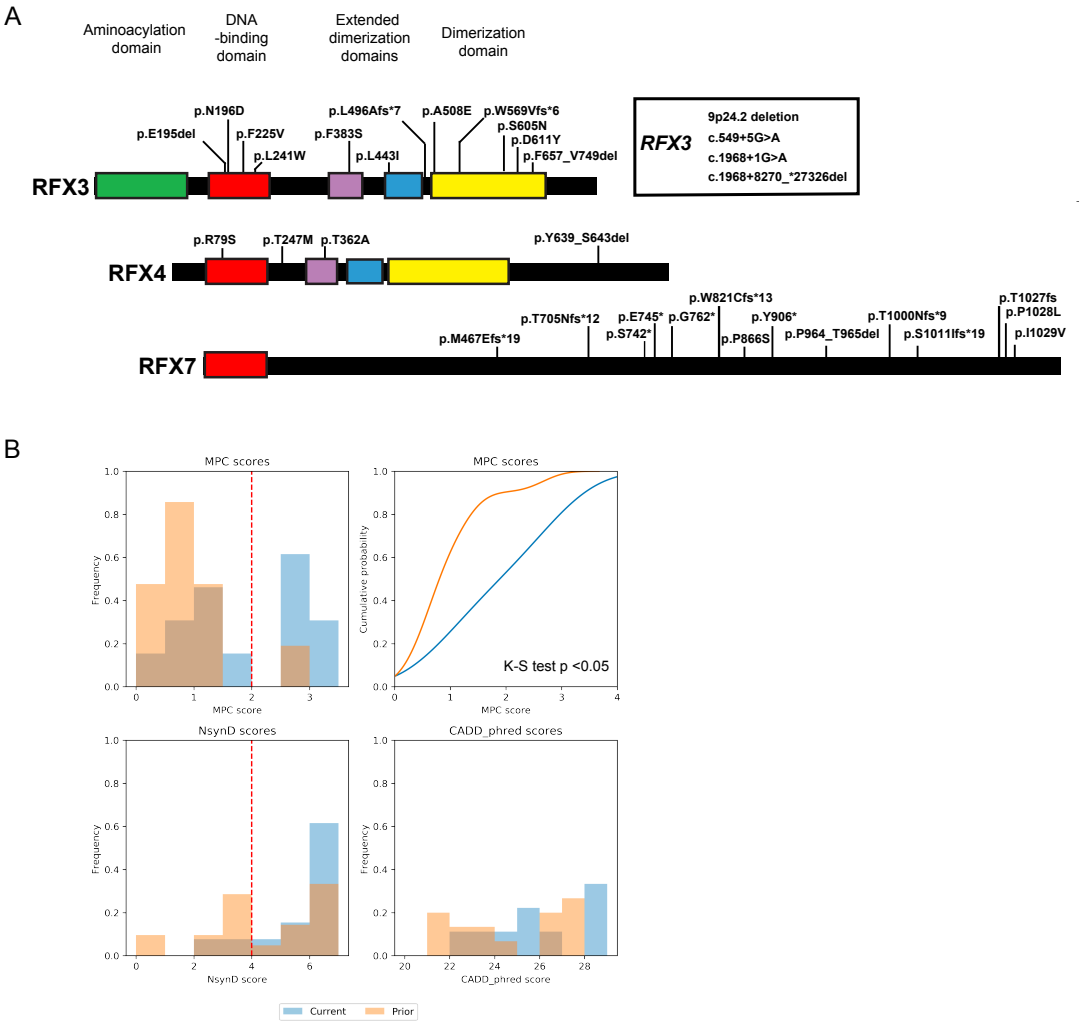


Figure 2.2: Distribution and predicted deleteriousness of RFX variants. A. Mapping of selected RFX variants to domains. Whole gene deletion and intronic variants are not illustrated. *RFX3* (NP_602304.1), *RFX4* (NP_998759.1), *RFX7* (NP_073752.5). B. Missense variant deleteriousness scores for the currently reported variants (current) and prior reported variants (prior) in *RFX3*, 4, and 7. The distribution of MPC scores for missense variants reported in this study is significantly different from that of prior reported missense variants, Kolmogorov-Smirnov (K-S) test p-value <0.05 (p-value=0.015). MPC, Missense badness, PolyPhen-2, and Constraint. NsynD, Nonsynonymous Damaging score. CADD, Combined Annotation Dependent Depletion.

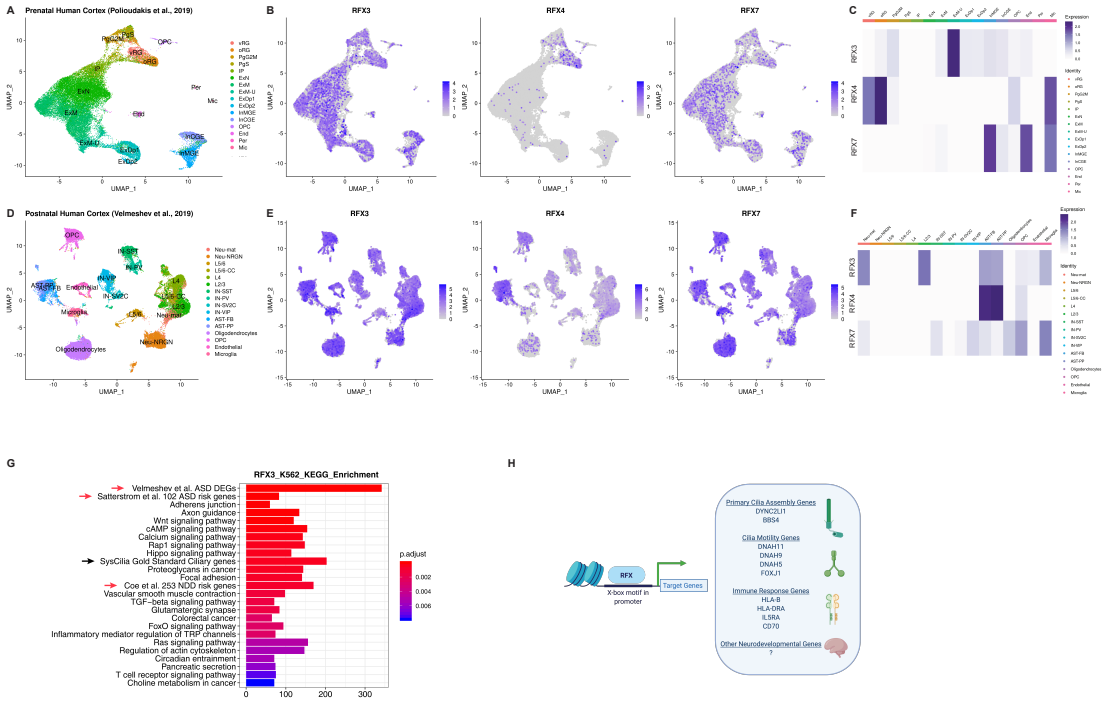


Figure 2.3: *RFX3*, *RFX4*, and *RFX7* expression patterns in human cortex and haploinsufficiency gene dosage model. A. Transcriptomic cell types in the prenatal human cortex identified by single-cell RNA- sequencing.24 B. *RFX3*, 4, and 7 expression patterns in single cells of the prenatal human cortex. C. Heatmap of *RFX3*, 4, and 7 expression levels among cell types in the prenatal human cortex. D. Transcriptomic cell types in the postnatal human cortex identified by single-cell RNA- sequencing.11 E. *RFX3*, 4, and 7 expression patterns in single cells of the postnatal human cortex. F. Heatmap of *RFX3*, 4, and 7 expression levels among cell types in the postnatal human cortex. G. The enrichment of KEGG pathways, ciliary genes, ASD risk gene sets, and ASD differentially expressed genes (DEGs) among *RFX3* ChIP-seq binding targets. Pathways and ASD gene sets are ranked by their statistical significance (p.adjust values, Benjamini-Hochberg's correction). Red arrows indicate ASD risk gene sets and ASD DEGs. X-axis shows the number of genes bound by RFX in their promoter regions. H. Binding of RFX family transcription factors bind to X-box motif in promoter regions of ciliary and immunologic genes. Target gene lists obtained from Piasecki, Durand, Reith, Sugiaman-Trapman.3,37-39 Model of *RFX* gene dose-dependent regulation of genes. In tissues with higher expression of *RFX* genes, ASD genes are activated. Lower levels of *RFX* genes are sufficient to activate ciliary genes. vRG, ventricular radial glia. oRG, outer radial glia. PgG2M, cycling progenitors G2/M phase. PgS, cycling progenitors S phase. IP, intermediate progenitors. ExN, migrating excitatory. ExM, maturing excitatory. ExM-U, maturing excitatory upper enriched. ExDp1, excitatory deep layer 1. ExDp2, excitatory deep layer 2. InMGE, interneuron MGE. InCGE, interneuron CGE. OPC, oligodendrocyte precursor cells. End, endothelial. Per, pericyte. Mic, microglia. Neu-mat, immature neurons. Neu-NRGN, NRGN expressing neurons. L5/6, layer 5/6 excitatory neurons. L5/6-CC, layer 5/6 excitatory cortico-cortical projection neurons. L4, layer 4 excitatory neurons. L2/3, layer 2/3 excitatory neurons. IN-SST, somatostatin interneurons. IN-PV, parvalbumin interneurons. IN-SV2C, SVC2 expressing interneurons. IN-VIP, VIP interneurons. AST- FB, fibrous astrocytes. AST-PP, protoplasmic astrocytes. OPC, oligodendrocyte precursor cells.

2.2.5 *RFX* binding motifs are present in ASD risk gene cis-regulatory regions

Dysregulated gene expression, especially in upper layer cortical neurons, has been implicated in ASD pathogenesis[275,256]. Given the expression of *RFX3* in layer 2/3 neurons and the autistic features of individuals reported here, we considered whether *RFX* family genes might be important transcriptional regulators of ASD risk genes. *RFX* family transcription factors bind to a characteristic consensus motif called an X-box (GTHNYY AT RRNAAC)[62] with individual family members having additional specificity for particular subsequences within this consensus. We therefore performed *RFX3*, 4, and 7 motif enrichment analysis in upstream regulatory sequences of 187 ASD risk genes (the union of 102 TADA genes from Satterstrom et al., 2020 and 124 genes meeting exome-wide significance from Coe et al, 2019)[232,45] and an additional set of 447 genes identified to be upregulated in ASD brains[275]. We found enrichment of X-box motifs (q-value <0.05) in human ESC-neuron specific enhancers for ASD risk genes (Table S2.5A). As a group, *RFX3* and *RFX4* motifs were particularly enriched (q-value <0.005), while the *RFX7* motif was not (q-value 0.48). X-box, *RFX3* and *RFX4* motifs were similarly enriched in the enhancer regions of genes upregulated in ASD brains (Table S2.5B). Enrichment of *RFX* motifs in promoter regions of ASD risk genes and DEGs did not emerge (data not shown). Last, we analyzed available *RFX* ChIP-seq data from the ENCODE project (Table S2.6) to determine enrichment for Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, ciliary genes, ASD risk gene sets, and ASD DEGs (Table S2.7). *RFX* functional binding genes from most ENCODE cell lines were significantly enriched in ASD risk genes and DEGs after multiple testing correction (p.adjust < 0.05; Benjamini Hochberg's correction; Figure 2.3G, Figure A.16, Table S2.8). Across cell lines, there was a positive correlation between enrichment in ASD genes and *RFX* expression levels in that cell type (Figure A.17, Table S2.9), indicating that higher *RFX* expression levels may be required to engage ASD relevant targets.

Finally, single gene analyses showed enrichment of *RFX3* and *RFX4* motifs in the promoters

of five ASD-associated genes (FIMO p-value <0.0001, q-value <0.1): *AP2S1, KDM6B, ANK2*, *NONO*, and *MYT1L* (Figure A.15B, D),[78] and *RFX3* ENCODE ChIP-seq data from HepG2 cells confirmed *RFX3* binding peaks in the promoters of *AP2S1, KDM6B,* and *NONO* (Figure A.15E-G). Notably, *de novo* loss-of-function variants in *KDM6B* (MIM 611577) cause a neurodevelopmental syndrome that has phenotypic overlap with *RFX3* haploinsufficiency as described in this report, namely mild global delays, delayed speech, hypotonia, and features of ASD and ADHD, while loss-of-function variants in *NONO* (MIM 300084) and *MYT1L* (MIM 613084) are a cause of X-linked and autosomal dominant intellectual disability, respectively (Table S2.4). These cases support the model that *RFX* members may be transcriptional activators of a subset of ASD risk genes via actions at both enhancer and promoter sites.

2.3 Discussion

Our results delineate a novel human neurobehavioral phenotype including ASD, ID and/or ADHD due to deleterious variants in *RFX* family transcription factors. While presence of neuroimaging findings, seizures, and dysmorphisms varied between different RFX family members, the behavioral phenotypes of individuals with *RFX3*, *RFX4*, and *RFX7* were strikingly similar, and often included sensory hypersensitivity and impulsivity. Like ID/DD and ASD more generally,[211] individuals with *RFX* variants also exhibited a male bias.

This report complements accumulating statistical genetic evidence for *RFX3* as an ASD risk gene[232] and extends these findings to the closely related *RFX* family members *RFX4* and *RFX7*. Two-thirds of individuals with *RFX3* variants in our series carried an ASD diagnosis, half had ADHD, and just over half of individuals had ID. Several individuals with *RFX3* variants also exhibited post-pubertal cognitive or behavioral regression sometimes accompanied by psychosis. *RFX3* CNVs have been previously reported in schizophrenia[278,225]. *RFX3* also lies within the region of

the chromosome 9p deletion syndrome (OMIM#158170), associated with developmental delay, ID, and ASD, although the size of the deletions in this syndrome make *RFX3* unlikely to be the sole contributor.

This report also implicates both *RFX4* and *RFX7* as causes of human neurodevelopmental disorders. Individuals with *RFX4* or *RFX7* variants were somewhat more severely affected than those with *RFX3* variants, with *RFX7* less likely to be associated with ASD or ADHD, but showing almost uniform diagnoses of language delay and ID (92%). There were fewer individuals identified with *RFX4* variants, but those identified had high rates of ASD and ID.

RFX family members have been previously known for their biological roles in cilia development. The *RFX3* transcription factor activates core components necessary for development and maintenance of both motile and primary cilia[64,43,25], and biallelic *RFX3* knockout in mice results in situs inversus, hydrocephalus, and deficits in corpus callosum formation[25,23,12]. This raises the question of whether the neurodevelopmental phenotypes reported here may be mechanistically related to cilia development – e.g., a hypomorphic human ciliopathy. The majority of described genetic ciliopathies are recessive, and therefore not due to haploinsufficiency, but some (e.g. Meckel syndrome, Joubert syndrome, Bardet-Biedl syndrome, oral-facial-digital syndrome type I) may be associated with neurodevelopmental abnormalities and/or brain malformations. On the other hand, the individuals described in this report lack systemic features of ciliopathies, suggesting the alternative hypothesis that *RFX* haploinsufficiency may directly dysregulate the expression of ASD risk genes while leaving ciliary genes intact (Figure 2.3H). Future work, which may include analyzing cilia morphology and function in cellular or animal models of *RFX* haploinsufficiency, or characterizing transcriptome-wide effects of *RFX* gene disruption, may prove helpful in distinguishing these hypotheses.

Enriched expression of *RFX3* in upper cortical layer neurons places this gene in cells that are involved in communication between regions of the cortex important for higher cognition and social

behavior[247], raising the possibility that haploinsufficiency may disrupt either the developmental specification, synaptic connectivity, or electrophysiological function of this set of neurons. Projection neurons in this layer have been implicated in ASD by analyses of co-expression networks of autism genes[284,201], and superficial cortical neurons exhibit the strongest amount of differential gene expression in ASD brains compared to controls[275,256]. Sun and colleagues in fact showed strong enrichment of *RFX* motifs in differentially acetylated peaks upregulated in ASD brains compared to controls[256]. Future studies aimed at understanding the downstream targets of *RFX* family members in human brain may shed new light on pathways important to the molecular pathogenesis of ASD, ADHD, and ID.

2.4 Acknowledgements

We thank Billie Lianoglou, MS, CGC (Fetal Treatment Center, University of California at San Francisco Fetal Treatment Center) for her contributions of an *RFX4* variant of unclear significance (see Supplementary). We thank the genetic counselors at GeneDx (Kirsty McWalter, MS, CGC and Erin Torti, MS, CGC) and Ambry Genetics (Meghan Towne, MS, CGC, LGC, Zoe Powis, MS, CGC, and Deepali Shinde, PhD) for their contributions including facilitation of clinician communication. JL was supported by award T32GM007753 from the National Institute of General Medical Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health. BZ was supported by the Manton Center Pilot Project Award and Rare Disease Research Fellowship. JEP and JRL were supported in part by the National Human Genome Research Institute (NHGRI) and National Heart Lung and Blood Institute (NHBLI) and the Baylor-Hopkins Center for Mendelian Genomics (BHCMG, UM1 HG006542). KU, SAB, CF, and SR were supported by the French Ministry of Health and the Health Regional Agency from Bretagne,

Pays de la Loire and Centre Val de Loire (HUGODIMS 2, 2017). The DDD study is supported by the Wellcome Trust and the UK Department of Health Innovation Challenge Fund [HICF-1009-003] and the Wellcome Trust Sanger Institute [grant no. WT098051] (Nature 2015;519:223-8). TWY was supported by grant nos. NIH/NIMH R01MH113761, NICHD/NHGRI/NIH U19HD077671 and NIH/NICHD U24HD0938487, and by a SFARI Pilot Research Award.

2.5 Author Contributions

TWY, HKH, and TN conceptualized the paper, collected case information from collaborators, conducted and managed all functional studies, drafted the initial manuscript, and edited and revised the manuscript. TWY provided research study oversight. JL designed and performed RFX motif analysis, analyzed published brain transcriptome and single-cell RNA-sequencing data, conducted variant analyses, contributed to the manuscript, and edited and revised the manuscript. BZ collected ChIP-seq data and performed over-representation analysis, contributed to the manuscript, and edited and revised the manuscript. NA and CSG conducted functional studies of the impact of *RFX3* variants on protein stability. AS performed RFX variant analyses. CAG assisted in research enrollment. LR, RP, TG, BBAdV, MEHS, KLIvG, EvB, C(N)MLV, AH, CDA, LLI, CB, MW, EF, TLT, KWG, LB, FV, PR, XW, JLA, MF, GET, JEP, JRL, EA, AN, RA, ARa, PB, CRF, MJL, MK, GL, AL, AP, KKP, LEW, KA, JB, CS, JM, CPB, GP, PG, MB, SK, MN, IGR, MYZ, CK, ARe, MI, KU, SA, CF, SR, MI, PDT, JB, YW, GZ, SS, IB, RAJ, WBD, AB, and LLC contributed clinical case information and/or analyzed exome data. PBA and AB supported research subject enrollment.

2.6 Supplemental Data

Main tables and all supplemental data can be found in the supplemental materials associated with the publication Harris et al., 2021 and include detailed clinical descriptions for each affected indi-

vidual.

2.7 Materials and Methods

2.7.1 Case Ascertainment and Data Collection

We obtained phenotypic data from 15 unrelated individuals with loss-of-function variants in *RFX3*, 4 unrelated individuals with loss-of-function variants in *RFX4*, and 14 unrelated individuals with loss-of-function variants in *RFX7*. Individual case summaries for all individuals are provided (Supplemental Data). Variants arose *de novo* with the exception of four related individuals from the same nuclear family with the same heterozygous loss-of-function variant in *RFX3*, and three other related cases in *RFX4* (homozygous for an inherited missense variant). Pedigree information and contributed photographs are shown in Figure 1. Diagnoses of ASD were reported in the medical record, but not uniformly evaluated by standardized measures such as the Diagnostic and Statistical Manual, Fourth or Fifth Edition (DSM-IV and 5), Autism Diagnostic Observation Schedule (ADOS), or Autism Diagnostic Interview, Revised (ADI-R). Similarly, ID and ADHD diagnoses were accepted per clinician report and not always accompanied by standardized cognitive or behavioral testing measures.

2.7.2 Exome Sequencing

Individuals included underwent exome sequencing on a clinical or research basis. Seven of the individuals were sequenced through GeneDx using genomic DNA from the proband or proband plus parents, captured using either the Clinical Research Exome kit (Agilent Technologies, Santa Clara, CA) or the IDT xGen Exome Research Panel v1.0, and sequenced on an Illumina system with 100bp or greater paired-end reads. Reads were aligned to human genome build GRCh37/UCSC hg19, and variants were analyzed and interpreted as previously described using variant classification

criteria publicly available on the GeneDx ClinVar submission page (see Web Resources). Two cases of *RFX7* were sequenced through Ambry Genetics whose gene and variant classification process are available on the AmbryGenetics web page. The remainder of the individual's exome sequencing was performed through the clinicians' institutions or an external laboratory or research program (see Acknowledgments).

2.7.3 Variant Analyses

Variant genomic coordinates are reported in relation to the Human Dec. 2013 (GRCh38/hg38) Assembly. The reference mRNA and protein sequences used are *RFX3* NM_134428.2, NP_602304.1; *RFX4* NM_213594.2, NP_998759.1; and *RFX7* NM_022841.5, NP_073752.5. The variant databases gnomAD v2.1.1 and v3 were examined for the presence of each variant[120] Predictions of the functional effects for all variants were assessed using MutationTaster, SIFT, PolyPhen2, PROVEAN, LRT, and MutationAssessor, and the total number of algorithms out of six with a deleterious prediction is referred to as the Nonsynonymous Damaging score (NsynD) as previously described[57]

2.7.4 Cell transfection and culture

Human *RFX3* (NM_134428.2; Human *RFX3* cDNA) was cloned into V5-tagged mammalian expression vectors using the Gateway cloning system (Thermo Fisher Scientific). Point mutations were introduced with the QuikChange Lighting Site-Directed Mutagenesis kit (Agilent Technologies) to incorporate variants from affected individuals. To quantify the expression level of exogenous *RFX3*, equal amounts of tagged-*RFX3* expression vectors were transfected into Hela cells using Lipofectamine 3000 (Thermo Fisher Scientific). The transfected cells were cultured for 48 hours before harvesting. Cell extracts were analyzed by immunoblotting, using antibodies raised

against *RFX3* (HPA035689, Sigma-Aldrich), V5 (R960-25, Thermo Fisher Scientific), or beta actin (ab6276, Abcam). Blots were scanned on a Li-Cor Odyssey imager (Li-Cor). Signal intensities were quantified using Image Studio Lite (Li-Cor). Each immunoblot analysis was replicated six times. One-way ANOVA by repeated measures was employed. Multiple comparison correction was performed by using Dunnett statistical testing.

2.7.5 KEGG pathway and ASD gene set over-representation analysis

ChIP-seq and eCLIP-seq narrowPeak bed files for RFX family members, CREBBP, EP300, FMR1, FXR1, and FXR2 were obtained from the ENCODE portal,[49] and additional ChIP-seq data for *RFX3*_K562 were obtained from RegulomeDB[28] (Table S2.6). Functional binding genes (1 kb upstream/downstream of TSS [transcriptional start site]) were annotated using ChIPseeker[295]. ASD risk gene lists included 102 TADA genes from Satterstrom et al. and 253 ASD/ID genes from Coe et al. (Table S2.7)[232,45] Differentially expressed genes (DEGs) in ASD brains were extracted from Velmeshev et al. excluding endothelial DEGs that could originate from vascular cells in the brain (Table S2.7)[275]. The SYSCILIA Gold Standard (SCGSv.1) was used as a gold standard of known ciliary genes in human[271]. Customized KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analysis was performed using clusterProfiler[294] to determine the enrichment for KEGG pathways, ciliary genes, ASD risk gene sets, and ASD DEGs. Multiple testing correction was performed using Benjamini-Hochberg correction (Table S2.8). Annotations, statistical analyses, and plots were implemented in R.

2.7.6 Motif Analyses

For motif occurrence analysis, FIMO was used to scan promoter sequences for individual occurrences of RFX motifs[78]. For all analyses, motif models were obtained from the JASPAR 2020

database[71]. Motifs searched for included *RFX3* (MA0798.1), *RFX4* (MA0799.1), and *RFX7* (MA1554.1). Promoter sequences were defined as -1000 base pairs and +500 base pairs relative to the transcription start site. Motif occurrences were classified as significant based on a reporting threshold of p-value <0.00001 and q-value (Benjamini) <0.10. For motif enrichment analysis, we used the HOMER findMotifs.pl and findMotifsGenome.pl scripts. Motifs were classified as enriched based on fold-enrichment > 1.5 over randomly selected background sequences with matched GC% content, and q-value (Benjamini) <0.01. Enhancer sequences associated with genes of interest were obtained from Enhancer Atlas 2.0[73]. ASD risk gene lists were obtained as noted previously to include 102 TADA genes and 124 ASD/ID genes reaching exome-wide significance[2,32,45].

"The mind is the result of the activities of the brain."

John Searle

3

Multi-omic dosage analysis of the ciliogenic transcription factor *RFX3* reveals a novel role in modulating activity-dependent responses via enhancing CREB binding in human neurons

This section is based on the following manuscript in which I made significant contributions during my PhD studies:

Lai J, Demirbas D, Phillips K, Zhao B, Wallace H, Seferian M, Nakayama T, Harris HK, Chatzipli A, Lee EA, Yu TW. Multi-omic dosage analysis of the ciliogenic transcription factor *RFX3* reveals a novel role in modulating activity-dependent responses via enhancing CREB binding in human neurons. Manuscript under peer review. 2023.

3.1 Introduction

Autism spectrum disorder (ASD) is a highly heritable and genetically heterogeneous condition characterized by deficits in social communication and repetitive behaviors. De novo loss-of-function variants in over 100 genes are recognized as significant contributors to ASD risk[232]. Implicated

genes are broadly involved in gene expression regulation and neuronal communication [232]. Epigenetic and transcriptional changes in postmortem brain tissue from individuals with ASD point to a functional convergence in synaptic transmission [256,275,72]. Yet, inaccessibility of brain tissue in human neurodevelopment limits the understanding of molecular processes that contribute to neurodevelopmental disorders. Functional neurobiological characterization of ASD risk genes in experimentally manipulable human model systems can provide insight to underlying disease mechanisms.

RFX3 encodes a member of the RFX family of transcription factors that is characterized by a highly conserved winged-helix DNA binding domain that recognizes the X-box motif [255]. Haploinsufficiency of *RFX3* leads to a human neurodevelopmental disorder with features including ASD, ADHD, and behavioral dysregulation [85]. However, the molecular and cellular basis for the impact of *RFX3* deficiency on human brain development and function are not understood. Previous studies of *RFX3* and its binding sites in worms and mouse models have shown that *RFX3* regulates the expression of ciliary genes, including those involved in recessively inherited human ciliopathies like Bardet-Biedl syndrome, Alstrom syndrome, and Meckel-Gruber [258,64,13]. Mouse *RFX3* knockouts exhibit defective ciliary assembly and function that results in hydrocephalus, situs inversus, and corpus callosum malformation [146,25,23,165]. *RFX3* knockout mice also fail to undergo proper differentiation of auditory hair cells as well as pancreatic beta-islet cells [3,65].

Of note, previous biological roles ascribed to RFX genes (ciliogenesis, hair cell development, beta-islet cell development) have emerged from studies of complete knockout animals (e.g., bearing disruptive mutations in both copies of the gene), while haploinsufficient phenotypes have not been described. This raises the question as to whether the neurodevelopmental syndrome observed with heterozygous *RFX3* mutations in humans can be ascribed to its previously described molecular and cellular functions – for instance, a "mild ciliopathy," since some recessive human ciliopathies are associated with intellectual disability. Alternatively, the neurobehavioral impacts of *RFX3* haploin-

sufficiency could reflect disruption of novel roles and downstream targets.

To address this question, we used human iPSC-derived neuronal cultures and organoids to investigate the impacts of altering *RFX3* gene dosage on human neuronal development and function. We present evidence that in addition to its known role as a direct transcriptional activator of ciliary genes, RFX3 binds cis-regulatory elements of synaptic genes and regulates their expression in a *RFX3* gene dosage-sensitive manner. Changes in *RFX3* gene dosage altered synaptic formation and spontaneous neural activity. We also present evidence that *RFX3* facilitates CREB binding at the promoters of activity-dependent genes, and is important for their induction in response to neuronal depolarization. This study provides insights into the regulatory roles of *RFX3* on neurodevelopment and synaptic activity and highlights how the tuning of activity dependent responses may underlie a subset of neurodevelopmental disorders.

3.2 Results

3.2.1 *RFX3* regulates neurogenesis in forebrain organoids

RFX3 gene expression is enriched in the developing brain in progenitors and maturing excitatory neurons[85], suggesting possible roles in neuron differentiation or function. To investigate this further, we generated dorsal forebrain organoids[24] from isogenic human iPSCs with CRISPR-Cas9 engineered heterozygous (HET) and homozygous (KO) loss-of-function (LoF) variants in *RFX3* in a control (WT) iPSC line (A.18A). We introduced frameshift variants into exon 5, upstream of the DNA binding domain (NM_001282116.2:c.313del/ins, c.335ins). Consistent with this resulting in a null allele, Western blot analysis demonstrated ~50% reduced RFX3 protein levels in the HET iPSCs and no detectable protein in the KO iPSCs (A.18B).

We then characterized forebrain organoids bearing 0, 1, or 2 copies of the *RFX3* frameshift variant at days 45 and 90 in culture using single-cell RNA-sequencing (scRNAseq) to identify alter-

ations in neurodevelopmental trajectories and cell type specific transcriptional changes (3.1A). We profiled the gene expression of 114,962 cells, from a total of 34 control and mutant forebrain organoids across the two time points (A.18C-D). We correlated the transcriptomes of the organoids at these timepoints with those of the developing human brain from the Brainspan database [119]. Overall, day 45 and day 90 organoid transcriptomes were most correlated with the early-mid gestation stage (8-21 post-conception weeks), which coincides with the period when *RFX3* expression is the highest, and corresponds to periods of progenitor proliferation and neurogenesis [7] (A.18E). *RFX3* was broadly expressed in progenitors and neurons in the organoids at these two timepoints (A.18F).

Following integration and annotation of cell types using an ensemble of references from human fetal brain and forebrain organoid atlases [208,218,274], we identified five major cell types, including cycling radial glia (cycling_RG), outer radial glia (oRG), intermediate progenitors (IP), deep layer excitatory neurons (ExDL), and upper layer excitatory neurons (ExUL) (3.1B). Over time, the organoids recapitulated the main stages of human corticogenesis, including progenitor expansion and neurogenesis. At day 45, the organoids had developed cycling and outer radial glia (*MKI67+, ASPM+, TNC+, PTPRZ1+*) and deep layer excitatory neurons (*FOXP2+*). By day 90, intermediate progenitors and upper layer neurons (*SATB2+, CUX1+*) emerged (3.1B). Each organoid reproducibly generated each cell type, and the expression of top cell type marker genes was highly similar across individual control organoids (3.1C).

3.2.2 Biallelic loss of *RFX3* leads to delayed neurogenesis

We asked whether *RFX3* disruption may impact developmental cell fates. Comparison of cell type proportions in control versus *RFX3* deficient organoids at day 45 and day 90 demonstrated delayed generation of neurons in *RFX3* KO organoids (3.1D). At day 45, *RFX3* KO organoids had a significantly increased proportion of cycling radial glia and outer radial glia, accompanied by a

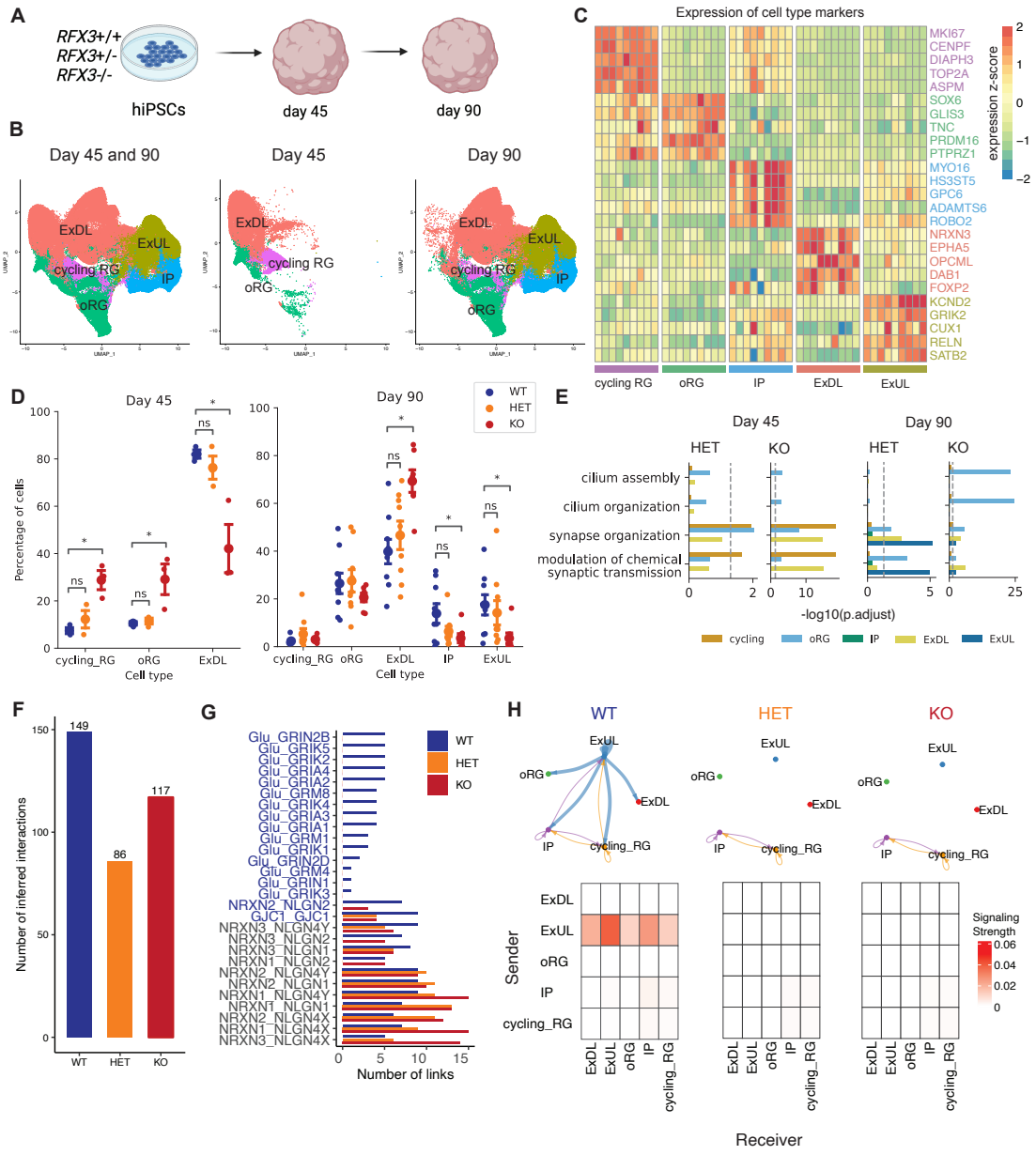


Figure 3.1: Altered glutamatergic signaling in *RFX3* deficient dorsal forebrain organoids. A. Schematic of dorsal forebrain organoid development. B. UMAP visualization of five major cell types identified in dorsal forebrain organoids over development. C. Top five marker genes per cell type in day 90 WT organoids. Each column corresponds to an individual organoid replicate (n=9). Color bar represents z-scaled expression. D. Cell type proportions per organoid at day 45 and day 90. Data represents mean +/- SEM. * 95% credible intervals do not overlap, indicating a significant difference in relative abundance. E. Ciliary Gene Ontology (GO) term enrichment among significantly downregulated genes in *RFX3* HET and KO cell types.

Figure 3.1 (Continued). F. Number of significant inferred interactions in WT, HET, and KO organoids (permutation test, n=1000 permutations, p-value<0.05). G. Ranking of signaling pathways by number of links in WT, HET, and KO organoids. Pathways in blue indicate significantly increased in WT organoids compared to HET and KO (Wilcoxon test, p-value<0.05). H. Top, circle plot showing aggregated number of interactions for the significant pathways in (G) in WT, HET, and KO organoids. Arrow width indicates number of interactions. Bottom, heatmap of aggregated communication strength for the significant pathways in (G) in WT, HET, and KO organoids.

significantly decreased proportion of deep layer neurons. At day 90, *RFX3* KO organoids had a significantly increased proportion of deep layer neurons and decreased proportion of upper layer neurons (3.1D). This suggests a temporal delay in corticogenesis in organoids with biallelic loss of *RFX3*. Notably, specification and development of cell types was not significantly altered in *RFX3* HET organoids compared to isogenic controls.

3.2.3 Cell type-specific transcriptional alterations by *RFX3* dosage

RFX3 expression is maintained in the postnatal brain into adulthood, and enriched in upper layer excitatory cortical neurons that engage in cortical-cortical connectivity and demonstrate altered gene expression and function in models of ASD [108,275,68,223,293,182,85]. Brain magnetic resonance imaging (MRI) of individuals with *RFX3* haploinsufficiency fail to reveal consistent gross anatomical abnormalities [85]. This raised the possibility *RFX3* haploinsufficiency may instead alter gene expression networks critical to postnatal neuronal function.

Given that *RFX3* is well-known for its role in ciliogenesis, we first asked whether ciliary gene expression is disrupted in organoids with single or biallelic LoF variants in *RFX3*. To understand the cell type specific transcriptional perturbations caused by *RFX3* deficiency, we identified genes differentially expressed in each cell type in *RFX3* HET or KO organoids, compared to WT (A.19,A.20, Table S3.1). Since *RFX3* is a transcriptional activator, we focused on downregulated genes as primary transcriptional alterations induced by *RFX3* deficiency. This revealed that at both day 45 and day 90, *RFX3* KO organoids demonstrated significant disruption of ciliary assembly and organiza-

tion genes, specifically in oRG, that was more pronounced by day 90.

In contrast, *RFX3* HET organoids did not show significant dysregulation of these ciliary processes in any cell type (3.1E, Table S3.2). Instead, genes downregulated in *RFX3* HET organoids were strongly enriched for axon development, synapse organization, and modulation of chemical synaptic transmission. This pattern was recapitulated across several cell types and at both developmental timepoints (Figure S2-S3, Table S3.2). Furthermore, ligand-receptor (L-R) analysis using NeuronChat[301] revealed that the overall number of significant L-R interactions was decreased in *RFX3* deficient organoids (3.1F). Specifically, glutamatergic signaling (ligand: glutamate, receptors: *GRIN2B, GRIK5, GRIK2, GRIA4, GRIA2, GRM8, GRIK4, GRIA3, GRIA1, GRM1, GRIK1, GRIN2D, GRM4, GRIN1, GRIK3*) was significantly decreased in *RFX3* HET and KO organoids (3.1G). In aggregate, these significantly altered glutamatergic L-R pairs predominantly originate from ExUL in WT organoids, while no significant interactions were present between ExUL and other cell types in *RFX3* deficient organoids (3.1H). These results suggest that glutamatergic signaling, specifically involving ExUL neurons, is disrupted in *RFX3* deficiency.

3.2.4 Synaptic genes are sensitive to *RFX3* dosage

We used *Ngn2* induction of *RFX3* WT, HET, and KO iPSCs to explore the potential influence of *RFX3* on glutamatergic signaling further[299]. *RFX3* WT neuronal cultures demonstrated robust RFX3 expression assessed by Western blot at day 14, while levels were decreased in HET and KO neurons as expected (A.21A-B).

We profiled gene expression patterns of *RFX3* WT, HET, and KO neurons at day 14 in culture to identify differentially expressed genes (3.2A, Table S3.3). Functional enrichment analysis revealed that, similar to the dorsal forebrain organoids, genes downregulated in *RFX3* KO neurons were significantly enriched for ciliary processes (cilium assembly, cilium organization), whereas genes downregulated in *RFX3* HET neurons were not (3.2B, Table S3.4-S3.5). As a gross measure, the fraction

of ciliated neurons (marked ARL13B) was similar between RFX3 WT, HET and KO (A.21C-D). However, expression of a ciliary gene set (CiliaCarta, n=956)[270] was on average significantly reduced in *RFX3* KO neurons, but not in *RFX3* HET neurons (3.2C). Western blot analysis for ciliary protein ARL13B also revealed significant decrease in RFX3 KO neurons, but not in RFX3 HET neurons (A.21E-F). This suggests that ciliogenesis is generally intact in RFX3 haploinsufficiency.

Instead, the top enriched pathways among downregulated genes in *RFX3* HET and KO neurons were related to neurodevelopment and synaptic function, including forebrain development, modulation of chemical synaptic transmission, synapse assembly, and synapse organization (3.2D, Table S3.4-S3.5). Synaptic genes[131], as a set (SynGO, n=1233), had significantly reduced expression in *RFX3* HET and KO neurons (3.2E). Genes downregulated in *RFX3* HET and KO neurons were also enriched for ASD risk genes, defined by SFARI (n=913) and a recent large exome sequencing study[232](ASD102, n=102) (3.2F). We also analyzed the overall expression of ASD risk gene sets as well as gene sets for rare genetic intellectual disability (ID) (ORPHA:183757, n=1953), ADHD (n=24)[51,126,297], and schizophrenia (SCZ, n=164)[159]. ASD risk genes (ASD102 and SFARI sets) were significantly downregulated in both *RFX3* HET and KO neurons, while ADHD, ID, and SCZ risk gene sets were not significantly downregulated in either (3.2G). ASD risk genes with decreased expression in *RFX3* HET and KO neurons were most enriched for synaptic transmission, signaling, and organization processes, suggesting that *RFX3* regulates a subset of ASD risk genes primarily involved in neuronal communication (3.2I). These results corroborate the altered glutamatergic synaptic signaling in *RFX3* deficiency demonstrated by the organoids, further supporting an underappreciated role for *RFX3* in regulating synaptic processes which may underlie the neurodevelopmental phenotype caused by *RFX3* haploinsufficiency.

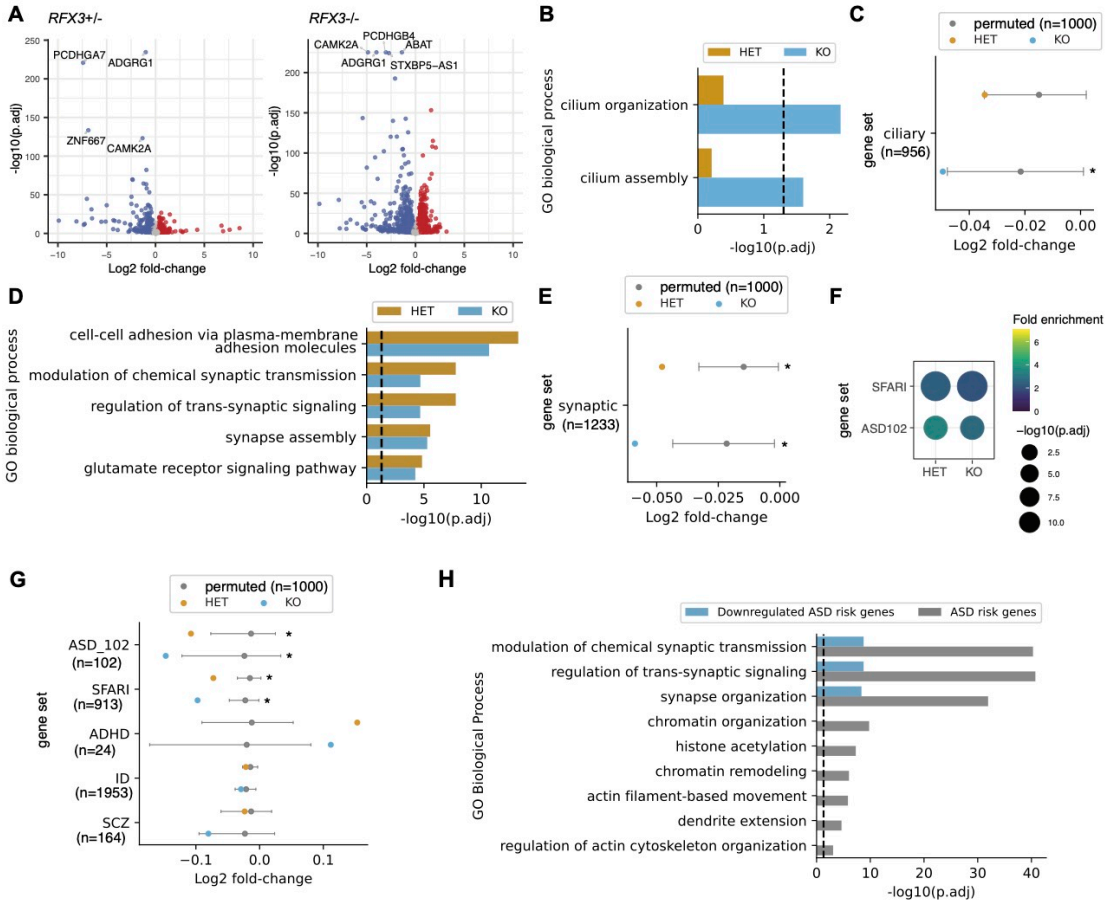


Figure 3.2: Synaptic gene expression is sensitive to *RFX3* dosage. A. Volcano plot showing differentially expressed genes (DEGs) between *RFX3* HET and WT neurons (left) and *RFX3* KO and WT neurons (right) at day 14 in culture. Genes significantly downregulated are in blue (log2FC<-0.25, FDR<0.05). Genes significantly upregulated are in red (log2FC>0.25, FDR<0.05). The top DEGs ranked by significance are labeled. B. Gene Ontology (GO) term enrichment analysis of significantly downregulated genes in *RFX3* HET and KO. Ciliary related GO terms shown. Dashed line indicates significance threshold FDR 0.05. C. Average observed log2 fold-change of ciliary gene set (n=956 CiliaCarta) in *RFX3* HET and KO neurons compared to permuted distribution for randomly sampled gene sets of equal size (n permutations=1,000, 90 percentile interval). *left-tailed p <0.05, permutation test. D. Gene Ontology (GO) term enrichment analysis of significantly downregulated genes in *RFX3* HET and KO. Top enriched GO terms related to neurodevelopment shown. Dashed line indicates significance threshold FDR 0.05. E. Average observed log2 fold-change of synaptic gene set (n=1233 SynGO) in *RFX3* HET and KO neurons compared to permuted distribution for randomly sampled gene sets of equal size (n permutations=1,000, 90 percentile interval). *left-tailed p <0.05, permutation test.

Figure 3.2 (Continued). F. Overrepresentation analysis of ASD risk genes among *RFX3* HET and KO downregulated genes. Dot size indicates significance of enrichment. Color indicates fold enrichment (fraction of overlapping genes compared to probability of obtaining overlap by chance). G. Average observed log2 fold-change of neuropsychiatric risk gene sets in *RFX3* HET and KO neurons compared to permuted distribution for randomly sampled gene sets of equal sizes (n permutations=1,000, 90 percentile interval). *left-tailed p <0.05, permutation test. ID, intellectual disability. SCZ, schizophrenia. ADHD, attention-deficit hyperactivity disorder. H. Gene Ontology (GO) term enrichment analysis of all ASD risk genes and those significantly downregulated in *RFX3* HET and KO neurons. Dashed line indicates significance threshold FDR 0.05.

3.2.5 RFX3 directly binds cis-regulatory elements of synaptic genes

Downregulation of synaptic genes in *RFX3* deficient neurons could occur as a direct effect of decreased RFX3 binding to the cis-regulatory elements of synaptic genes, or could be the result of indirect effects of *RFX3*, for instance on neuronal differentiation pathways. Morphological analysis did not reveal differences in total neurite length (marked by TUJ1) in *RFX3*+/- neurons (A.22A). Furthermore, RNA deconvolution analysis (aimed at estimating the maturity of progenitors and neurons in *RFX3* WT, HET and KO iPSC-derived neuronal cultures)[35] did not show significant differences in fractions of progenitors and neurons between genotypes over multiple time points, suggesting that differentiation of Ngn2 neurons was not impaired (A.22B).

To identify *RFX3* targets, we profiled RFX3 binding genome-wide in human neurons using CUT&RUN-sequencing. This yielded 4,024 RFX3 binding peaks in WT neurons that were not present in KO neurons (3.3A, Table S3.6). Motif enrichment analysis revealed that the RFX3 motif was the most enriched among these binding regions and was present in 86% of peaks (RFX3 motif E-value 3.70e-922, 3.3B). Motif co-occurrence enrichment analysis using SpaMo[283] identified binding motifs for SIX (AACCTGA) (E-value 5.11e-7) and CREB (TGACGTCA) (E-value 2.02e-6); these were the top enriched motifs that co-occur with the RFX3 motif with zero gap between the motifs (3.3C). SIX and RFX3 have been shown to cooperatively interact to regulate gene expression programs in auditory sensory epithelium[151], and RFX5 and CREB cooperatively bind the X2 box in MHC class II promoters[187]. This suggests that RFX3 may also form a complex with these

transcription factors to regulate gene expression in human neurons.

RFX3 target genes were enriched for ciliary related and synaptic related genes (3.3D). Promoter peaks (defined as those within 3000 nt of the adjacent gene transcription start site) were enriched for both ciliary processes and synaptic processes, while distal peaks (>3000 nt) were enriched for synaptic processes (3.3D). Consistent with RFX3 acting as a direct transcriptional activator, RFX3 target genes exhibited significantly decreased expression in *RFX3* deficient neurons (3.3E).

3.2.6 RFX3 exhibits dosage sensitive binding near synaptic genes

Next, we asked why synaptic gene expression is vulnerable to loss of one copy of *RFX3* while ciliary gene expression is more robust. To investigate potential mechanisms, we characterized RFX3 binding patterns, hypothesizing that differences in ciliary and synaptic gene expression patterns may simply reflect differential RFX3 binding strength for these genes. Overall, 947/4024 RFX3 binding sites (24%) demonstrated significantly decreased binding (FDR<0.2) in *RFX3* HET neurons compared to WT (referred to as "dosage sensitive sites") (3.3F, Table S3.7). Moreover, fold-changes in RFX3 peak intensity were significantly correlated with fold-changes in target gene expression (HET v. WT, Pearson r 0.108, p-value 1.02e-09, n=3,351 target genes) (3.3G). Functional enrichment analysis revealed that dosage sensitive targets (defined as genes with decreased RFX3 binding and decreased expression in *RFX3* HET neurons) were most enriched for axon guidance (i.e., *PLXNC1, NOG, BMPR1B*) and synapse organization (i.e., *CACNB2, SRGAP2, GRIN2B, CBLN2*), and were not enriched for ciliary pathways (3.3H).

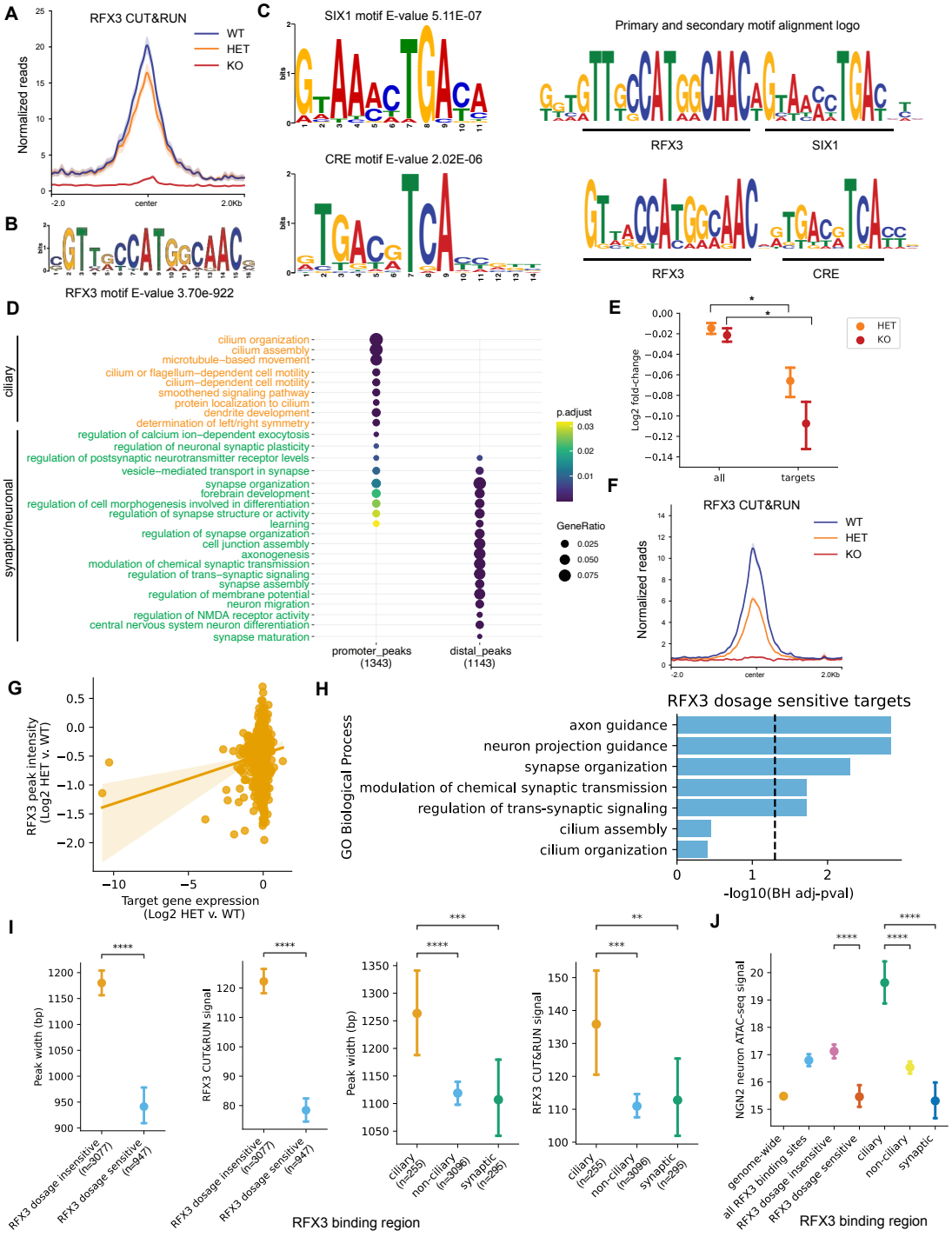


Figure 3.3: RFX3 exhibits dosage sensitive binding near synaptic genes. A. Aggregate plot of RFX3 CUT&RUN reads within RFX3 binding peaks called in WT neurons (n=4024 peaks). Trace shows mean +/- SEM of n=7 WT replicates, n=6 HET replicates, n=6 KO replicates.

Figure 3.3 (Continued). B. RFX3 binding motif enriched within RFX3 binding peaks. C. SIX1 binding motif enrichment near RFX3 binding motifs in RFX3 binding peaks. Primary (RFX3) and secondary (SIX1) motif alignment. CRE binding motif enrichment near RFX3 binding motifs in RFX3 binding peaks. Primary (RFX3) and secondary (CRE) motif alignment. D. Gene Ontology (GO) enrichment analysis of genes with an RFX3 peak in the promoter (promoter_peaks) or distal regions (distal_peaks). E. Log2 fold-change of all genes or RFX3 target genes in *RFX3* HET or KO neurons compared to WT. *p<0.05, t-test with Bonferroni correction. F. Aggregate plot of RFX3 CUT&RUN reads within RFX3 binding peaks significantly decreased in *RFX3* HET neurons (n=947 peaks, FDR<0.2). Trace shows mean +/- SEM of n=7 WT replicates, n=6 HET replicates, n=6 KO replicates. G. Log2 fold-change in RFX3 peak intensity v. log2 fold-change in target gene expression in *RFX3* HET neurons. Linear regression model fit with 95% confidence interval (CI) shown. H. Gene Ontology (GO) enrichment analysis of RFX3 target genes with significantly decreased peak and expression in *RFX3* HET neurons. Dashed line indicates significance threshold FDR 0.05. I. Peak width and RFX3 CUT&RUN signal among *RFX3* dosage sensitive, *RFX3* dosage insensitive, ciliary, non-ciliary, and synaptic regions. *p<0.05, **p<0.01, ***p<0.005, ****p<0.0001, t-test. J. ATAC-seq signal in Ngn2 neurons in different sets of RFX3 binding regions. Data represent mean +/- 95% CI. *p<0.05, **p<0.01, ***p<0.005, ****p<0.0001, one-way ANOVA with t-test with Bonferroni correction.

3.2.7 RFX3 has weaker binding near synaptic genes associated with decreased chromatin accessibility

We next explored why RFX3 binding sites near synaptic targets might be especially sensitive to genetic dosage, relative to binding sites near ciliary genes. Overall, in WT neurons, *RFX3* dosage sensitive sites had decreased peak width and height compared to dosage insensitive sites (3.3I). Correspondingly, ciliary genes had RFX3 peaks with increased peak width and height compared to non-ciliary genes and synaptic genes (3.3I). This suggested that RFX3 binding sites that are sensitive to genetic dosage may simply be those that bind less RFX3 at baseline. To break this down further, we considered whether RFX3 binding strength may be influenced by differences in motif sequences, or differences in chromatin accessibility. We found that ciliary and synaptic targets contained a similar number of RFX3 motifs (2 RFX3 motifs per binding region; A.22C), and exhibited comparable RFX3 motif scores (A.22D). In contrast, we analyzed ATAC-seq data from Ngn2 neurons[228] and found that dosage sensitive sites had significantly decreased chromatin accessibility compared to dosage insensitive sites (3.3J). RFX3 binding sites near synaptic genes also had significantly decreased chromatin accessibility compared to binding sites near ciliary genes (3.3J). These binding

patterns suggested that decreased chromatin accessibility is one of the factors responsible for decreased RFX3 binding affinity near synaptic targets (compared to ciliary targets), rendering synaptic targets more sensitive to *RFX3* gene dosage.

3.2.8 RFX3 binding motifs near ciliary and synaptic genes are evolutionarily conserved

The role of RFX transcription factors in ciliogenesis is evolutionarily conserved down to early metazoans[44], but less is known about the role of *RFX3* in synaptic processes. We asked whether *RFX3* regulation of synaptic genes in neurons is evolutionarily conserved or primate-specific. To address this question, we converted RFX3 bound regions into their homologous counterparts in the genomes of other species, and scanned for the presence of RFX3 binding motifs. We defined primate-specific RFX3 motif occurrences as those found in humans, chimpanzees, or macaques, and conserved RFX3 motif occurrences as those found in humans, primates, and at least one other non-primate species. We found that over 90% of RFX3 binding motifs in neuronal RFX3 binding sites were evolutionarily conserved (3291/3644) while 10% were specific to primates (353/3644). Conserved motifs were enriched near both ciliary and synaptic genes, suggesting that *RFX3* regulation of ciliogenesis and synaptic processes is evolutionarily conserved (A.22E). In contrast, primate-specific motifs were only enriched near genes involved in neurogenesis (A.22E). We next asked whether the conserved and primate-specific motif occurrences were enriched in different genomic regions. We found that conserved motifs had an increased proportion in promoter regions while primate-specific motifs had an increased proportion in distal regions (A.22F, p. adj <0.05, proportions z-test). This suggests that *RFX3* may have acquired primate-lineage specific roles at distal regulatory elements for additional spatiotemporal control of gene expression [130].

3.2.9 Transcriptional dysregulation of synaptic genes in *RFX3* haploinsufficiency is functionally associated with immature neuronal networks

Given that *RFX3* deficient neurons demonstrated decreased expression of genes involved in synaptic transmission and plasticity, we asked whether this transcriptional dysregulation has functional consequences on synaptic assembly and activity. We first quantified the number of excitatory synapses marked by SYN1 and PSD-95 colocalization. Synapse number was significantly decreased in *RFX3* deficient neuronal cultures compared to control (3.4A). We then used multielectrode arrays (MEA) to measure spontaneous synaptic activity of *RFX3* deficient neurons compared to isogenic WT controls from one week to eight weeks in culture. MEA captures complex neuronal network activity over time, where neuronal activity progresses from random spikes, to electrode bursts, to organized patterns of network wide bursts. Synchronized firing represents strong synaptic connections and plays a role in activity-dependent organization of cortical networks[242,268,33]. Synchronized neural networks increase over brain development, and are a hallmark of mature cortical circuits[125,189,202]. We found that overall activity, measured by mean firing rate, did not differ between *RFX3* WT, HET, and KO (3.4B). However, the number of spikes per network burst and synchrony were decreased in *RFX3* deficient neurons, indicating decreased synaptic strength and immature network activity (3.4C-D). We also adopted an established approach to summarize the metrics measured by the MEA into a single objective neural activity score (NAS) that correlates with time in culture and reflects network maturation[202]. *RFX3* HET and KO neurons had significantly decreased NAS compared to WT from day 39 onwards (3.4E). These results demonstrate that dysregulated synaptic gene expression in *RFX3* deficient neurons is functionally associated with altered synaptic number, decreased synaptic strength, and immature network activity.

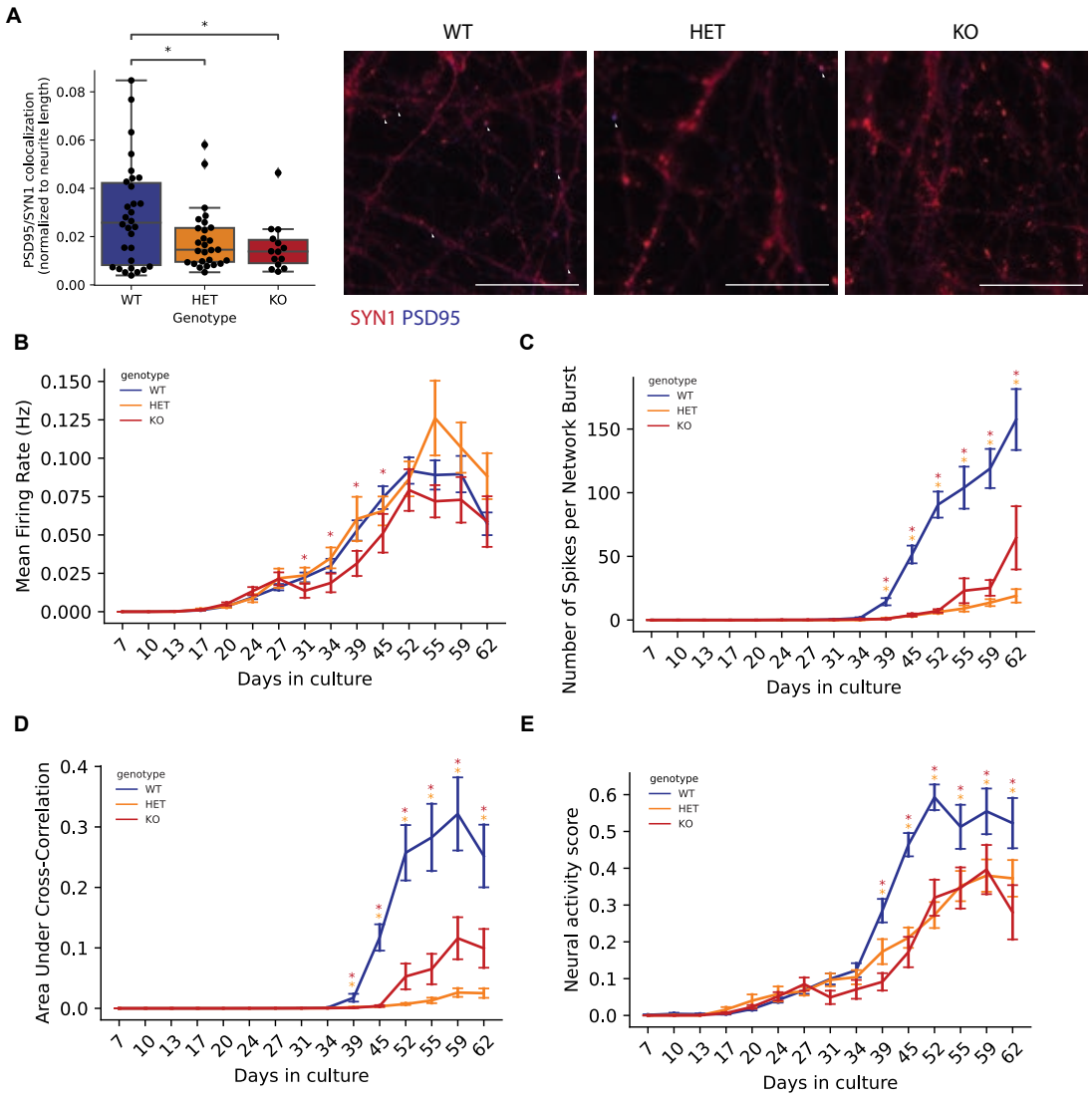


Figure 3.4: Decreased synchronized neural networks in *RFX3* deficient neurons. A. Quantification and representative images of SYN1 and PSD95 colocalized puncta on neurites in day 14 neuronal cultures. n=32 wells per genotype. White arrows mark representative colocalized puncta. Scale bar=100 um.

Figure 3.4 (Continued). B. Mean firing rate (Hz) of neuronal cultures from day 7 to day 62. C. Average number of spikes per network burst of neuronal cultures from day 7 to day 62. D. Area under cross-correlation (measure of synchrony) of neuronal cultures from day 7 to day 62. E. Neural activity score of neuronal cultures from day 7 to day 62. (A-D) Data normalized to the number of covered electrodes per well. Data represented as mean +/- SEM. n=32 wells per genotype from two independent MEA plates. *p<0.05, one-sided t-test, red * indicates KO v. WT, yellow * indicates HET v. WT (A-D), t-test (E) with Bonferroni correction.

3.2.10 *RFX3* modulates CREB-dependent activity-dependent transcriptional responses

Activity dependent responses are implicated in synaptic plasticity and synaptogenesis which underlies the development of functional neuronal networks, and are transcriptionally dysregulated in ASD[72,263]. *RFX3* binding sites were enriched for the binding motif of CREB, a key regulator of activity dependent gene expression programs[237] (3.3C). *RFX3* and CREB also showed evidence of protein-protein interaction in a large-scale mammalian two hybrid screen conducted in human cells[217]. Notably, *RFX3* exhibited binding near CREB target genes involved in activity-dependent signaling pathways, including the immediate early genes FOSB, JUN, and JUNB (3.5A). These findings suggested that *RFX3* may be a direct regulator of activity-dependent gene expression programs. Consistent with this notion, when we compared *RFX3* binding sites with KCl inducible H3K27ac regions in *Ngn2* neurons that mark activity-dependent regulatory elements[228], we found both the *RFX3* binding motif and *RFX3* binding sites to be significantly enriched (A.23A-B).

3.2.11 *RFX3* deficiency leads to blunted CREB-dependent activity induced gene expression

To test whether *RFX3* modulates activity-dependent transcription, we used RNA-seq to profile activity-dependent gene expression responses of day 14 *RFX3* WT, HET, and KO neurons after depolarizing with 55 mM potassium chloride (KCl) (3.5B). Previous studies showed that this treat-

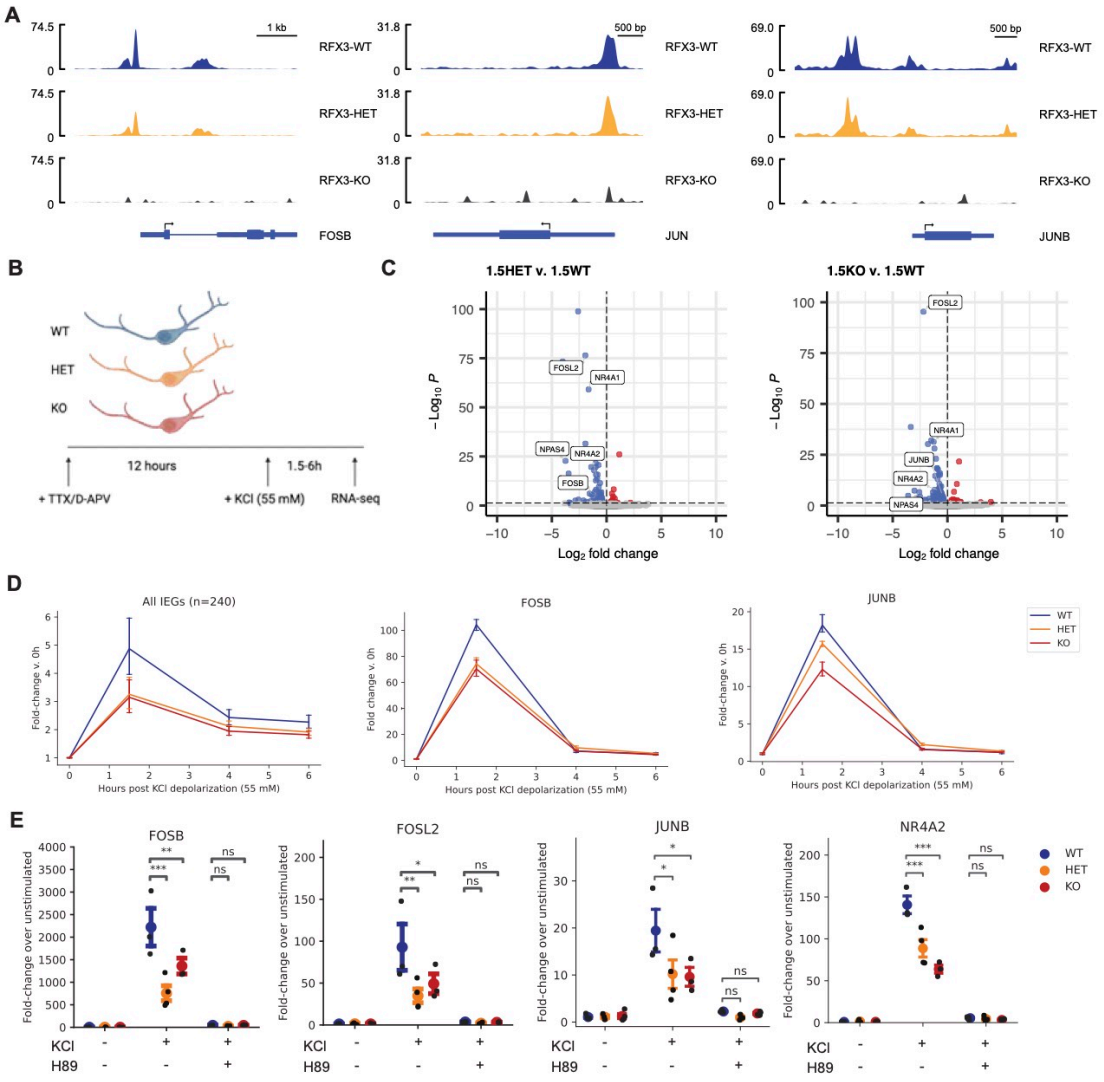


Figure 3.5: *RFX3* modulates CREB-dependent activity-dependent transcriptional responses. A. *RFX3* binding peaks in the promoter regions of FOSB, JUN, and JUNB in *RFX3* WT, HET, and KO day 14 neurons.

Figure 3.5 (Continued). B. Schematic of neuronal stimulation with KCl. C. Volcano plot of differentially expressed genes in 1.5 hour KCl stimulated HET v. WT or KO v. WT neurons at day 14 in culture. Genes significantly downregulated are in blue (FDR<0.05). Genes significantly upregulated are in red (FDR<0.05). Select immediate early genes are labeled. D. Fold-induction of all immediate early genes, FOSB, and JUNB at 1.5, 4, and 6 hours post KCl depolarization compared to unstimulated neurons. Data represented as mean and 95% CI, *p<0.05 t-test, n=3 independent neuron differentiations per genotype. E. Fold-induction of FOSB, FOSL2, JUNB, NR4A2 at 1.5 hours post KCl depolarization or H89 pretreatment and KCl depolarization assessed by RT-qPCR. Data represented as mean +/- SEM, n=3-4 independent wells per condition. *p<0.05, **p<0.01, ***p<0.005, ns not significant, one-way ANOVA with t-test with Bonferroni correction.

ment paradigm results in a peak in immediate-early gene (IEG) induction after 1.5 hours and late response gene (LRG) induction after 6 hours in human iPSC-derived neuronal cultures[26,228]. We used this system to ask whether *RFX3* regulates the magnitude or the timing of IEG induction. Comparison of stimulated and unstimulated WT neurons allowed us to identify 240 IEGs and 481 LRGs that were upregulated with KCl stimulation (fold change of at least 2, FDR<0.05) (A.23C, Table S3.8). Analysis of *RFX3* deficient neurons demonstrated many genes with decreased response to KCl compared to WT (3.5C, A.23D). IEGs as a group showed significantly decreased induction by 40% in *RFX3* deficient neurons 1.5 hours post depolarization (3.5D). 40/240 (17%) IEGs, including AP-1 complex members *FOSB, FOSL2, JUNB*, and *JUND*, had significantly impaired induction in both *RFX3* HET and KO neurons (3.5D, Table S3.9). The degree of impaired induction was comparable to levels observed in neurons with knockdown of *CRTC1*, a CREB transcriptional coactivator[26]. *RFX3* deficient neurons also had a significantly decreased induction of LRGs (15%) 6 hours post depolarization (A.23D-E).

We next sought to understand how *RFX3* regulates induction of activity dependent genes. *RFX3* expression itself was not altered in response to activity, and it localized to the nucleus in both unstimulated and stimulated neurons (A.23F-G), suggesting that *RFX3* has a constitutive role in regulating activity dependent transcription. Next, we considered whether *RFX3* regulation of activity dependent responses may be CREB dependent or independent. To test whether impaired IEG induction in response to KCl in *RFX3* deficient neurons depends on CREB activation by phosphory-

lation at Ser-133, we measured IEG induction with quantitative real-time PCR (qPCR) on neurons stimulated for 1.5 hours with either KCl alone or KCl + H89, a protein kinase A (PKA) inhibitor that blocks CREB Ser-133 phosphorylation[82]. Consistent with the RNA-seq data, *RFX3* deficient neurons had decreased induction of IEGs in response to KCl compared to WT. This phenotype was not present when CREB activation was blocked with H89 (3.5G). We also treated neurons with forskolin (FSK), a compound that activates CREB via activation of adenylate cyclase. *RFX3* deficient neurons showed decreased induction of IEGs in response to FSK (A.23H). These results demonstrate that the impaired IEG induction phenotype is dependent on CREB activation, confirming that *RFX3* is involved in the CREB signaling pathway.

3.2.12 *RFX3* promotes CREB binding at cis-regulatory elements of activity-dependent genes

This raises two potential mechanisms by which *RFX3* deficiency might lead to decreased IEG induction: (1) *RFX3* could be required for proper CREB activation, such that decreased *RFX3* leads to decreased CREB phosphorylation in response to stimulation, or (2) *RFX3* may act in parallel or downstream of activated CREB to modulate IEG induction. To distinguish between these two possibilities, we assessed CREB phosphorylation at Ser-133 (P-CREB) levels in response to KCl in *RFX3* deficient compared to WT neurons. By both Western blot and immunocytochemistry, P-CREB levels at 15 min and 1.5 hours post KCl depolarization were similar between WT and *RFX3* deficient neurons (A.24A-D), indicating that CREB phosphorylation was unaffected by *RFX3* deficiency.

This result suggested that despite normal activation of CREB, CREB targets cannot respond appropriately without RFX3. We reasoned that RFX3 may act through promoting CREB binding. Alternatively, RFX3 could enhance CREB function via other mechanisms, such as facilitating the recruitment of other factors (e.g., CBP). To begin to test whether CREB binding is directly altered

in *RFX3* deficiency, we profiled CREB binding with CUT&RUN-seq in *RFX3* WT, HET, and KO unstimulated neurons. We identified 3,095 CREB peaks in WT neurons. Motif enrichment analysis revealed that the CRE motif was the most enriched among these binding regions (CRE motif E-value 5.57e-200, 3.6A-B). In addition, the RFX3 motif was significantly enriched and present in 10.4% of peaks (RFX3 motif E-value 4.14e-44, 3.6B), suggesting that CREB and RFX3 might co-bind at least a subset of sites. 567 binding sites were shared between CREB and RFX3, including the promoters of AP-1 components *FOSB, JUNB,* and *JUN* (3.6C, Table S3.10). On average, RFX3 co-occupied sites had 50% increased CREB CUT&RUN signal compared to sites without RFX3 (3.6D), consistent with *RFX3* promoting CREB binding. In addition, 126/567 (22.2%) of the CREB sites co-occupied by RFX3 were very sensitive to *RFX3* dosage, wherein a CREB binding peak was not called in the *RFX3* HET and KO neurons (Table S3.11); these included the AP-1 complex components *FOSB* and *JUNB* (3.6G). Fold-change in CREB peak intensity was also weakly correlated with fold-change in RFX3 peak intensity (HET v. WT, Pearson r 0.08, p-value 0.045, n=567 sites) (3.6E). The average CREB peak intensity at co-bound sites was significantly decreased in *RFX3* HET neurons compared to the permuted distribution (n=1000 permutations) (3.6F). This suggests that RFX3 directly promotes CREB binding at co-bound sites. (Note that the average CREB peak intensity at co-bound sites was even further decreased in *RFX3* KO neurons, but it was not significantly different from the permuted distribution (n=1000 permutations), suggesting either a floor effect, or that *RFX3* KO neurons may have additional indirect perturbations that reduce CREB binding independent of RFX3 binding). Overall, these data support a mechanism whereby RFX3 and CREB are binding partners at the promoters of a subset of CREB targets. At these sites, RFX3 promotes CREB binding, such that *RFX3* deficiency leads to decreased CREB binding and impaired induction of CREB targets in response to stimulation (3.6H, 3.7).

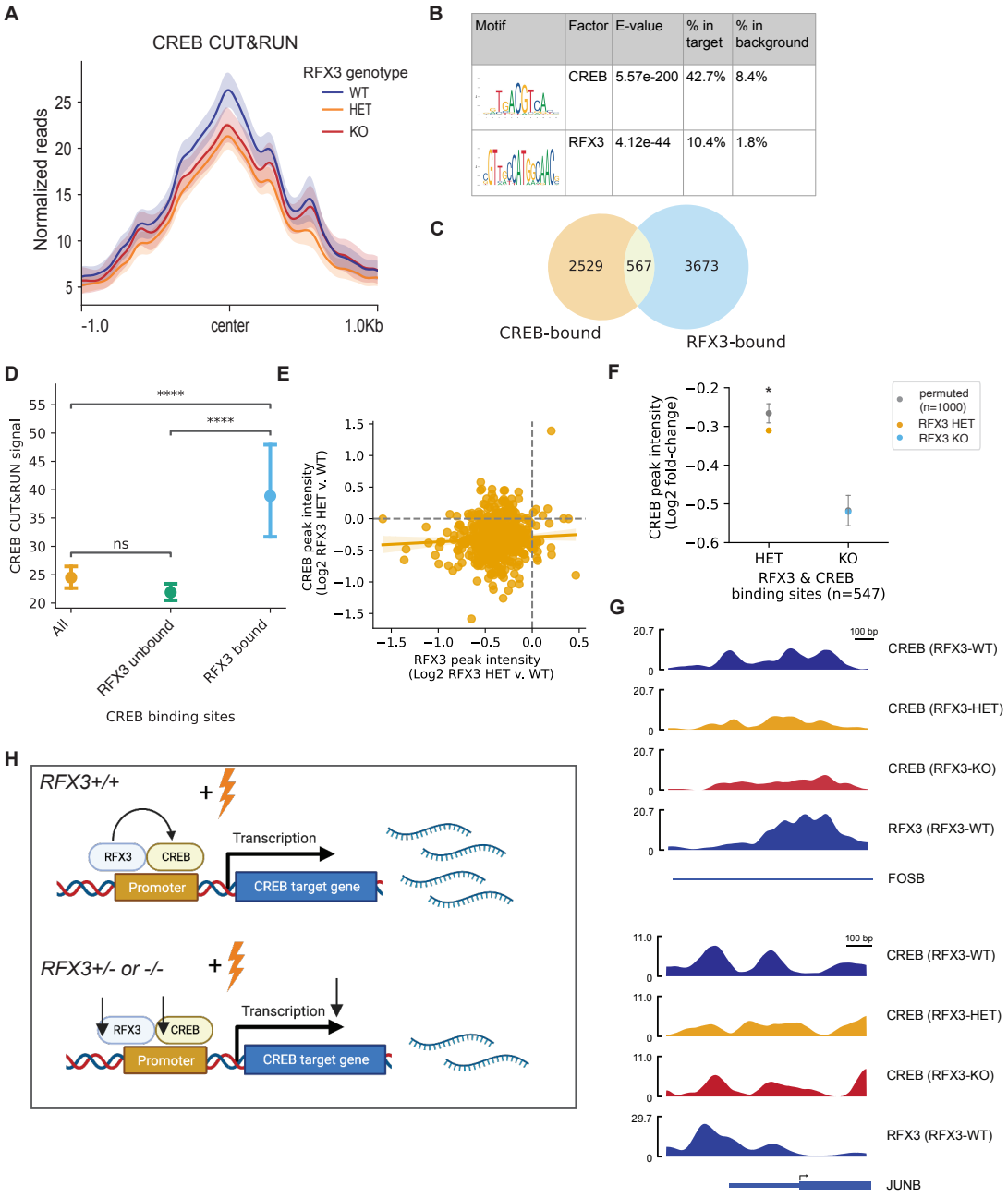


Figure 3.6: *RFX3* promotes CREB binding in unstimulated neurons. A. Aggregate plot of CREB CUT&RUN reads within CREB binding peaks called in WT neurons (n=3095 peaks). Trace shows mean +/- SEM of n=2 WT replicates, n=2 *RFX3* HET replicates, n=2 *RFX3* KO replicates.

Figure 3.6 (Continued). B. CREB and RFX3 binding motifs enriched within CREB binding regions in WT neurons. C. Overlap between CREB-bound and RFX3-bound regions in WT neurons. D. CREB CUT&RUN signal at all, RFX3 unbound, and RFX3-bound CREB binding sites in WT neurons. ****p<0.001, ns not significant, one-way ANOVA with t-test with Bonferroni correction. E. Log2 fold-change in CREB peak intensity v. log2 fold-change in RFX3 peak intensity in *RFX3* HET neurons. Linear regression model fit with 95% confidence interval (CI) shown. F. Average observed log2 fold-change in CREB peak intensity at regions co-bound by RFX3 in *RFX3* HET and KO neurons compared to permuted distribution (n permutations=1,000, 90 percentile interval). *left-tailed p <0.05, permutation test. G. CREB binding peaks in the promoter regions of FOSB, and JUNB in *RFX3* WT, HET, and KO day 14 neurons. RFX3 binding peaks in the promoter regions of FOSB, and JUNB in *RFX3* WT day 14 neurons. H. Schematic of proposed mechanism where RFX3 promotes CREB binding to modulate expression of activity dependent genes.

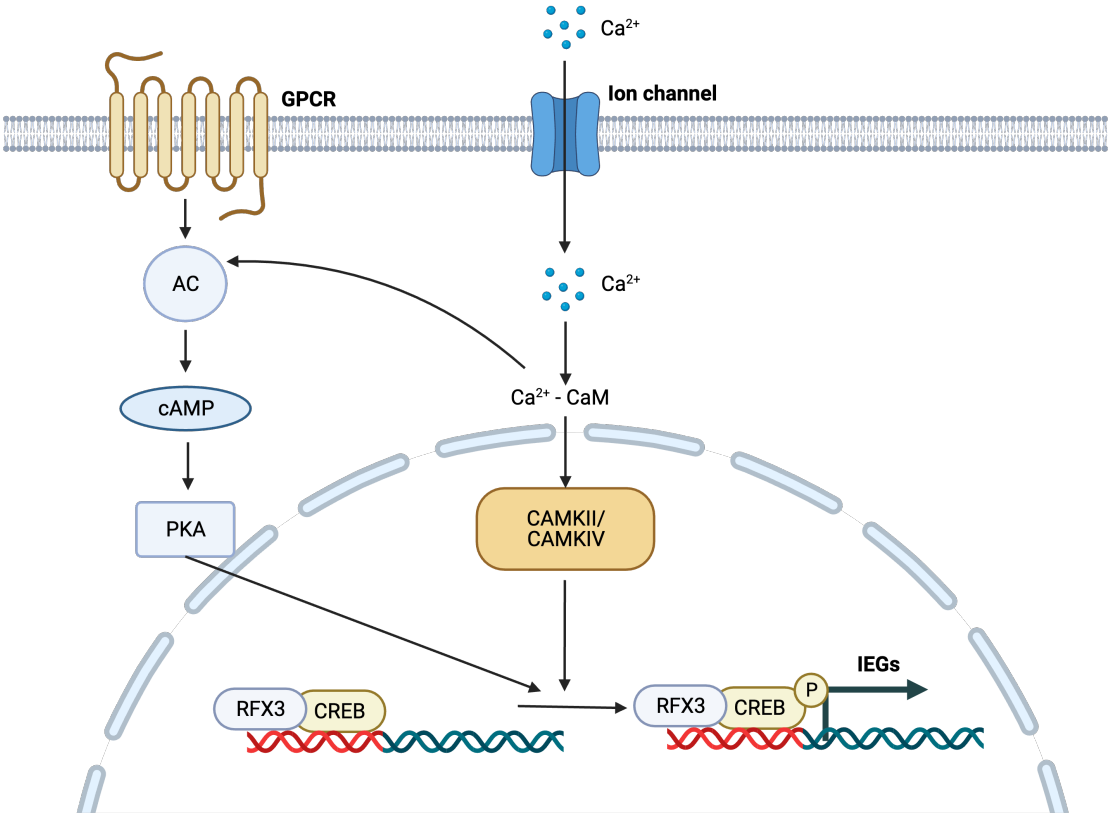


Figure 3.7: Proposed model for role of RFX3 in CREB-mediated activity-dependent signaling pathway. (Adapted from BioRender figure template "CREB Signaling Pathway"

3.3 Discussion

RFX3 is a brain-enriched transcription factor with important roles in ciliogenesis and neurodevelopment, and haploinsufficiency of *RFX3* is associated with a neurodevelopmental disorder. Here, we performed genome-wide profiling of RFX3 binding and transcriptomic alterations in *RFX3* deficiency in human neurons to provide molecular insights into transcriptional programs regulated by *RFX3* and enhance our understanding of the neurodevelopmental disorder associated with *RFX3* haploinsufficiency. Using dorsal forebrain organoids as a model of human corticogenesis, we found that biallelic loss of *RFX3* leads to altered neurogenesis associated with downregulation of ciliary genes while this phenotype is not observed in *RFX3* HET organoids. Disruption to primary cilia has been linked to impaired neurogenesis related to sonic hedgehog signaling and cell proliferation [31,74,66,240]. For example, loss of the ciliary gene *Inpp5e* (mutated in some forms of Joubert Syndrome) altered neurogenesis with overproduction of layer V neurons in mouse cortex [88], and mutations in the gene encoding AHI1, a ciliary transition zone protein, can present with polymicrogyria characterized by abnormal cortical lamination [56,75]. We show that in excitatory neurons, RFX3 binds to regulatory elements of ciliary and synaptic genes, and *RFX3* haploinsufficiency leads to decreased expression of synaptic genes, including other ASD risk genes, while ciliary gene expression is intact. This suggests that disruption of synaptic processes (rather than ciliary processes) is the primary contributor to the neurodevelopmental disorder caused by *RFX3* haploinsufficiency. *RFX3* deficiency was functionally associated with immature development of neuronal network activity, including decreased synchrony and number of spikes per network burst. Notably, decreased spontaneous inter-hemispheric synchrony was found in toddlers with ASD [55]. This phenotype is also consistent with functional phenotyping of neurons derived from individuals with idiopathic ASD and other single-gene models of ASD that found decreased spiking activity and burst frequency (*MEF2C, AFF2/FMR2, ANOS1, ASTN2, ATRX, CACNA1C, CHD8, DL-*

GAP2, KCNQ2, SCN2A, TENM), suggesting a convergence on reduced functional connectivity in ASD [53,52,6,169,186].

CREB-mediated activity dependent transcription is important for synapse maturation and dendritic arborization involved in learning [281], and dysregulation of activity dependent signaling has been previously implicated in neurodevelopmental disorders [61]. For example, variants in genes encoding several key components of activity dependent signaling lead to neurodevelopmental disorders including *L-VSCC* (Timothy syndrome), *RSK2* (Coffin-Lowry), *CBP* (Rubenstein-Taybi syndrome), *UBE3A* (Angelman), *MECP2* (Rett syndrome), and BAF complex subunits that mediate AP-1 transcriptional activity [61,221]. Studies of ASD patient iPSC-derived neurons and postmortem cerebral cortex have also identified downregulated transcriptional signatures of ASD that are enriched for activity-dependent genes and synaptic signaling, but transcriptional responses to neuronal depolarization were not tested [263,72]. In this study, using human neuronal models of *RFX3*-associated ASD, we demonstrated that *RFX3* deficiency leads to impaired induction of activity dependent genes, including *FOSB, FOSL2, JUNB, NR4A1*, and *NR4A2*. Haploinsufficiency of *NR4A2* and *FOSL2* have recently been associated with neurodevelopmental disorders [163,46]. Symptoms of neurodevelopmental disorders appear in early postnatal life, coinciding with periods of activity-dependent synapse development and refinement [281]. Thus, impairment in activity-dependent synapse maturation may be common to genetically heterogeneous neurodevelopmental disorders.

The impact of *RFX3* on activity-dependent signaling may contribute to ASD phenotypes in patients with *RFX3* deficiency, but might have mechanistic implications for idiopathic ASD as well. ASD epigenetic signatures marked by increased histone acetylation (H3K27ac) were most enriched for the RFX motif and AP-1 motif [256]. In addition, CREB and RFX family motifs were enriched in activity-inducible promoters associated with elevated ASD-heritability [26]. Here, we found that RFX3 and CREB co-occupy promoters of core activity dependent response genes, and that RFX3

promotes CREB binding at these sites, suggesting that RFX3 is a direct component of activity-dependent transcriptional complexes. RFX3 co-occupation of a subset of CREB binding sites may reflect cell-type specific modulation of activity-dependent gene programs[99,291].

This study has several limitations that indicate potential areas of further research. The human neuronal models employed only consist of excitatory neurons, so the impact of *RFX3* deficiency in inhibitory neurons remains uncharacterized. Future studies may include models of *RFX3* deficiency in inhibitory neurons, and co-culture systems of excitatory and inhibitory neurons. In addition, to refine our understanding of *RFX3* target genes in the brain, future research may leverage technologies that can map long-range physical interactions in the genome, or single-cell CUT&RUN profiling to identify cell type specific targets[184,21].

This study highlights an underappreciated role of *RFX3* in regulating synaptogenesis and synaptic plasticity. Our finding that *RFX3* deficiency is associated with impaired induction of activity-dependent genes in depolarized neurons suggests a role for dysregulated activity dependent gene expression in neurodevelopmental disorders. Understanding the role of *RFX3* in gene regulation in neurodevelopment illuminates molecular pathways that may underlie ASD and highlights modulation of functional deficits in synapses and activity-dependent responses as a potential therapeutic target in the postnatal brain.

3.4 Acknowledgements

We thank Mandovi Chatterjee, Ph.D. from the Harvard Single Cell Core for her assistance with scRNA-seq library preparation, Lee Barrett, Ph.D., Elizabeth Buttermore, Ph.D., and Nina Makhortova, Ph.D. from the Boston Children's Human Neuron Core for their assistance with high-content imaging and multi-electrode array assays, the Harvard Biopolymer Facility for bulk RNA-sequencing library preparation and sequencing, Christopher A. Walsh, M.D.-Ph.D., Michael E. Greenberg,

Ph.D., Lisa V. Goodrich, Ph.D., GiHun Choi, Ph.D, and members of the Yu and Lee labs for helpful scientific discussions. Illustrations were created with BioRender.com.

This work was supported by the National Institute of Health grant R01MH113761 and the G. Harold & Leila Y. Mathers Foundation. J.L is supported by award Number T32GM007753 and T32GM144273 from the National Institute of General Medical Sciences, and the Ruth L. Kirschstein NRSA F30 Fellowship (1F30MH128995). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health.

3.5 Author Contributions

Conceptualization, T.W.Y, E.A.L, and J.L; Methodology, T.W.Y, E.A.L, J.L, D.C; Investigation, J.L, D.C, K.P, H.W, B.Z, M.S, T.N, H.H; Formal Analysis, J.L, B.Z, A.C; Validation, K.P, H.W; Supervision, T.W.Y, E.A.L; Writing – Original Draft, J.L; Writing – Review & Editing, T.W.Y, E.A.L, J.L., D.C

3.6 Methods

3.6.1 Maintenance and culture of iPSCs

Parental PGP1-SV1 iPSC clones and *RFX3* CRISPR-Cas9 edited clones were obtained from Synthego. All iPSC lines were mycoplasma negative, karyotypically normal, and expressed pluripotency markers OCT4, SOX2, NANOG, SSEA4, TRA-1-60. iPSCs were maintained in 6-well or 10 cm plates (Corning) coated with LDEV-free hESC-qualified Matrigel (Corning Catalog #354277) in feeder-free conditions with complete mTeSR-plus medium (STEMCELL Technologies) in a humidified incubator (5% CO_2, 37°C). iPSCs were fed fresh media daily and passaged every 3-4 days.

3.6.2 Differentiation of iPSCs to neurons

Human iPSC-derived neurons (iNs) were generated from iPSCs transduced with the lentiviruses Tet-O-Ngn2-Puro and FUW-M2rtTA based on a previously published protocol[299]. On day -1 of differentiation, iPSC colonies were dissociated into single-cells using Accutase (STEMCELL Technologies) and 4-8 million cells were plated on Matrigel coated 10 cm dishes with complete mTeSR plus medium supplemented with Y-27632 (10 uM). On day 0, cells were fed N2 media (DMEM/F-12 media, 1X N2, 1X Nonessential Amino Acids) supplemented with doxycycline (2 μg/mL), BDNF (10 ng/mL), NT3 (10 ng/mL), and laminin (0.2 μg/mL). On day 1, media was replaced with N2 media with puromycin (1 μg/mL) in addition to the supplements listed above. On day 2, cells were fed B27 media (Neurobasal media, 1X B27, 1X Glutamax) supplemented with puromycin (1 μg/mL), doxycycline (2 μg/mL), Ara-C (2 uM), BDNF (10 ng/mL), NT3 (10 ng/mL), and laminin (0.2 μg/mL). On day 3, cells were dissociated into single cells using Accutase and replated onto polyethylenimine (PEI)/laminin (0.07% PEI and laminin diluted 1:200 in sterile water) coated 96-well plates (10,000 cells/well) with B27 media supplemented with Y-27632 (10 μM), doxycycline (2 μg/mL), Ara-C (2 μM), BDNF (10 ng/mL), NT3 (10 ng/mL), and laminin (0.2 μg/mL). On days 5, 7, 10, and 14, cells were fed with a half-media change of Conditioned Sudhof Neuronal Growth Medium (1:1 ratio of Astrocyte Conditioned Media and Neurobasal Media, 1X B27, Glutamax, $NaHCO_3$, and transferrin) supplemented with BDNF (10 ng/mL), NT3 (10 ng/mL), and laminin (0.2 μg/mL). Cells were fed with a half-media change of Conditioned Sudhof Neuronal Growth Medium as described above once a week. Cells were maintained in a humidified incubator (5% CO_2, 37°C).

3.6.3 Dorsal forebrain organoid culture

Human dorsal forebrain organoids were generated from iPSCs using the STEMdiff Dorsal Forebrain Organoid Differentiation Kit (STEMCELL Technologies Catalog # 08620) according to the manufacturer instructions. On day 0 of organoid formation, iPSCs from each line were dissociated into single cells using Gentle Cell Dissociation Reagent (STEMCELL Technologies Catalog # 100-0485) and 3 x 10^6 cells were seeded into one well of an AggreWell 800 plate (Catalog # 34811) with Seeding Medium. On day 1-5, uniform embryoid bodies were visible and were fed a half-media change of Forebrain Organoid Formation Medium. On day 6, organoids were transferred to 6-well Ultra-Low Adherent plates with Forebrain Organoid Expansion Medium (15 organoids per well) and evenly dispersed to ensure no contact between individual organoids and incubated on a stationary level surface. On day 8-25, full media changes with Forebrain Organoid Expansion Medium were performed every 2 days. On day 25, full media changes with Forebrain Organoid Differentiation Medium were performed every 2 days. From day 43 onwards, full media changes with Forebrain Organoid Maintenance Medium were performed every 3-4 days. Organoids were maintained in a humidified incubator (5% CO_2, 37°C).

3.6.4 Bulk RNA-sequencing

Total RNA was harvested from samples using the PureLink RNA Mini kit (Life Technologies Catalog # 12183018A). Libraries were prepared with the KAPA mRNA prep. Sequencing was performed on Illumina NovaSeq, 2x150bp configuration at approximately 100X depth per sample. Reads were aligned to hg38 using STAR v2.7.9a[58] followed by processing with featureCounts (-p -s 2) to obtain a gene-level counts matrix[156]. Differential gene expression analysis was performed using DESeq2[162] with WT set as the reference condition and results coefficient set as 'condition_HET_vs_WT' or 'condition_KO_vs_WT', and lfcShrink(type="apeglm"). Differentially

expressed genes with FDR<0.05 and |log2 fold-change| >0.25 were considered significant and used in downstream analysis.

3.6.5 Isolation of single cells and scRNA-seq library preparation from dorsal forebrain organoids

Individual organoids were dissociated into single cells using 500 uL dissociation solution per organoid (30 units/mL papain, 125 units/mL DNaseI diluted in Hank's Balanced Salt Solution (HBSS)) for 30 minutes on a shaker in an incubator at 37C. Organoids were triturated with a P1000 pipette 10-15 times and then returned to the incubator for 15 minutes. Organoids were triturated with a P1000 pipette 10-15 times again to obtain single cells and added to a 15 ml conical tube containing 1 ml of ovomucoid protease inhibitor (10 mg/mL in HBSS). Cells were centrifuged 300g x 5 minutes at room temperature and resuspended in 1 mL HBSS. Cells were counted and assessed for morphology, doublets, and debris. Cell fixation was performed using the Parse Cell Fixation Kit v1 according to manufacturer instructions. Barcoding and library preparation were performed using the Parse Single Cell Whole Transcriptome Kit v1 or Parse Evercode Whole Transcriptome Mega Kit v1 according to manufacturer instructions at the Harvard Single Cell Core. Libraries were paired-end sequenced using Illumina NovaSeq 6000 S2 with target 40,000 reads per cell.

3.6.6 scRNA-seq analysis

scRNA-seq data for each individual organoid was preprocessed and aligned to the human reference genome Homo_sapiens.GRCh38.93 using the Parse Biosciences computational pipeline split-pipe –mode all. The resulting count matrices were used to create Seurat objects per organoid using Seurat v4[253]. Further quality control removed cells with less than 200 genes per cell, greater than 10000 genes per cell, and greater than 15% mitochondrial content.

SCTransform was used to perform normalization of expression values per sample. After normalization, samples were integrated to align common clusters across individual datasets using Seurat's integration method[253]. Dimensionality reduction was performed using RunPCA and RunUMAP. A shared nearest neighbor graph was constructed using FindNeighbors with dims 1:30, k.param 100, and cosine metric. Clusters were then identified using the Leiden algorithm at a resolution of 1.5.

Marker genes for each cluster were obtained using FindAllMarkers with a minimum log2 fold-change of 0.25. Cell type identities were assigned to clusters based on reference-based annotations and validation with cell type markers from literature. SingleR was used for reference-based annotation using reference datasets from the human developing cortex and dorsal forebrain organoids[275,208,274,218].

To identify changes in cell type proportions in *RFX3* mutant organoids compared to controls, we used a Bayesian statistics approach to compositional data analysis previously described in the context of microbial ecology[86]. Compositional data has a negative covariance structure that is accounted for by the multinomial and Dirichlet probability distributions. We implemented Dirichlet multinomial modeling (DMM) using Hamiltonian Monte Carlo (HMC) using the R packages rstan and bayestestR[166]. The input to the model was a matrix of counts where the rows correspond to replicates (individual samples) and the columns correspond to cell types. We then modeled the replicates using the multinomial distribution for the probability of each cell type per organoid. The Dirichlet distribution was used to model the multinomial parameters. The prior on the Dirichlet parameters was another Dirichlet distribution with a fixed parameter alpha 10-7, which gives uniform cell type proportions in expectation. The resulting posterior distributions were Dirichlet distributions of the cell type proportions in mutant and control. We then subtracted the posterior probability distribution (95% credible interval) of control from mutant to see whether there are significant differences in relative cell type composition. A cell type shift in log2 abundance ratio (*RFX3* mutant/control) was considered significant if the 95% credible did not include 0.

3.6.7 Differential gene expression and Gene Ontology analysis

Differential gene expression analysis was performed on each cell type using edgeR[220]. The integrated Seurat object was subset to obtain Seurat objects for each cell type. The read counts were modeled and normalized using a negative binomial distribution with the trimmed mean of M-values (TMM) normalization method. To ensure robust signals, only genes that were expressed in at least 5% of one cell type were included in the analysis. The design matrix formula was genotype + cellular detection rate. Differentially expressed genes between *RFX3* HET or KO and control were identified using the likelihood ratio test (glmLRT). Genes with FDR < 0.05 and a log2 fold-change greater than 0.25 were used for downstream analysis.

Gene Ontology (GO) enrichment analysis was performed on the significantly downregulated differentially expressed genes in each cell type using clusterProfiler[294]. GO biological processes with FDR < 0.05 were considered significant.

3.6.8 CUT&RUN-sequencing

Neurons were dissociated into single cells with DNase/papain dissociation solution prepared by combining 20 units/mL papain (Worthington) and 1 unit/mL DNase (Worthington) in DMEM/F12 (STEMCELL Technologies) and incubated at 37C for 15 minutes. 500,000 single cells per sample were used as input for the CUTANA ChIC/CUT&RUN kit (EpiCypher) according to the manufacturer instructions. Cells were washed twice with CUT&RUN wash buffer and then bound to concanavalin A-coated (conA) beads for 10 minutes at room temperature. ConA bead-bound cells were resuspended in CUT&RUN antibody buffer with the following antibodies: anti-RFX3 (Sigma-Aldrich Cat# HPA035689 diluted 1:100 and Abcam Cat# ab168475 diluted 1:100), anti-H3K4me3 (EpiCypher Cat# 13-0041 diluted 1:100), anti-IgG (EpiCypher Cat# 13-0042 diluted 1:100), anti-CREB (Sigma-Aldrich Cat# 06-863 diluted 1:100) and incubated overnight at 4C on

a shaking nutator with capped ends elevated. After incubation, the conA bead-bound cells were washed twice with Cell Permeabilization Buffer. To each sample, 2.5 uL of pAG-MNase (20X stock) was added followed by incubation for 15 minutes at room temperature. After incubation, the conA bead-bound cells were washed twice with Cell Permeabilization Buffer. On ice, 1 uL of 100 mM Calcium Chloride was added to each sample with gentle vortexing and incubated on a shaking nutator for 2 hours at 4C. After incubation, 33 uL of Stop Buffer was added to each sample with gentle vortexing. 1 uL (0.5 ng) of E. coli Spike-in DNA was added to each sample and incubated for 10 minutes at 37C in a thermocycler. Samples were placed on a magnetic stand and the supernatant containing CUT&RUN enriched DNA was transferred to clean tubes. DNA was purified from the supernatants using the included DNA Cleanup Columns according to the manufacturer instructions. Purified CUT&RUN-enriched DNA was quantified using Qubit dsDNA HS Assay kit (Thermo Fisher Scientific Cat# Q32851) according to the manufacturer instructions. CUT&RUN-sequencing libraries were prepared using KAPA HyperPrep Kit (Kapa Biosystems Cat# KK8502) according to the manufacturer instructions with the following modifications. 25 uL of purified CUT&RUN-enriched DNA was used as input per sample for End Repair and A-Tailing. To enrich sub-fragments during End Repair and A-Tailing, the thermocycling parameters were 20C for 30 minutes followed by 50C for 1 hour. 15 uM adapter stock was used for the Adapter Ligation reaction. Post-ligation cleanup was performed three times with 1X, 1.2X, and 1.2X AMPure beads (Beckman Coulter Cat# A63881). 16 PCR cycles were performed for library amplification. Post-amplification cleanup was performed three times with 1.1X AMPure beads. Purified amplified libraries were quantified by qPCR and library quality and size was assessed using Agilent TapeStation. Libraries were paired-end sequenced with Illumina NextSeq 2x150 bp reads with target 10 million reads per sample. Adapters were trimmed using cutadapt -a AGATCGGAAGAGCACACGTCTGAACTCCAGTCA -A AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT[172]. Alignment to hg38 was performed using bowtie2 –local –very-sensitive-local –no-

mixed –no-discordant –phred33 –dovetail -I 10 -X 700 [142], followed by sorting and indexing with samtools (Danecek et al., 2021). Peak calling was performed using macs2 callpeak for each sample with -c set as the Rabbit IgG negative control and -f BAMPE -p 0.05 [300]. Reproducible peak sets among >3 biological replicates were called using ChIP-R with parameter -m 3 [196]. Reproducible peak sets among 2 biological replicates were called using IDR [152]. Peaks were annotated to nearest genes and genomic regions using ChIPseeker [295]. Gene Ontology analysis was performed using clusterProfiler as described above [294]. For categorization of targets as ciliary genes, we used the lists of ciliary genes from CiliaCarta [270]. For categorization of targets as synaptic genes, we used the list of synaptic genes from SynGO [131]. Average peak profiles were generated from RPGC normalized bigwig files using deeptools computeMatrix and plotProfile [216]. Motif enrichment analysis was performed using the MEME suite [16]. Differential motif enrichment analysis was performed using Homer getDifferentialPeaksReplicates.pl [89]. For evolutionary conservation analysis, human RFX3 binding sites were converted into homologous sequences in other species using UCSC LiftOver [122] followed by RFX3 binding motif scanning using Homer findMotifsGenome.pl [89].

3.6.9 Reverse transcription quantitative PCR (RT-qPCR)

RNA was reverse transcribed into cDNA using Cells-to-CT RT buffer and RT enzyme (Thermo Fisher Scientific Cat# A35378) consisting of 22.5 ul RNA and 27.5 ul of Cells-to-CT RT Master Mix. The reverse transcription reactions were incubated at 37C for 1 hour, and 95C for 5 min. qPCR was performed using TaqMan Fast Advanced Master Mix (Thermo Fisher Scientific Cat# A35378) and Gene Expression Assays (Thermo Fisher Scientific Cat# 4331182) with 2 ul input cDNA. The thermocycling parameters were 50C for 2 min, 95C for 2 min, 40 cycles of 95C for 1 second and 60C for 20 seconds. TaqMan Gene Expression Assays included RFX3 Hs01060440_m1, RFX3 Hs01060430_m1, FOSB Hs00171851_M1, FOSL2 Hs01050117_m1, JUNB Hs00357891_s1, NR4A2 Hs00428691_m1, Human GAPDH Endogenous Control (VIC/MGB probe, primer lim-

ited, Thermo Fisher Scientific Cat# 4326317E).

3.6.10 Western blotting

Cell lysates were collected using RIPA buffer (Boston BioProducts Cat# BP-115D-250) supplemented with Roche cOmplete Protease Inhibitor (Millipore Sigma Cat# 5892791001) and PhosSTOP (Millipore Sigma Cat# 4906845001). Lysates were homogenized with a hand-held spin homogenizer for 10 seconds, then frozen at -80C for 20 minutes. Lysates were then thawed on ice, centrifuged at 20,000 g for 20 minutes at 4C, and supernatant was transferred to a clean tube on ice. Lysates were incubated with 4x Laemmli buffer (BioRad), placed in a heat block at 95C for 5 minutes, then loaded onto 4-15% precast gradient protein gels (BioRad) and separated by electrophoresis (50V for 5 min, 150V for 45 min). Protein samples were then transferred to PVDF membranes and incubated at 4C overnight with the following primary antibodies: anti-RFX3 (Sigma-Aldrich Cat# HPA035689 diluted 1:100), anti-PCREB (Invitrogen Cat# MA1-114 diluted 1:250), anti-CREB (Sigma-Aldrich Cat# 06-863 diluted 1:500 and Cell Signaling Technology Cat# 4820S diluted 1:500), anti-ARL13B (Proteintech, Cat# 17711-1-AP diluted 1:500), anti-GAPDH (Abcam Cat# ab9485 diluted 1:1,000), anti-alpha Tubulin (Sigma-Aldrich Cat# ab1543 diluted 1:10,000). Membranes were incubated with secondary antibodies diluted 1:5,000 for the target and 1:10,000 for the loading control (GAPDH or alpha Tubulin) and visualized with the Li-Cor Odyssey system and quantified with EmpiriaStudio software (Li-Cor).

3.6.11 Neuronal depolarization

12 hours prior to depolarization, neuronal cultures were silenced with a full medium replacement consisting of medium supplemented with 1 uM TTX (Abcam Cat# Ab120055) and 100 uM D-APV (Tocris Cat# 0106). Neurons were then left in the silenced condition (unstimulated) or de-

polarized with KCl depolarization buffer (170 mM KCl, 2 mM $CaCl_2$, 1 mM $MgCl_2$, 10 mM HEPES, solution pH 7.4) to a final KCl concentration of 55 mM. For pretreatment with H89 dihydrochloride (Tocris Cat# 2910), H89 was added to a final concentration of 10 uM for 1 hour prior to KCl depolarization. For cultures treated with Forskolin (FSK), FSK was added to a final concentration of 10 uM.

3.6.12 Multielectrode array

48-well CytoView MEA plates (Axion BioSystems) were coated with 0.07% PEI coating buffer at 37C overnight, washed three times with sterile water, air dried, and then coated with Laminin (diluted 1:200 in sterile water) at 37C overnight. On day 3 of neuron differentiation, wells were plated with 100,000 neurons and 20,000 human iPSC-derived astrocytes (Ncardia Ncyte Astrocytes) per well. Spontaneous neural activity and viability was recorded using the Axion Maestro Pro MEA (Axion BioSystems) with default burst detection settings twice a week for 10 minutes per recording for the following 8 weeks. Half-media changes were performed once a week on a non-recording day with unconditioned neuronal growth medium (consisting of Neurobasal media, B27, $NaHCO_3$, Glutamax, Transferrin). Neural activity recordings were processed and analyzed using Axion Axis Navigator and AxIS Metric Plotting Tool softwares with default settings (Axion BioSystems). Data from wells with >0 active electrodes per well were included and were normalized to the number of covered electrodes per well. The number of biological replicates were n=32 wells per genotype from two independent MEA plates.

3.6.13 Immunostaining

Cells were fixed with 4% paraformaldehyde for 20 minutes at room temperature and stored at 4C until used for immunostaining. Cells to be stained for PSD95 were fixed with ice cold methanol

at -20C for 20 minutes. Cells were incubated with CellO-IF (Cellorama Tech) for 10 minutes at 37C, then incubated with primary antibodies diluted in CellO-IF for 1 hour at 37C. Following incubation, cells were washed with PBS two times, then incubated with secondary antibodies diluted in CellO-IF for 1 hour at 37C in the dark. Cells were incubated with Hoechst diluted 1:1,000 in PBS for 10 minutes in the dark. After two PBS washes, cells were mounted with glycerol solution (1:1 glycerol and PBS). Primary antibodies included anti-RFX3 (Sigma-Aldrich Cat# HPA035689, diluted 1:100), anti-phospho-CREB (Invitrogen Cat# MA1-114, diluted 1:250), anti-CREB (Sigma-Aldrich Cat# 06-863, diluted 1:500), anti-PSD95 (Neuromab Cat# 75-028, diluted 1:500), anti-SYN1 (Sigma-Aldrich Cat# ab1543, diluted 1:1,000), anti-beta III Tubulin (GeneTex Cat# GTX85469, diluted 1:1,000), anti-ARL13B (Proteintech Cat# 17711-1-AP, diluted 1:800), anti-MAP2 (Abcam Cat# ab5392, diluted 1:2,000). Alexa fluor-conjugated secondary antibodies were used accordingly (diluted 1:500).

3.6.14 High-content imaging and analyses

96-well plates (Greiner Bio-One Cat# 655090) were imaged with ImageXpress MicroXLS Widefield High-Content Molecular Device microscope (Molecular Device, LLC, MetaXpress v6.6.2.46) under 20X or 40X magnification with 9 evenly spaced fields per well. For Z-stack, five steps and 3 um thickness/step were used and stacked images were combined with maximum resolution projection. Molecular Devices MetaXpress software was used to design algorithms to systematically identify and quantify selected features from each image. ImageJ software was used to change the LUT for representative images.

3.6.15 Statistical analyses

All statistical tests were performed in R (v4.0.1) or Python (v3.8) with multiple hypothesis correction using the Benjamini-Hochberg procedure (FDR<0.05) or Bonferroni correction as specified in figure legends. Statistical tests are noted in figure legends, and include ANOVA, Fisher's exact test, Student's t-test, permutation test, and Mann-Whitney U-test.

3.6.16 Data

Human iPSC-derived NGN2 neuron ATAC-seq data were obtained from GEO: GSE196854 from Sanchez-Priego et al., 2022.[228] H3K27ac CUT&RUN data were obtained from GEO: GSE196207 from Sanchez-Priego et al., 2022.[228]

RNA-seq, scRNA-seq, and CUT&RUN data will be deposited at GEO and publicly available on the date of publication. This paper does not report original code. Additional information to reanalyze the data in this paper is available from the corresponding author upon reasonable request.

3.7 Supplemental Data

Supplementary figures are included in the Thesis Appendix. Tables are included in the supplementary data of the manuscript (Lai et al., 2023) upon which this chapter is based are available from the author upon request.

"We are witnessing a new era of medicine where genetic information becomes the foundation for targeted and effective treatments."

Jennifer Doudna

4

Harnessing mRNA processing inefficiencies for single gene-based therapies

4.1 Introduction

There are over 7,000 rare diseases caused by genetic variants in single genes; in aggregate these conditions are estimated to affect 1 in 10 people in the United States[79]. Often, targeting the causal

genes (or even mutations) can be effective treatments for the genetic disorder at the source. While there are several different approaches to gene therapy, two major therapeutic approaches that we will discuss in this chapter include (1) gene editing to correct the mutation, or (2) RNA-targeting approaches to therapeutically modulate the expression of the disease gene: knocking down gene expression in the case of diseases associated with gain-of-function genetic variants, or restoring gene expression in the case of diseases associated with loss-of-function genetic variants.

4.1.1 CRISPR-Cas9

CRISPR-Cas9 based therapies utilize gene-editing technology to permanently alter DNA to correct deleterious mutations, disrupt genes (for instance to reduce expression of gain-of-function alleles), or insert new sequences in the genome[67]. Several advantages of CRISPR based therapies include their permanent nature, such that a one-time administration of the therapy has the potential to cure the disease, the precision of gene editing based on complementary sgRNA, and their relatively easy ability to be developed for many targets (4.1). A major challenge of the current generation of CRISPR therapeutics is the need for delivery vectors that can target the cell type of interest and carry genetic cargo. For example, one of the most common delivery vectors is adeno-associated viruses (AAVs), which have a 4.7 kb packaging limit[286]. This puts constraints on the size of donor templates for precise knock-in by homology directed repair. Viral vectors also pose immunogenicity risks due to the viral capsid or genome that may activate host innate immunity[241]. Furthermore, while potential off-target effects can be minimized through computational predictions, any unintended off-target editing is permanent.

Table 4.1: Pros and Cons of Different Gene-based Therapeutic Modalities

Consideration	CRISPR-Cas9	Antisense oligonucleotides (ASOs)
Treatment administration	Permanent one-time administration. Effect cannot be titrated.	Regular dosing (~every 4 months), effect can be titrated.
Specificity	Precise targeting with complementary sgRNA sequence.	Precise targeting with complementary ASO sequence.
Programmability	Can be developed for many gene targets by altering the sgRNA sequence.	Can be developed for many gene targets by altering the ASO sequence.
Delivery	Requires a delivery vector (i.e., AAVs, nanoparticles).	Freely taken up by cells via endocytic mechanisms.
Off-target effects	Can computationally predict off-target effects. Potential off-target effects are permanent.	Can computationally predict off-target effects. Potential off-target effects are reversible.
Target	For homology directed repair or precise gene knock-in, donor template length may be limited by packaging capacity of the delivery vehicle.	No limits on target transcript length.
Cell types	Depends on the delivery vehicle.	Evidence for broad uptake by many cell types.

4.1.2 Antisense oligonucleotides

Antisense oligonucleotides (ASOs) are short (typically less than 25 nucleotides) single-stranded oligonucleotides that bind complementary sequences on RNA (typically pre-mRNA or mRNA) to modulate gene expression[219]. Similar to CRISPR-Cas9 approaches, ASOs can be designed to bind specific genomic targets and can be easily customized to different targets (4.1). Moreover, ASOs do not require delivery vectors, since they are freely taken up by cells through endocytic mechanisms referred to as gymnosis[260] with broad biodistribution. A recent study on single-cell ASO activity throughout the brain demonstrated that ASOs have broad activity across individual cells[188]. This is in contrast with viral-vector based gene therapies which may have variable transduction efficiency, although work to engineer AAV capsid variants with improved and specific transduction is on-going[76]. Finally, while ASOs require regular dosing every few months (which may be considered an inconvenience), this also does allow for drug dose titration during a patient's clinical course to maximize the benefits and minimize any side effects.

4.1.3 Types of ASOs

ASOs can modulate gene expression by multiple distinct mechanisms including 1) splicing modulation and 2) RNase H-activation. Splice-switching ASOs work by steric hindrance of splicing cis-regulatory elements (i.e., exon or intron splicing enhancers and silencers) and can be used to rescue the impact of pathogenic splice-altering genetic variants. Knockdown ASOs lead to degradation of their complementary targets by recruiting RNase H to a central gap region (hence referred to as "gapmer" ASOs) and inducing cleavage of the target RNA[177]. A specialized application of splice-switching ASOs that deserves special attention is to use them to boost expression of some genes by targeting naturally occurring mRNA transcripts that would otherwise undergo nonsense-mediated decay (NMD). For example, the splicing of some genes can result in a subset of transcripts that exhibit intron retention, skipping of a critical exon ("vulnerable exon"), or inclusion of a poison exon; these in turn can lead to NMD (4.1). Splice-switching ASOs can alter patterns of splicing to disfavor formation of nonproductive isoforms and promote productive isoforms, thus increasing gene expression. These non-productive (NMD-destined) transcripts are a reservoir that can be harnessed to increase gene expression. This is especially useful as a therapeutic strategy for diseases caused by haploinsufficiency of a single gene[158]. Finally, ASOs can also increase gene expression by targeting untranslated regions (UTRs) of transcripts to modulate the use of specific translation initiation sites or binding of microRNAs[231,2,155,251].

4.1.4 ASOs for neurogenetic disease

ASOs are a useful therapeutic modality for neurogenetic diseases in particular, given their broad biodistribution in the brain and spinal cord following intrathecal injection that circumvents the blood-brain barrier[111,188]. Over 15 ASOs have entered clinical development for rare neurogenetic diseases, with nusinersen for spinal muscular atrophy (SMA) receiving FDA approval in

2016[103,69,180], tofersen for amyotrophic lateral sclerosis (ALS) receiving FDA approval in 2023, and others in ongoing clinical trials for Dravet Syndrome and Angelman Syndrome[177]. Individualized/small-N ASO trials led by investigators are also underway for Batten disease, Ataxia Telangiectasia, and other ultra rare, serious neurogenetic conditions[127,128,133,266]. Recent work establishing a framework for identifying variants that are amenable to ASO intervention estimates about 15% of individuals with a potentially ASO amenable variant in the recessive neurogenetic disorder Ataxia Telangiectasia[128].

4.1.5 ASOs for haploinsufficiency disorders

In contrast to mutation-specific ASOs, ASOs for haploinsufficiency disorders can be mutation-agnostic, offering the potential to treat patients with haploinsufficiency of a single gene regardless of their genetic variants. These ASOs can target mRNA processing inefficiencies in the wildtype allele of the disease-causing gene to increase its expression. This may restore overall expression of the gene to normal levels to modify the disease course or even reverse phenotypes[81].

There are over 100 known genes with *de novo* heterozygous loss-of-function variants that increase the risk for Autism spectrum disorders (ASD)[232]. Monogenic forms of ASD that are caused by genetic haploinsufficiency, such as *RFX3* haploinsufficiency, may be candidates for mutation-agnostic, gene-specific ASOs that boost expression of the wildtype allele. Here, we identified mRNA processing inefficiencies in ASD risk genes in human iPSC-derived neurons to define opportunities for ASO-based restoration. Using *RFX3* as an example case, we designed and tested splice-modulating ASOs in human neuronal models of *RFX3* haploinsufficiency and identified candidates that boosted *RFX3* expression. These ASOs serve as tool compounds to study the function of *RFX3* in human neurons, define critical windows for restoration of *RFX3*, and potentially alleviate symptoms associated with *RFX3* haploinsufficiency.

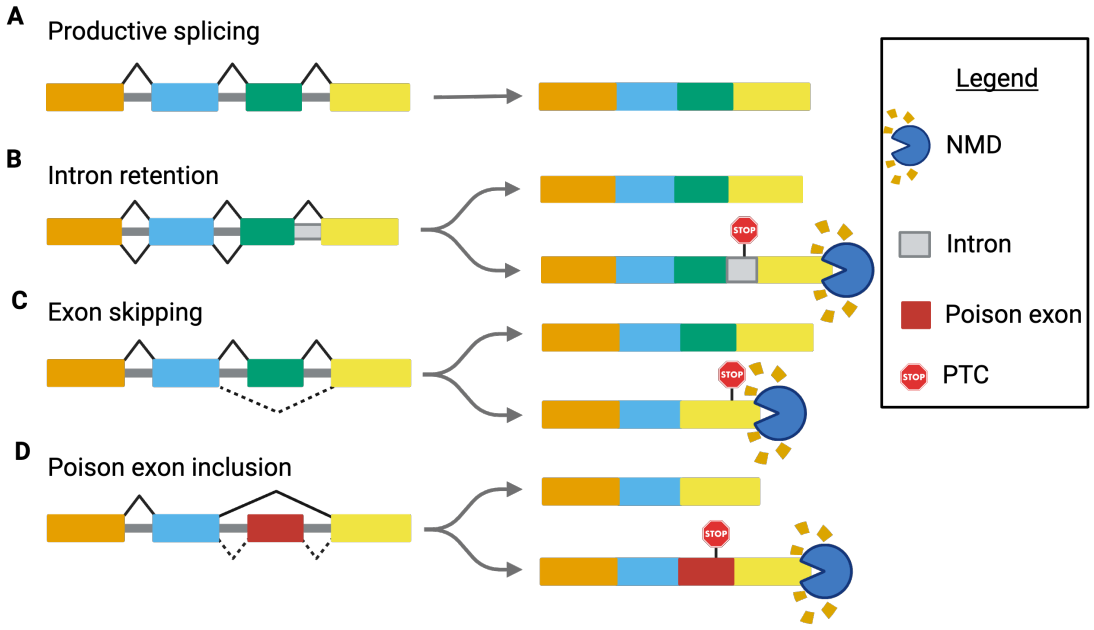


Figure 4.1: mRNA processing inefficiencies subject to nonsense-mediated decay (NMD). A. Schematic of productive splicing. B. Schematic of intron retention. Premature termination codon (PTC) within the retained intron triggers nonsense-mediated decay (NMD). C. Schematic of exon skipping. Skipping of a critical "vulnerable" exon triggers NMD. D. Schematic of poison exon inclusion. Poison exons contain a PTC that triggers NMD.

4.2 Results

4.2.1 Identification of intron retention in ASD risk genes

Individuals with heterozygous loss-of-function variants in monogenic forms of ASD have a single wild-type allele that can be leveraged to increase levels of productive mRNA and full-length protein. The expression of these genes can be increased by ASOs targeting naturally occurring mRNA processing inefficiencies that are subject to NMD (4.1).

To identify ASD risk genes with NMD-sensitive transcripts in human neurons, we differentiated control human iPSCs into NGN2 neurons[299], inhibited NMD using cycloheximide (CHX), and performed deep RNA-sequencing of neurons with and without NMD inhibition for alter-

native splicing analysis. Intron retention may introduce premature termination codons and is one of the major triggers of NMD [14,183] (4.1). To identify intron retention events in ASD risk genes, we estimated intron retention genome-wide in human iPSC-neurons using the bioinformatic tool IRFinder (IRratio) [161] and an additional method to calculate percent intron retention (PIR) [29]. IRratio and PIR levels were highly correlated (4.2) (Pearson r=0.97, p-value <0.05). We then compared the intron retention levels in neurons with and without NMD blockade and only considered those with increased IRratio and PIR in NMD inhibited neurons. As an estimate of gene expression that may be recovered by targeting intron retention, we stratified ASD risk genes by their intron retention levels: high (>=0.30), medium (0.30 > and >=0.10), and low (< 0.10). This resulted in two genes (GIGYF1, SHANK2) with high intron retention that are potentially strong candidates for a therapeutic strategy involving intron retention (tab:asdIR). There were nine genes (*ASH1L, CACNA1E, DSCAM, GIGYF1, KDM6B, MKX, NR3C2, NSD1, RORB*) with medium intron retention that may be potential candidates, and 85 genes with low intron retention such that this approach is unlikely to sufficiently increase gene expression levels. These findings demonstrate that ASOs targeting intron retention may be a potential therapeutic strategy for 10% (10/102) of highly penetrant ASD risk genes.

4.2.2 Identification of RNA processing inefficiencies in *RFX3* that can be targeted by ASOs

We next asked whether *RFX3* expression could be increased by ASOs that target naturally occurring mRNA processing inefficiencies in *RFX3*. Since there were no candidate retained introns in *RFX3* identified, we next examined the presence of poison and skipped ("vulnerable") exons using a custom pipeline in our lab (Yu Lab, unpublished) that quantifies junctional reads at exon/intron boundaries. Using this approach, we identified an exon in *RFX3* that is subject to frequent exon skipping (24%) at baseline, leading to a transcript that is predicted to lead to NMD (NM001282116.2

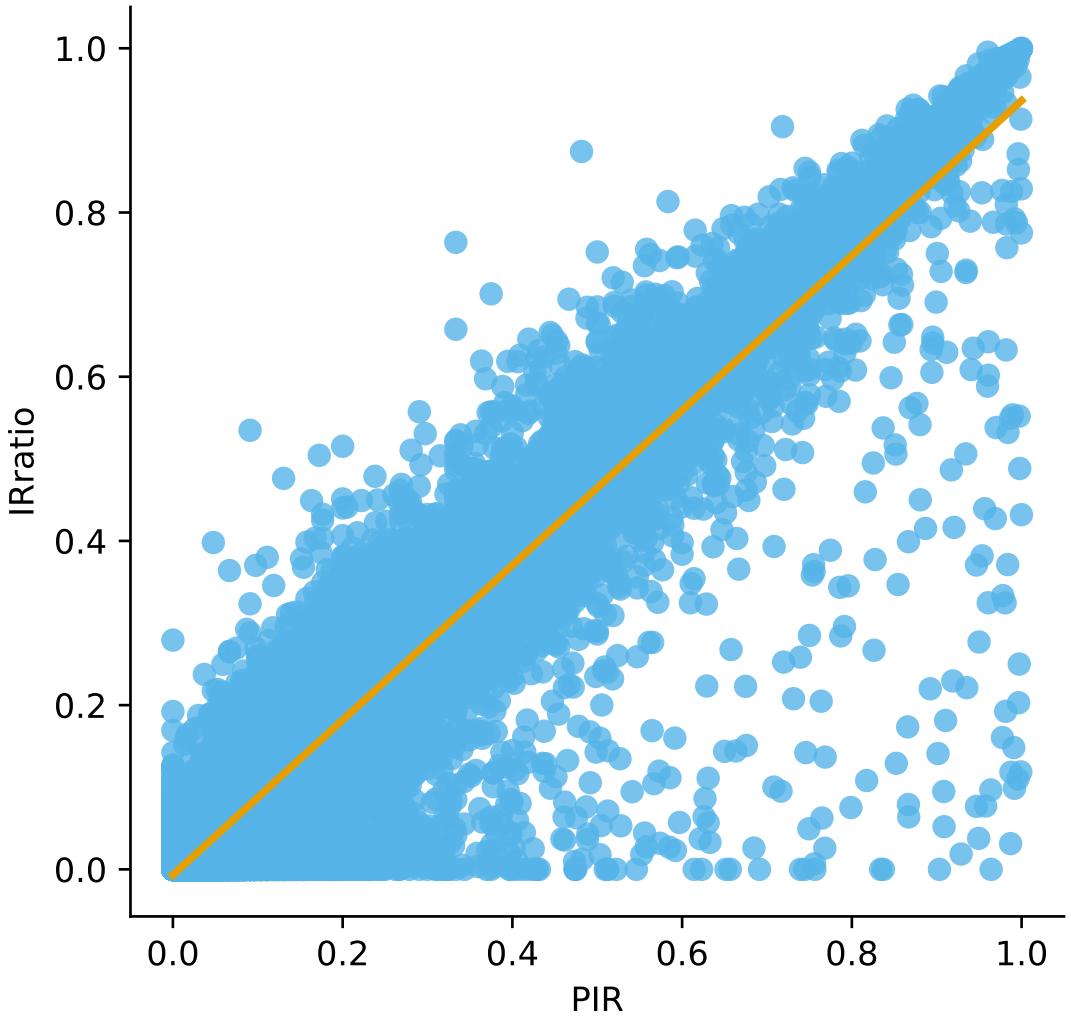


Figure 4.2: Correlation between IRratio and PIR. IRratio and PIR levels are highly correlated (Pearson r=0.97, p-value<0.05).

Table 4.2: ASD risk genes with high or medium levels of intron retention

Chr	Start	End	Gene	PIR	IRratio	Treatment
11	70661678	70698687	*SHANK2*	0.74	0.00	CHX
7	100681771	100681863	*GIGYF1*	0.42	0.39	CHX
1	155521618	155562152	*ASH1L*	0.18	0.00	CHX
1	181776228	181781426	*CACNA1E*	0.18	0.13	CHX
10	27734791	27735220	*MKX*	0.11	0.10	CHX
17	7851549	7851647	*KDM6B*	0.16	0.12	CHX
21	40085765	40087169	*DSCAM*	0.16	0.00	CHX
4	148194862	148259977	*NR3C2*	0.18	0.00	CHX
5	177135200	177135940	*NSD1*	0.20	0.21	CHX
7	100683444	100683549	*GIGYF1*	0.23	0.16	CHX
7	100685146	100685343	*GIGYF1*	0.15	0.14	CHX
9	74497983	74630281	*RORB*	0.15	0.00	CHX

exon 8, 4.3A). To validate the presence of the vulnerable exon in control and *RFX3*+/- human neurons, we performed RT-PCR using forward primers designed to *RFX3* exon 7 and reverse primers designed to *RFX3* exon 9 on neurons with and without pharmacologic NMD blockade (treatment with CHX). Neurons without NMD blockade showed a single band at the expected size for the amplified region containing exon 7, exon 8, and exon 9. Neurons with NMD blockade displayed an additional smaller band at the expected size for a region with exon 8 skipped (4.3B). Relative densitometry analysis demonstrated that 13% of the transcripts in CHX treated neurons had exon 8 skipped. The presence of the vulnerable exon in *RFX3*+/- neurons suggested that it could be a promising therapeutic target to increase *RFX3* levels in haploinsufficient individuals.

4.2.3 *RFX3* expression can be restored with an ASO in human neurons

We designed 25 ASOs tiling the 50 nucleotides flanking exon 8 to promote its inclusion by blocking exon splicing silencers (ESS) and intron splicing silencers (ISS) that typically reside in these re-

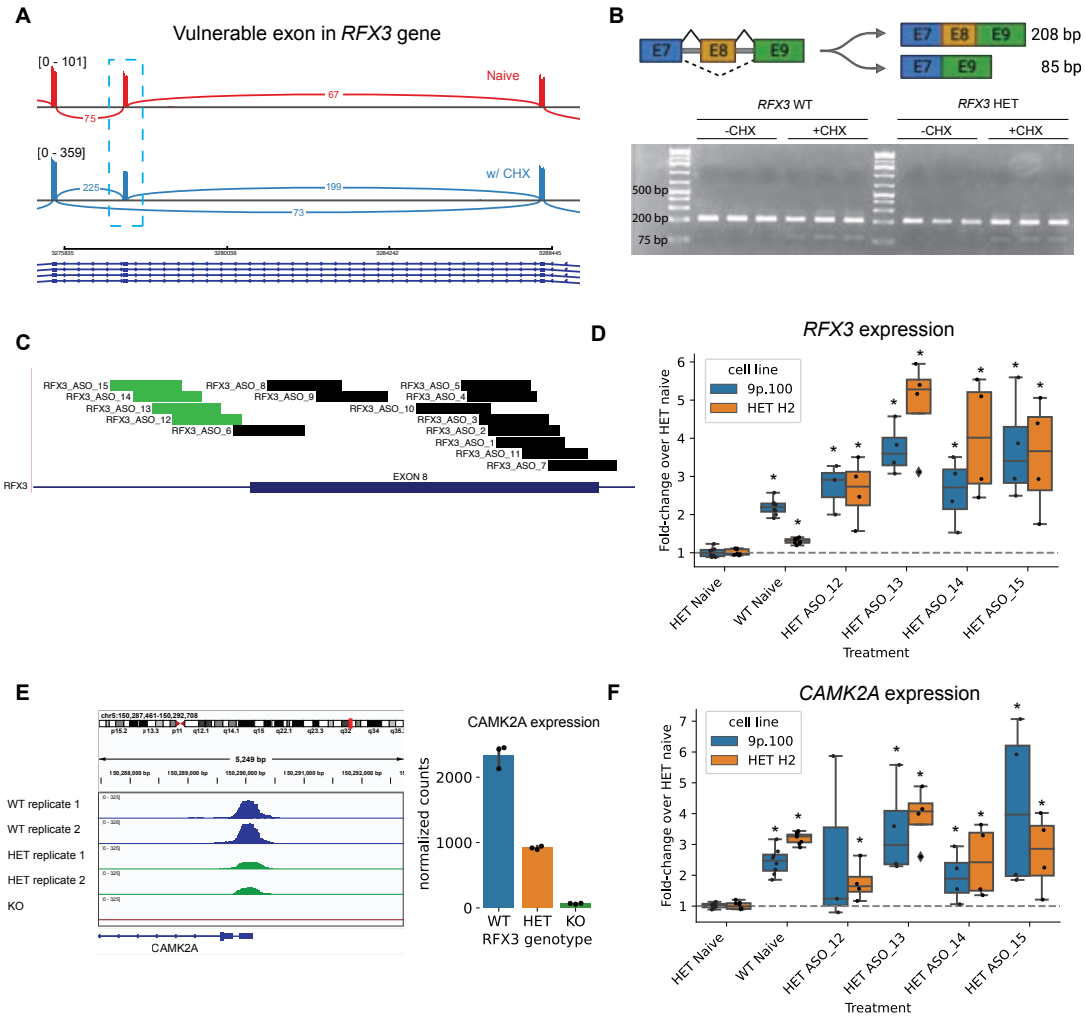


Figure 4.3: ASOs targeting *RFX3* exon 8 increase productive splicing. A. Skipping of *RFX3* exon 8 in iPSC-derived neurons revealed by NMD blockade with cycloheximide (CHX) with RNA-seq. B. Validation of *RFX3* exon 8 skipping in iPSC-derived neurons with NMD blockade by RT-PCR gel visualization. C. Tiled ASO designs. ASOs in green are the four lead ASOs. D. ASOs targeting exon 8 increase *RFX3* expression in haploinsufficient neurons as measured by VE-specific RT-qPCR. n=8 independent wells for naive condition, n=3-4 independent wells for ASO treated conditions. E. *RFX3* binds to the promoter of CAMK2A and is necessary for CAMK2A expression. F. ASOs targeting *RFX3* exon 8 boost expression of the *RFX3* target gene CAMK2A. n=8 independent wells for naive condition, n=3-4 independent wells for ASO treated conditions. *p<0.05, two-way ANOVA with Bonferroni correction. HET: *RFX3* heterozygous line. KO: *RFX3* homozygous knockout line. 9p.100: 9p-minus syndrome iPSC-derived neurons. HET H2: *RFX3* heterozygous iPSC-derived neurons, clone H2. WT: corresponding control line for HET H2 and 9p.100 lines.

gions[102]. ASOs were filtered to remove those with predicted off-target binding, secondary structure formation, and self-complementarity that may lead to oligo dimer formation. This resulted in 15 final ASOs for screening (4.3C). ASOs were synthesized using 2-methoxyethyl nucleotide modifications and uniform phosphorothioate backbones (PS-2'MOE chemistry), the chemical subclass employed in nusinersen, for which there is most clinical experience in patients.

To detect an increase in exon 8 included transcripts specifically, we utilized a TaqMan Gene Expression assay where the probe sits on the boundary of *RFX3* exon 8 and exon 9. We introduced ASOs (10 uM) into human neuronal cultures via gymnotic delivery and identified several candidates that could boost expression of *RFX3* exon 8 containing transcripts by up to 5 fold in *RFX3* haploinsufficient human neurons (4.3D). Four lead ASOs proved capable of boosting *RFX3* expression in human neurons from a patient with 9p-minus syndrome (a chromosomal deletion overlapping *RFX3*) as well(4.3D). These effects were accompanied by increased expression of *CAMK2A*, a downstream target gene of *RFX3* (4.3E-F), thus demonstrating that ASOs can correct transcriptional dysregulation caused by *RFX3* haploinsufficiency.

4.3 Discussion

Haploinsufficiency of hundreds of genes is linked to neurodevelopmental disorders. Restoration of normal gene expression levels is a promising therapeutic approach for haploinsufficiency disorders. In this study, we describe a framework for identifying mRNA processing inefficiencies in human iPSC-derived neurons to catalog events that may be amenable to ASOs. We apply this framework to *RFX3*, a gene that is among the most commonly mutated in Autism Spectrum Disorder (ASD) and attention deficit/hyperactivity disorder (ADHD)[85]. We found that *RFX3* has naturally occurring exon 8 (NM001282116.2) skipping that subjects these transcripts to NMD. We identified several ASOs that could restore *RFX3* beyond wildtype levels, establishing a proof-of-concept that *RFX3*

expression can be modulated using ASOs in human neurons.

These ASOs serve as useful tool compounds for future research on the role of *RFX3* in human neurodevelopment that could have therapeutic implications. Since *RFX3* has a large coding size (9.3 kb), traditional overexpression methods with cDNA and open reading frame (ORF) clones delivered by lentiviral transduction or plasmid transfection (or AAV packaging) are challenging. Therefore, ASOs that can restore or overexpress *RFX3* are an alternative strategy that can further our mechanistic understanding of *RFX3* in neurons and neurodevelopmental disorders.

Future directions include optimizing the ASOs by microwalking and iterative design to maximize potency. Dose response curves will allow us to select the most potent ASOs, which is important for minimizing required doses and potential side effects. Lead ASOs can then be leveraged to define critical windows of *RFX3*'s role in neurodevelopment. First, it would be important to determine whether restoring *RFX3* expression with an ASO at the iPSC stage followed by differentiation into neurons is sufficient to rescue *RFX3* haploinsufficiency-associated transcriptional, cellular, and functional phenotypes. Second, it would be clinically relevant to determine the latest age at which neurons could be treated with an ASO, and which phenotypes, if any, are reversible in mature neurons. Lastly, establishing a mouse model of *RFX3* haploinsufficiency with detailed neurobehavioral characterization will allow us to eventually test whether ASOs can rescue and reverse *RFX3* haploinsufficiency-associated behavioral phenotypes.

ASOs (which can be given as infrequently as every 3-4 months) may be appropriate for debilitating forms of neurodevelopmental disorders with severe cognitive and behavioral consequences. These could include individuals with haploinsufficiency of *RFX3* but may not be limited to those individuals. For instance, we note that *RFX3* is located on chromosome 9p, and may be a strong contributor to 9p-minus syndrome, a severe neurodevelopmental disorder in which over 80% of individuals have a deletion spanning *RFX3* [227]. Ongoing work involves studying iPSC neurons from 9p-minus individuals to examine this overlap. Individuals with idiopathic ASD were found

to have broad, cortex-wide downregulation of synaptic plasticity genes compared to controls[72]. Our data suggest that synaptic related genes and genes critical to ASD pathogenesis are targets of *RFX3* (Chapter 3). The fact that *RFX3* is positioned upstream of ASD relevant genes makes it a useful therapeutic target that may be beneficial not only for individuals with *RFX3* haploinsufficiency but potentially other forms of ASD, since *RFX3* upregulation may reactivate expression of synaptic pathways dysregulated in ASD.

4.4 Acknowledgements

This work was supported by the National Institute of Health grant R01MH113761 and the G. Harold & Leila Y. Mathers Foundation. J.L is supported by award Number T32GM007753 and T32GM144273 from the National Institute of General Medical Sciences, and the Ruth L. Kirschstein NRSA F30 Fellowship (1F30MH128995). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health.

4.5 Author Contributions

Conceptualization, T.W.Y, J.L, B.Z, L.F.B, D.C; Methodology, T.W.Y, J.L, B.Z, L.F.B, D.C; Investigation, J.L, B.Z, L.F.B, D.C; Formal Analysis, T.W.Y, J.L, L.F.B; Validation, K.P, H.W, J.L; Supervision, T.W.Y; Writing – Original Draft, J.L; Writing – Review & Editing, T.W.Y, J.L

T.W.Y: Timothy W. Yu; J.L: Jenny Lai; B.Z: Boxun Zhao; L.F.B: Luca Fusar Bassini; D.C: Didem Demirbas Cakici.

4.6 Methods

4.6.1 Differentiation of iPSCs to neurons

Human iPSC-derived neurons (iNs) were generated from iPSCs transduced with the lentiviruses Tet-O-Ngn2-Puro and FUW-M2rtTA based on a previously published protocol[299]. On day -1 of differentiation, iPSC colonies were dissociated into single-cells using Accutase (STEMCELL Technologies) and 4-8 million cells were plated on Matrigel coated 10 cm dishes with complete mTeSR plus medium supplemented with Y-27632 (10 uM). On day 0, cells were fed N2 media (DMEM/F-12 media, 1X N2, 1X Nonessential Amino Acids) supplemented with doxycycline (2 μg/mL), BDNF (10 ng/mL), NT3 (10 ng/mL), and laminin (0.2 μg/mL). On day 1, media was replaced with N2 media with puromycin (1 μg/mL) in addition to the supplements listed above. On day 2, cells were fed B27 media (Neurobasal media, 1X B27, 1X Glutamax) supplemented with puromycin (1 μg/mL), doxycycline (2 μg/mL), Ara-C (2 uM), BDNF (10 ng/mL), NT3 (10 ng/mL), and laminin (0.2 μg/mL). On day 3, cells were dissociated into single cells using Accutase and replated onto polyethylenimine (PEI)/laminin (0.07% PEI and laminin diluted 1:200 in sterile water) coated 96-well plates (10,000 cells/well) with B27 media supplemented with Y-27632 (10 μM), doxycycline (2 μg/mL), Ara-C (2 μM), BDNF (10 ng/mL), NT3 (10 ng/mL), and laminin (0.2 μg/mL). On days 5, 7, 10, and 14, cells were fed with a half-media change of Conditioned Sudhof Neuronal Growth Medium (1:1 ratio of Astrocyte Conditioned Media and Neurobasal Media, 1X B27, Glutamax, $NaHCO_3$, and transferrin) supplemented with BDNF (10 ng/mL), NT3 (10 ng/mL), and laminin (0.2 μg/mL). Cells were fed with a half-media change of Conditioned Sudhof Neuronal Growth Medium as described above once a week. Cells were maintained in a humidified incubator (5% CO_2, 37°C).

4.6.2 Cycloheximide treatment

1000X stock of cycloheximide was prepared by dissolving cycloheximide in DMSO (final concentration 0.1 mg/uL). 3 hours prior to RNA harvest, neurons were treated with cycloheximide (final concentration 100 ug/mL) by full media replacement with cycloheximide containing media. Cells were incubated for 3 hours at 37C.

4.6.3 ASO delivery

Neuronal cultures were treated with ASO by gymnosis. On the day of treatment, media was fully replaced with media containing ASO (final concentration 10 uM). Cells were incubated for 4 days at 37C.

4.6.4 Bulk RNA-sequencing and analysis

Total RNA was harvested from samples using the PureLink RNA Mini kit (Life Technologies Catalog # 12183018A). Libraries were prepared with the KAPA mRNA prep. Sequencing was performed on Illumina NovaSeq, 2x150bp configuration at approximately 100X depth per sample. For identifying poison exons and skipped exons, reads were aligned to hg38 using STAR v2.7.9a[58] followed by processing with a custom built pipeline in our lab (Yu Lab, unpublished) that identifies differential splicing events by quantifying junctional reads at exon/intron boundaries using LeafCutter[154] and SAMtools[47]. To estimate intron retention, we applied the bioinformatic tool IRFinder[161]. This tool calculates IRratio using (IntronDepth / (IntronDepth + Max(Splice left, Splice right)), where IntronDepth = median depth of the intronic region (number of reads that map over a given bp), not considering excluded regions (bases with the top and bottom 30% of intronic depth), Splice left = number of reads that map the 5' flanking exon to another exon in the gene , and Splice right = number of reads that map the 3' flanking exon to another exon in the gene. Using the

output from IRFinder, we also computed Percent Intron Retention as described in Braunschweig et al., 2014[29]. where PIR = percent intron retention, E1I = exon 1 - intron junction reads, IE2 = intron - exon 2 junction reads, and E1E2 = exon 1 - exon 2 junction reads.

$$PIR = \frac{average(E1I, IE2)}{average(E1I, IE2) + E1E2}$$

4.6.5 RT-PCR

100 ng RNA per sample was reverse transcribed into cDNA using Cells-to-CT RT buffer and RT enzyme (Thermo Fisher Scientific Cat# A35378) consisting of 22.5 ul RNA and 27.5 ul of Cells-to-CT RT Master Mix. The reverse transcription reactions were incubated at 37C for 1 hour, and 95C for 5 min. The PCR protocol consisted of Phusion Hot Start II DNA Polymerase (Thermo Fisher Scientific), 5x Phusion GC Buffer, 10 mM dNTPs, 10 uM Forward primer, 10 uM Reverse primer, and 2 ul cDNA, followed by visualization on a 2% agarose gel. The forward primer was 5'-GACAACAACCCATGCAACAG-3'. The reverse primer was 5'-CCATCTGGCAGAGAAGAGATTT-3'. The expected amplicon length was 208 bp for exon 8 included amplicons and 85 bp for exon 8 excluded amplicons.

4.6.6 Quantitative RT-PCR

RNA was reverse transcribed into cDNA using Cells-to-CT RT buffer and RT enzyme (Thermo Fisher Scientific Cat# A35378) consisting of 22.5 ul RNA and 27.5 ul of Cells-to-CT RT Master Mix. The reverse transcription reactions were incubated at 37C for 1 hour, and 95C for 5 min. qPCR was performed using TaqMan Fast Advanced Master Mix (Thermo Fisher Scientific Cat# A35378) and Gene Expression Assays (Thermo Fisher Scientific Cat# 4331182) with 2 ul input cDNA. The thermocycling parameters were 50C for 2 min, 95C for 2 min, 40 cycles of 95C for 1

second and 60C for 20 seconds. TaqMan Gene Expression Assays included RFX3 Hs01060440_m1, RFX3 Hs01060430_m1, CAMK2A Hs00947041_m1, Human GAPDH Endogenous Control (VIC/MGB probe, primer limited, Thermo Fisher Scientific Cat# 4326317E).

5

Conclusion

5.1 Thesis summary

This thesis sought to understand the molecular and cellular underpinnings of human neurodevelopment and degeneration through functional genomic analysis of human brain tissue and iPSC-derived models of two pediatric neurogenetic disorders, Ataxia Telangiectasia (A-T) and *RFX3* haploinsufficiency associated autism spectrum disorder (ASD). Through functional genomic dissection

of human-based disease models, we have uncovered mechanisms by which pathogenic variants in the genes *ATM* and *RFX3* lead to their respective disease phenotypes.

Chapter 1 reveals novel insights into the role of microglia in the neurodegenerative disease Ataxia Telangiectasia (A-T) leveraging single-nucleus RNA-sequencing of patient postmortem brain and a human patient induced pluripotent stem cell (iPSC)-derived microglia and neuron model system. While *ATM* loss-of-function has long been identified as the genetic cause of Ataxia Telangiectasia (A-T), precisely how this genetic mutation leads to selective and progressive degeneration of cerebellar Purkinje and granule neurons remains unclear. *ATM* expression is enriched in microglia, the resident immune cell of the central nervous system, throughout cerebellar development and adulthood. Microglial activation has been strongly implicated in neurodegenerative disease and has been observed in rodent and cellular models of *ATM* deficiency. Here, we find evidence of prominent inflammation of microglia in cerebellum from A-T patients using single-nucleus RNA-sequencing. We find that A-T microglia exhibit transcriptomic signatures of aging and neurodegenerative disease associated microglia. Pseudotime analysis revealed that activation of A-T microglia preceded upregulation of apoptosis related genes in granule and Purkinje neurons, and microglia exhibited increased neurotoxic cytokine signaling to granule and Purkinje neurons in A-T. To confirm these findings experimentally, we studied microglia and neurons that we generated from A-T patient vs. control induced pluripotent stem cells (iPSCs). Transcriptomic profiling of A-T iPSC-derived microglia revealed cell-intrinsic microglial activation of cytokine production and innate immune response pathways compared to controls. Furthermore, adding A-T microglia to co-cultures with either control or A-T iPSC-derived neurons was sufficient to induce cytotoxicity. Taken together, these studies reveal that cell-intrinsic microglial activation may play a critical role in the development and progression of neurodegeneration in Ataxia Telangiectasia.

Chapter 2 delineates a novel neurodevelopmental disorder associated with deleterious variants in the *RFX* family of transcription factors. The *RFX* family of transcription factors are evolutionarily

conserved with known roles in ciliogenesis. While homozygous loss of *RFX* genes leads to severe ciliary defects and is generally embryonic lethal in rodent models, a heterozygous phenotype has not been previously described. We identified 38 individuals with *de novo* variants in *RFX3*, *RFX4*, and *RFX7*, and demonstrated that haploinsufficiency of these genes is associated with a neurobehavioral phenotype including ASD, ADHD, intellectual disability, and dysregulated behavior. Examination of single-cell transcriptomes from developing and adult human cortex revealed enriched expression of *RFX3* in upper layer glutamatergic neurons, *RFX4* in astrocytes, and *RFX7* in inhibitory and excitatory neurons. Moreover,the RFX binding motif was enriched in enhancer regions associated with ASD risk genes. These results establish a likely role of deleterious variation in *RFX3*, *RFX4*, and *RFX7* in cases of monogenic intellectual disability, ADHD and ASD, and position these genes as potentially critical transcriptional regulators of neurobiological pathways associated with neurodevelopmental disease pathogenesis.

Chapter 3 delves into the molecular and cellular mechanisms of *RFX3* in human neurodevelopment using genetic models of *RFX3* dosage in iPSC-derived dorsal forebrain organoids and neurons. Organoid single-cell RNA sequencing revealed delayed neuronal differentiation with biallelic loss of *RFX3*, but not monoallelic loss of *RFX3* as is observed in patients. Transcriptomic and binding analyses across *RFX3* dosages demonstrated disruption of ciliary target gene expression in *RFX3*-/- neurons, but not *RFX3*+/- neurons. Instead, *RFX3*+/- neurons demonstrated disruption of synaptic target gene expression associated with decreased synchronization of neural activity. Binding of *RFX3* co-occurred with CREB near activity-dependent genes, and *RFX3* deficiency led to impaired induction of CREB targets in response to neuronal depolarization associated with decreased CREB binding. This chapter highlights a novel role of the ciliogenic transcription factor *RFX3* in shaping activity-dependent responses in human neurons that is disrupted by haploinsufficiency.

Chapter 4 highlights the potential of gene based therapies for neurogenetic disorders and describes a framework for using human neurons to identify antisense oligonucleotide (ASO) amenable

targets that may be relevant for neurodevelopmental conditions like ASD. ASOs are short nucleotide sequences that can be designed to bind specific RNA targets to modulate gene expression. Monogenic forms of ASD are often caused by heterozygous loss-of-function mutations, leaving the other allele intact. We identify ASD risk genes with wildtype transcripts degraded by nonsense-mediated decay due to mRNA processing inefficiencies (poison exon inclusion, intron retention, vulnerable exon exclusion) in human iPSC-derived neurons. Using *RFX3* as a model, we demonstrate that ASOs can modulate splicing to correct mRNA processing inefficiencies to increase gene expression. This chapter provides a map of potential therapeutic opportunities for monogenic forms of ASD.

5.2 Future directions

5.2.1 Future directions in understanding neurodegeneration in A-T

Our finding that microglia are activated in A-T and contribute to increased cytotoxicity of neuronal cultures raises several key questions that would be of great interest for future research. One central remaining question is why the cerebellum specifically degenerates in A-T while the cortex is generally intact. We found that A-T microglia from the cerebellum have greater upregulation of pro-inflammatory genes than A-T microglia from the prefrontal cortex (PFC), but it remains unclear whether this observation is completely due to a cell-intrinsic microglia phenotype or may also be influenced by microglial activation by differential disease pathology in the cerebellum compared to cortex. Selective cerebellar degeneration may arise from 1) intrinsic differences between cerebellar and PFC microglia, such that A-T cerebellar microglia induce more cytotoxicity, or 2) differential neuronal vulnerability between cerebellum and PFC. One method to compare the impact of *ATM* deficiency between cerebellar and PFC microglia is to establish primary cultures of freshly isolated rodent or human cerebellar and PFC microglia, induce *ATM* deficiency with *ATM* siRNA or small

molecule ATM inhibitors (such as AZD1390), and assess their levels of activation by comparing their transcriptomes and cytokine secretion. By inducing *ATM* deficiency in homogeneous cultures of cerebellar and PFC microglia, we can look for brain region specific differences in cell-intrinsic phenotypes. Alternatively, to determine whether differential neuronal subtype vulnerability may be a contributing mechanism, co-culture of A-T patient iPSC-derived microglia with dorsal forebrain organoids and cerebellar organoids combined with single-cell transcriptome and spatial analysis could provide insight on differential microglia-neuronal interactions in the two brain regions. For example, microglia may have a greater predisposition to associate with synapses of particular neuron types based on differential expression of "eat me" signals, or pathologically activated microglia may secrete neurotoxic cytokines with receptor expression that varies between neuron types. However, one present limitation is that organoid models for the cerebellum are relatively less established and do not develop neurons with the same complexity characteristic of Purkinje neurons. Until organoid models of the cerebellum become more advanced, an alternative approach is to transplant human A-T patient iPSC-derived microglia into rodent models to understand their impact on cortical and cerebellar degeneration in a physiological context.

Another key question is whether restoration of *ATM* expression can modify disease progression, and whether restoration in particular cell types is sufficient. To understand the role of loss of *ATM* specifically in microglia, we can rescue *ATM* expression in our A-T patient iPSC-derived microglia. In fact, this is experimentally feasible today, since our laboratory has already developed a splice-switching ASO that successfully rescues the pathogenic *ATM* c.7865C>T mutation found in the patient iPSC-derived microglia used in our experiment [128]. This experimental ASO is currently in clinical trial [128], and could be used to determine whether microglial rescue is sufficient to prevent further cytotoxicity of neuronal co-cultures. We can also co-culture A-T microglia with A-T iPSC-derived neurons that are treated with ASO to determine whether *ATM* rescue in neurons is sufficient to reduce neurotoxicity. These complementary experiments will reveal the microglia and

neuronal mechanisms contributing to degeneration in A-T. These tractable systems also allow modeling of disease progression by sampling over time to determine the order in which cell type specific perturbations occur and define the critical windows and cell types when genomic interventions may be beneficial. Taken together, these future experiments could illuminate the precise mechanisms by which microglia-neuron interactions may contribute to selective cerebellar degeneration in A-T.

5.2.2 Future directions in understanding RFX-associated neurodevelopmental disorders

Since our initial study on deleterious variants in RFX associated with neurodevelopmental disorders, we have been in contact with 9 additional cases (4 *RFX3*, 4 *RFX7*, 1 *RFX4*). As a result, there are 47 known cases of neurodevelopmental disorders associated with *RFX3*, *RFX4*, or *RFX7* variants. These RFX members share a DNA binding motif, but have enriched expression in distinct cell types in the adult human cortex. While *RFX3* and *RFX7* are both expressed in cortical glutamatergic neurons, *RFX7* also has enriched expression in inhibitory neurons, and *RFX4* has enriched expression in astrocytes. This raises the question of how deficiency of *RFX3*, *RFX4*, and *RFX7* converge on similar neurobehavioral phenotypes despite expression and therefore potentially disrupted function in distinct cell types. Moreover, while we have characterized *RFX3* binding sites in human neurons in this thesis, the binding sites of *RFX4* and *RFX7* in the human brain remain unstudied. Therefore, an area of future research is to extend the *RFX3* functional genomic studies described in Chapter 3 to iPSC-derived models of *RFX4* and *RFX7* haploinsufficiency to identify whether there is a common set of targets with disrupted expression in *RFX3*, *RFX4* and *RFX7* deficiency that may contribute to the resulting neurodevelopmental disorder.

In addition, future work remains to further understand precisely how *RFX3* shapes activity dependent responses. While we have demonstrated that *RFX3* constitutively promotes CREB binding to a subset of promoters in the genome of glutamatergic cortical neurons, it remains un-

clear whether *RFX3* promotes binding at only these specific sites and if so, why. One hypothesis is that *RFX3* modulates CREB binding to these sites in a cell type specific manner, such as in L2/3 neurons specifically. Future advances in single-cell CUT&RUN-sequencing to profile *RFX3* and CREB binding in a diversity of cell types, such as excitatory neurons, inhibitory neurons, and glia, could provide insight on this question. It also remains unknown whether *RFX3* plays a permissive or instructive role in activity dependent responses. We found that *RFX3* deficiency leads to decreased CREB binding associated with impaired induction of immediate early genes in response to neuronal depolarization with KCl, which is consistent with either a permissive or an instructive role. To determine whether *RFX3* is permissive or instructive, depolarization induced immediate early gene expression can be assessed in neurons with *RFX3* overexpression. If *RFX3* has an instructive role, we would expect to see enhanced activity dependent responses in neurons with *RFX3* overexpression. This would provide insight into whether *RFX3* overexpression may be an approach to synthetically restore activity dependent responses in neurodevelopmental disorders caused by pathogenic variants in activity dependent components. Furthermore, resolving the activity dependent complex components that RFX3 binds through co-immunoprecipitation studies in neurons with and without stimulation can illuminate whether *RFX3* also has roles in recruiting or stabilizing binding of other key activity dependent complex factors. Finally, while this thesis has focused on human-based models of neurogenetic disorders, detailed phenotyping of a *RFX3*+/- mouse model in light of our current human iPSC-derived findings would be useful for linking alterations at the molecular and cellular level to behavioral phenotypes and in a more physiological context. For example, we could investigate whether activity dependent responses to physiological stimuli, such as light, sound, or touch, are blunted in *RFX3*+/- mice and how this correlates with neurobehavioral phenotypes.

5.3 Future directions in ASO development for monogenic neurological disorders

In Chapter 4, we cataloged genome-wide intron retention events in highly penetrant ASD risk genes as potential candidates for ASO therapy development. Future work involves experimental validation of the intron retention events with RT-PCR, design and screening of ASOs to enhance intron splicing, and demonstration that the ASO can increase productive mRNA and protein levels in neuronal models of haploinsufficiency.

We also identified several ASOs that could increase productive *RFX3* splicing and rescue expression of a downstream target gene. Expanding this work to characterize which *RFX3* deficiency associated transcriptional alterations are 1) rescuable and 2) reversible with ASO treatment at different neuronal stages will help define critical windows when ASO therapy may be most beneficial. Future studies may also involve determining whether restoring *RFX3* expression in mature neurons can rescue activity-dependent gene induction in response to depolarization, and rescue neuronal network activity phenotypes. The ASOs identified serve both as scientific tools to elucidate *RFX3*'s role by providing a means to restore *RFX3* expression (since *RFX3* is too large for traditional expression methods such as cDNA and ORF constructs), and as a potential therapeutic avenue for understanding which deficits associated with *RFX3* deficiency may be reversible.

5.4 Future directions for the field of functional genomics in iPSC-derived model systems

The past decade has experienced a tremendous burgeoning of human disease modeling with iPSCs and molecular understanding enabled by multi-omic technologies. Technological improvements in both iPSC modeling and functional genomics promise to deepen our understanding of human

genetic diseases. There are several key areas where advances may transform our ability to study neurogenetic disorders. First, improvements in organoid modeling of specialized brain regions, such as the cerebellum, and iPSC differentiation into specific cell types, can enable the study of genetic variants on disorders that affect specific regions or cell types in the brain. The rapid growth of single-cell atlases of the human brain in development and adulthood can elucidate critical cell type enriched transcription factors that may be overexpressed to efficiently differentiate iPSCs into a variety of cell types. In addition, a major challenge in organoid modeling is difficulty of maintaining long-term culture due to internal cellular necrosis from lack of oxygen and nutrient delivery. Innovations to incorporate vascularization and nutrient delivery to organoids can enable long-term culture which may be useful for modeling age-associated and neurodegenerative diseases. Moreover, advances in 3D organoid phenotyping, including electrophysiology, live organoid imaging, and single-cell omics will improve our understanding of cell type specific involvement in neurogenetic disease.

Lastly, continuing innovations in single cell technologies and genomics has the potential to address a significant challenge in neuroscience: the vast diversity of cell types and states within the brain, ranging from transcriptomes, to morphology, anatomical localization, and synaptic targets. Recent developments in single-synapse sequencing [197], single-cell epigenomic profiling [21], and spatial technologies [276] combined with established single-cell RNA-sequencing applied to human brain tissue and iPSC-derived models holds the promise of dissecting critical underlying human-specific biology of neurogenetic disorders.

5.5 Concluding remarks

The work presented in this thesis illustrates the power of combining functional genomic studies in patient-derived tissues with neurobiologic experimentation in human iPSC-derived models to understand the neurogenetic basis of human brain development and degeneration.

A

Appendix

A.1 Chapter 1 Supplemental Figures

Chapter 1 contains the following supplemental figures: A.1 A.2 A.3 A.4 A.5 A.6 A.7 A.8 A.9 A.10 A.11

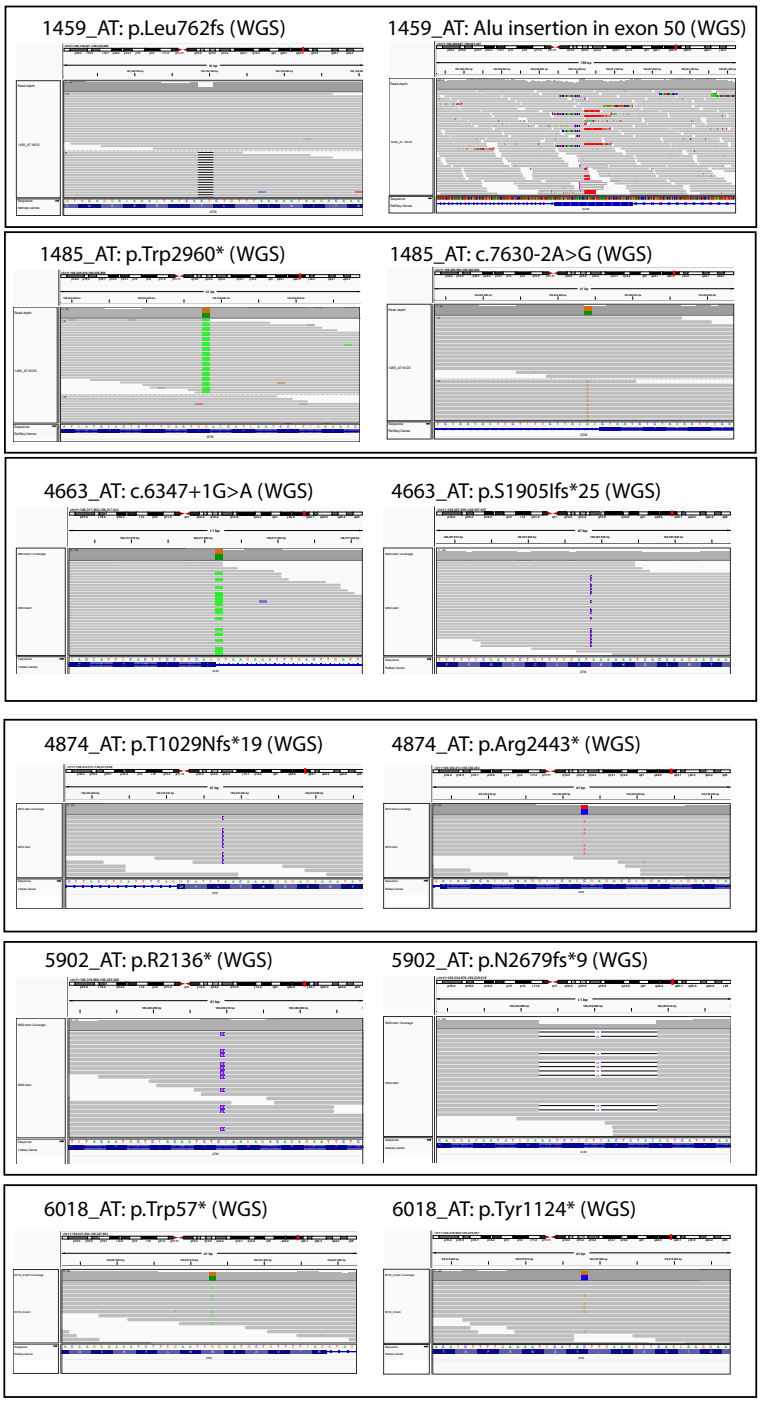


Figure A.1: Pathogenic ATM variants in A-T cases. Alignment tracks of whole genome sequencing (WGS) reads at the ATM locus showing pathogenic variants in A-T cases.

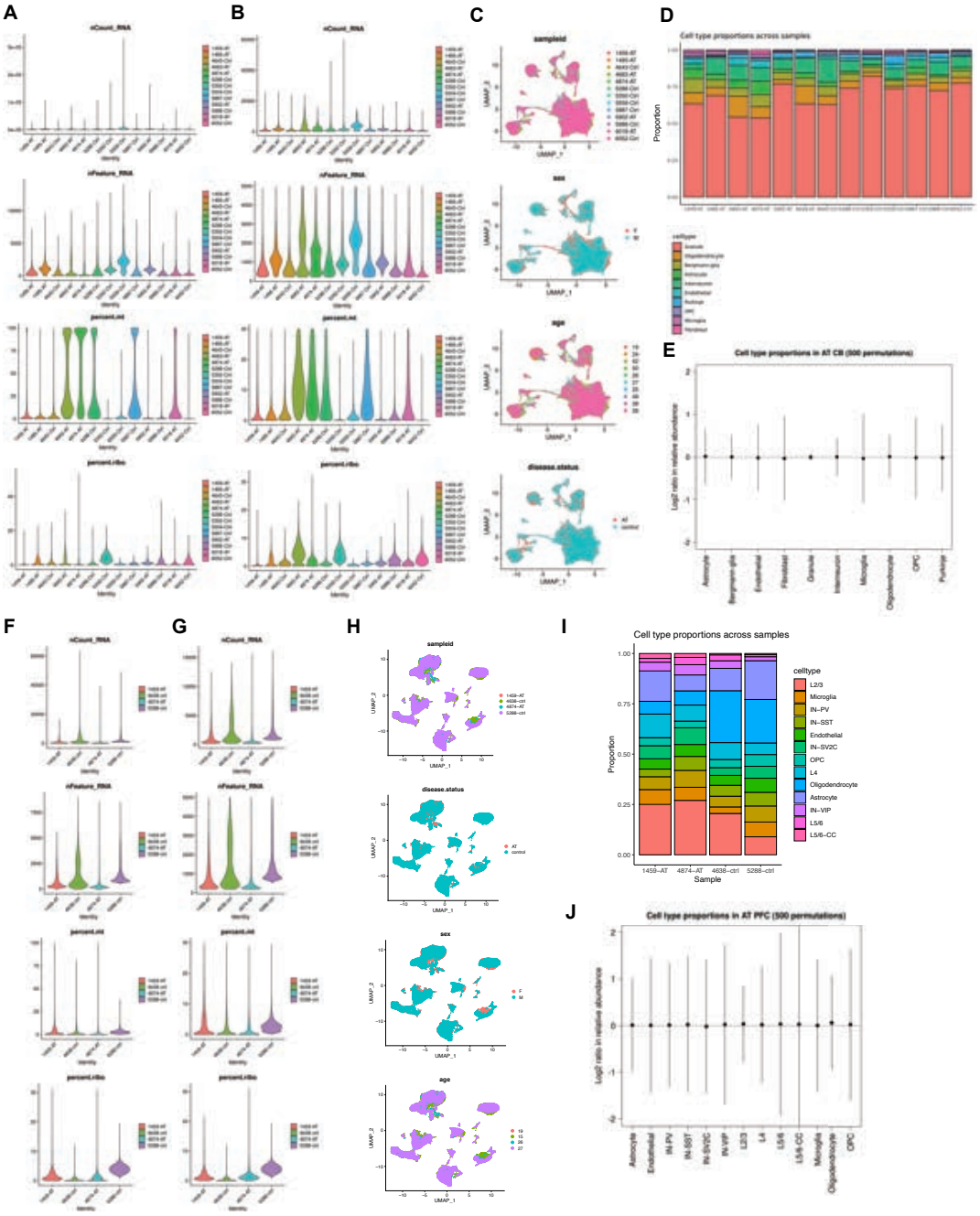


Figure A.2: Quality control information for snRNA-seq data from cerebellum and PFC. A. Transcript (nCount_RNA), gene (nFeature_RNA), percent transcripts from mitochondrial genes, and percent transcripts from ribosomal genes per cerebellar sample before quality control filtering. B. Transcript (nCount_RNA), gene (nFeature_RNA), percent transcripts from mitochondrial genes, and percent transcripts from ribosomal genes per cerebellar sample after quality control filtering.

Figure A.2 (Continued). C. UMAP plots of cerebellar data showing integration of samples, sex, age, and disease status. D. Stacked barplots showing proportion of each cell type per cerebellar sample. E. Posterior distribution of log2 ratio of cell type proportions in A-T cerebellum after 500 random permutations. F. Transcript (nCount_RNA), gene (nFeature_RNA), percent transcripts from mitochondrial genes, and percent transcripts from ribosomal genes per PFC sample before quality control filtering. G. Transcript (nCount_RNA), gene (nFeature_RNA), percent transcripts from mitochondrial genes, and percent transcripts from ribosomal genes per PFC sample after quality control filtering. H. UMAP plots of cerebellar data showing integration of samples, disease status, sex, and age. I. Stacked barplots showing proportion of each cell type per PFC sample. J. Posterior distribution of log2 ratio of cell type proportions in A-T PFC after 500 random permutations.

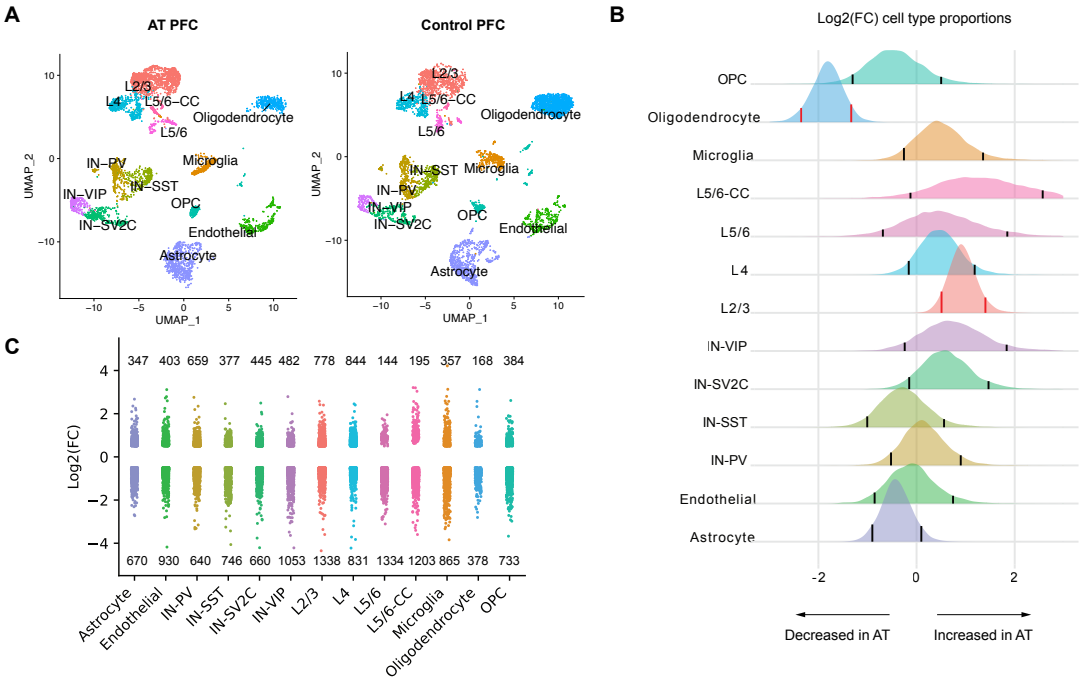


Figure A.3: snRNA-seq data of A-T PFC. A. UMAP plot of major cell types in AT and control human PFC, downsampled to 10,000 cells per condition only for visualization purpose. B. Relative abundance of cell types in AT versus control PFC, shown as the posterior distribution of Log2(proportion in AT/proportion in control) with 89% credible interval. Red bars highlight credible intervals that do not overlap 0. Bolded cell type labels indicate a significant difference in relative abundance. C. Differentially expressed genes (DEGs) in each cell type with FDR<0.05, |Log2FoldChange|>0.50. Each dot represents a significantly differentially expressed gene.

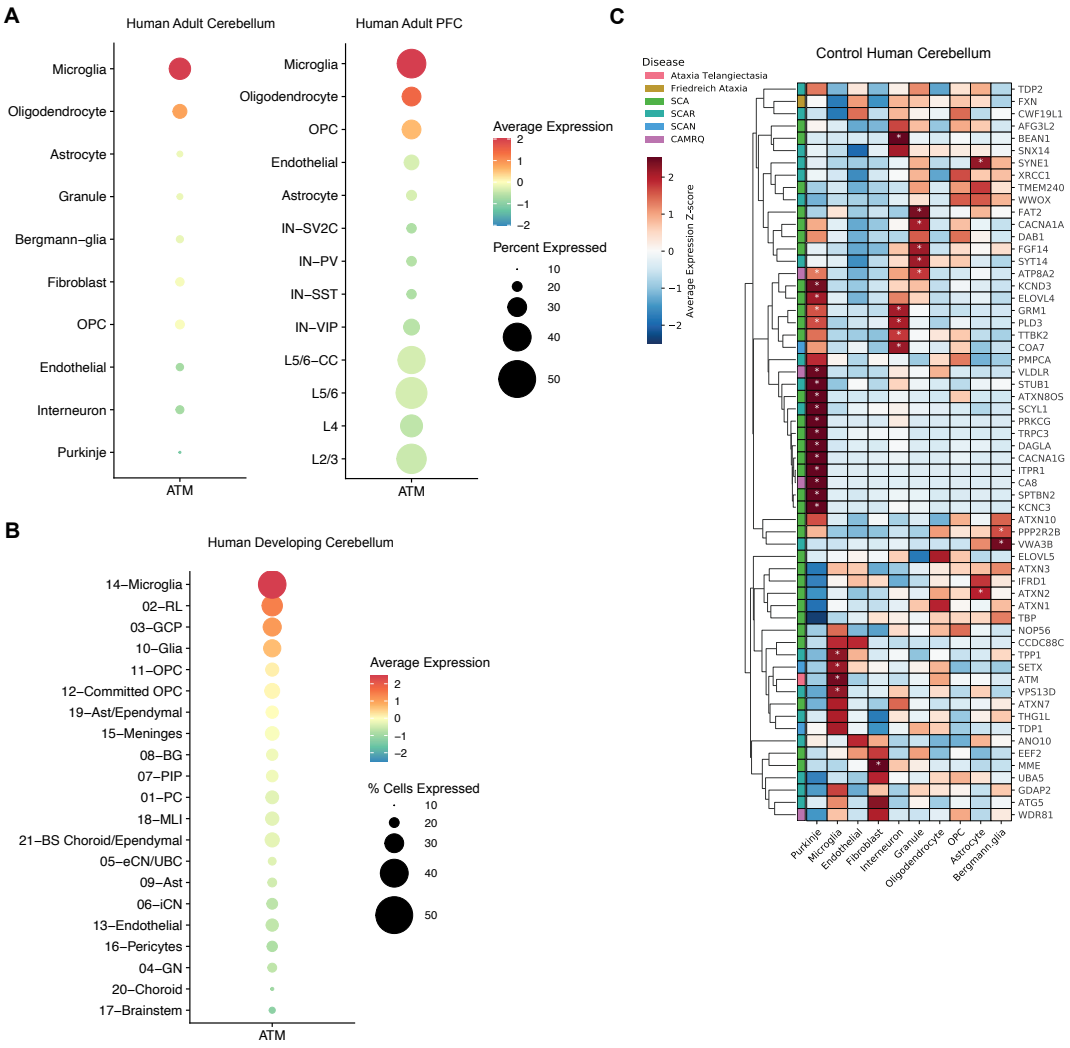


Figure A.4: Enriched expression of monogenic cerebellar disease genes in specific cell types. A. Dotplot of average scaled expression of ATM in cell types of the adult human cerebellum and prefrontal cortex, and B. developing human cerebellum (Aldinger et al., 2021). Microglia have the highest expression of ATM out of all cell types in the cerebellum and PFC. Size of dot represents percentage of single cells expressing the gene. Expression scaled to mean of 0 and standard deviation of 1. C. Heatmap of cerebellar ataxia disease gene expression across cell types in control human cerebellum. Color represents centered and scaled log-normalized expression. *Enrichment p-value <0.05.

Figure A.4 (Continued). OPC: oligodendrocyte precursor cell, IN-SV2C: SV2C expressing inhibitory neuron; IN-PV: PVALB expressing inhibitory neuron; IN-SST: SST expressing inhibitory neuron; L5/6-CC: layer 5/6 cortico-cortical neuron; L5/6: layer 5/6 neuron; L4: layer 4 neuron; L2/3: layer 2/3 neuron; 01-PC: Purkinje cells; 02-RL: Rhombic lip; 03-GCP: Granule cell progenitors; 04-GN: Granule neurons; 05-eCN/UBC: Excitatory cerebellar nuclei neurons/Unipolar brush cells; 06-iCN: Inhibitory cerebellar nuclei neurons; 07-PIP: PAX2+ interneuron progenitors; 08-BG: Bergmann glia; 09-Ast: Astrocytes; 10-Glia: Glia; 11-OPC: Oligodendrocyte precursor cells; 12-Committed OPC: Committed oligodendrocyte precursor cells; 13-Endothelial: Endothelial cells; 14-Microglia: Microglia; 15-Meninges: Meninges; 16-Pericytes: Pericytes; 17-Brainstem: Brainstem; 18-MLI: Molecular layer interneurons; 19-Ast/Ependymal: Astrocytes/ependymal cells; 20-Choroid: Choroid plexus; 21-BS Choroid/Ependymal: Choroid plexus/ependymal cell; SCA: spinocerebellar ataxia; SCAR: spinocerebellar ataxia, recessive; SCAN: spinocerebellar ataxia, autosomal recessive, with axonal neuropathy; CAMRQ: cerebellar ataxia, impaired intellectual development, and dysequilibrium syndrome.

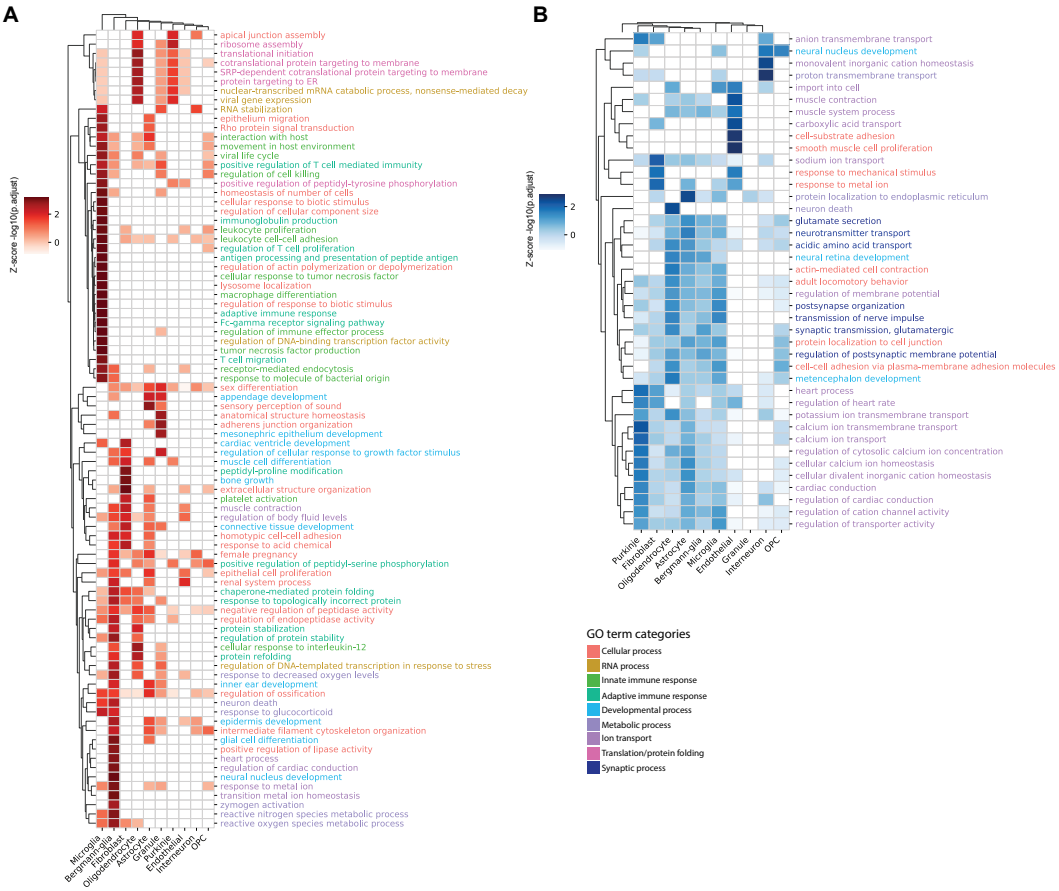


Figure A.5: Gene Ontology enrichment of upregulated and downregulated genes across cell types in A-T cerebellum. A. Heatmap of enriched pathways among upregulated DEGs in cell types from A-T cerebellum. B. Heatmap of enriched pathways among downregulated DEGs in cell types from A-T cerebellum. Color represents z-score of pathway significance (-log10(p. adjusted)) and only significant pathways (p. adj <0.05) are colored.

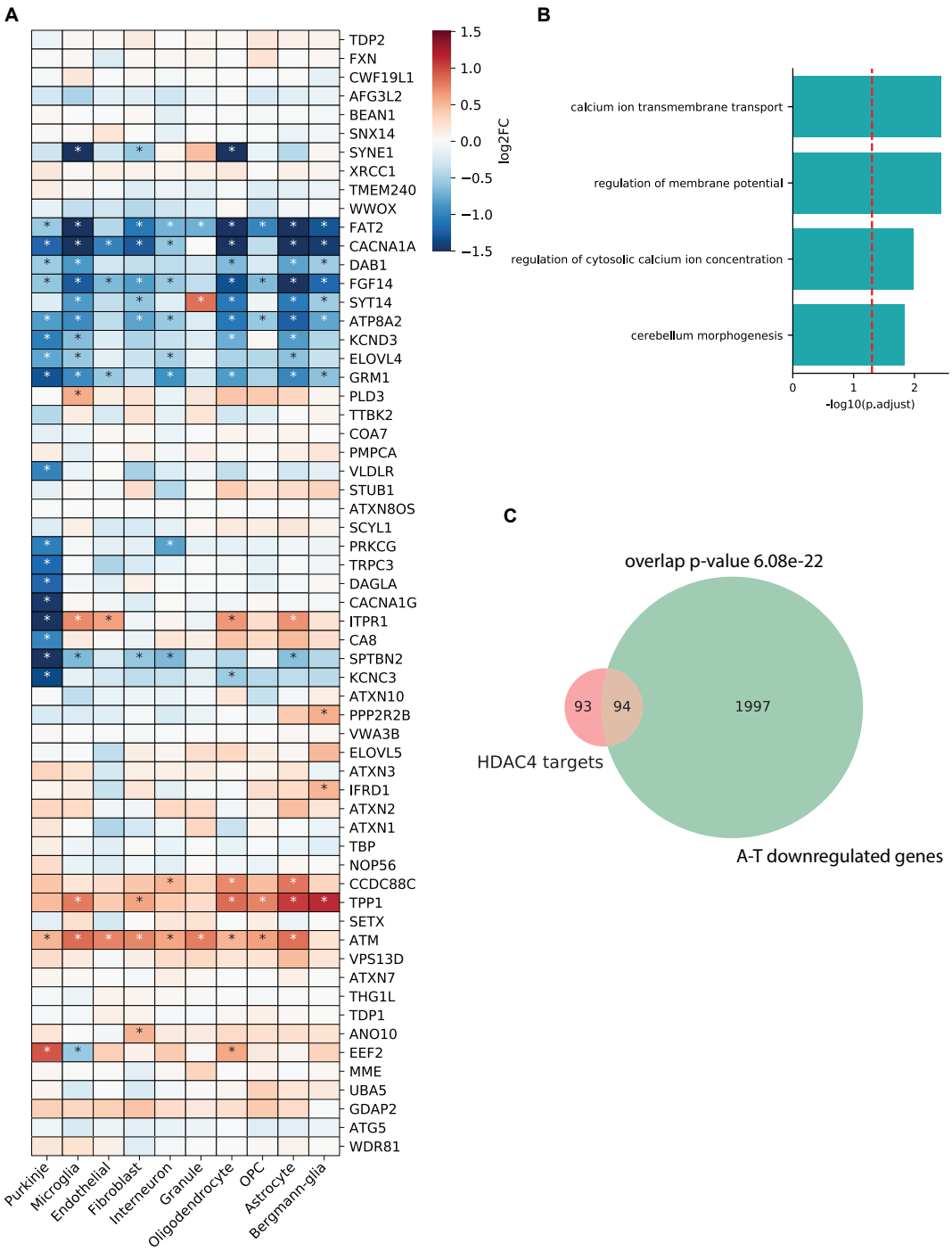


Figure A.6: Dysregulation of hereditary ataxia genes in A-T.

Figure A.6 (Continued). A. Heatmap showing log2 fold-change of hereditary ataxia genes in A-T cerebellum. *FDR<0.05. ATM expression increased in several cell types in A-T cerebellum, suggesting that lack of ATM function induces compensatory increases in transcription of the ATM locus. B. Enriched Gene Ontology (GO) pathways among disease genes with enriched expression in Purkinje cells and downregulated in A-T Purkinje cells. C. Overlap between HDAC4 neuronal target genes and downregulated DEGs in A-T cerebellum.

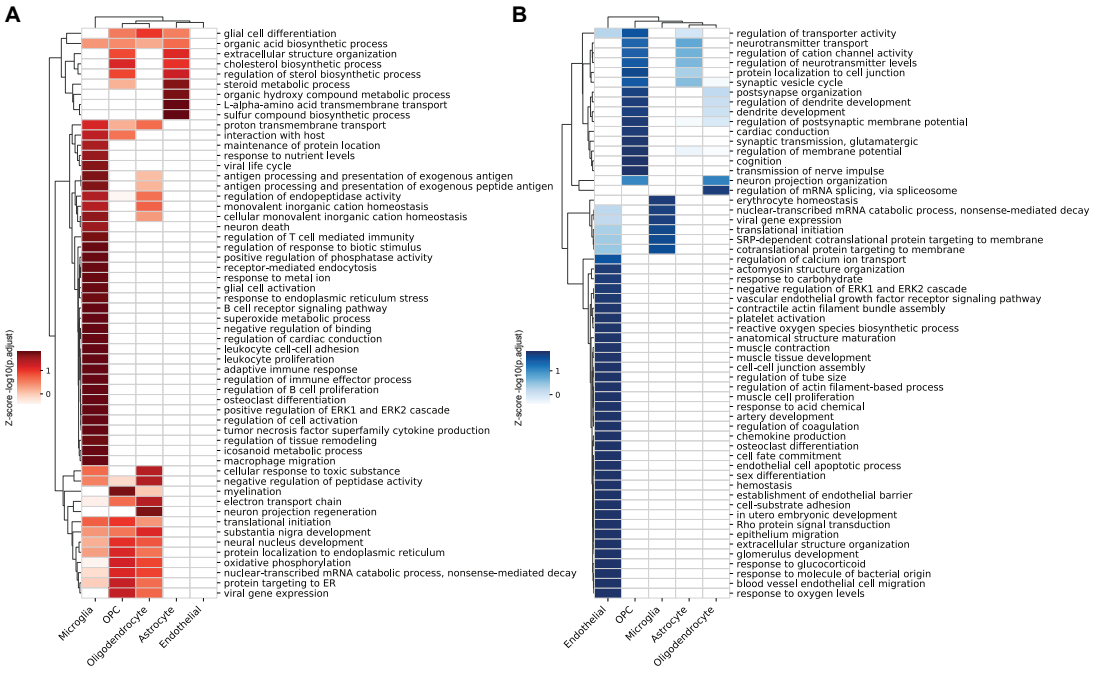


Figure A.7: Enrichment of pathways among genes with greater dysregulation in A-T cerebellum than A-T PFC. A. Heatmap showing log2 fold-change of hereditary ataxia genes in A-T cerebellum. *FDR<0.05. ATM expression increased in several cell types in A-T cerebellum, suggesting that lack of ATM function induces compensatory increases in transcription of the ATM locus. B. Enriched Gene Ontology (GO) pathways among disease genes with enriched expression in Purkinje cells and downregulated in A-T Purkinje cells. C. Overlap between HDAC4 neuronal target genes and downregulated DEGs in A-T cerebellum.

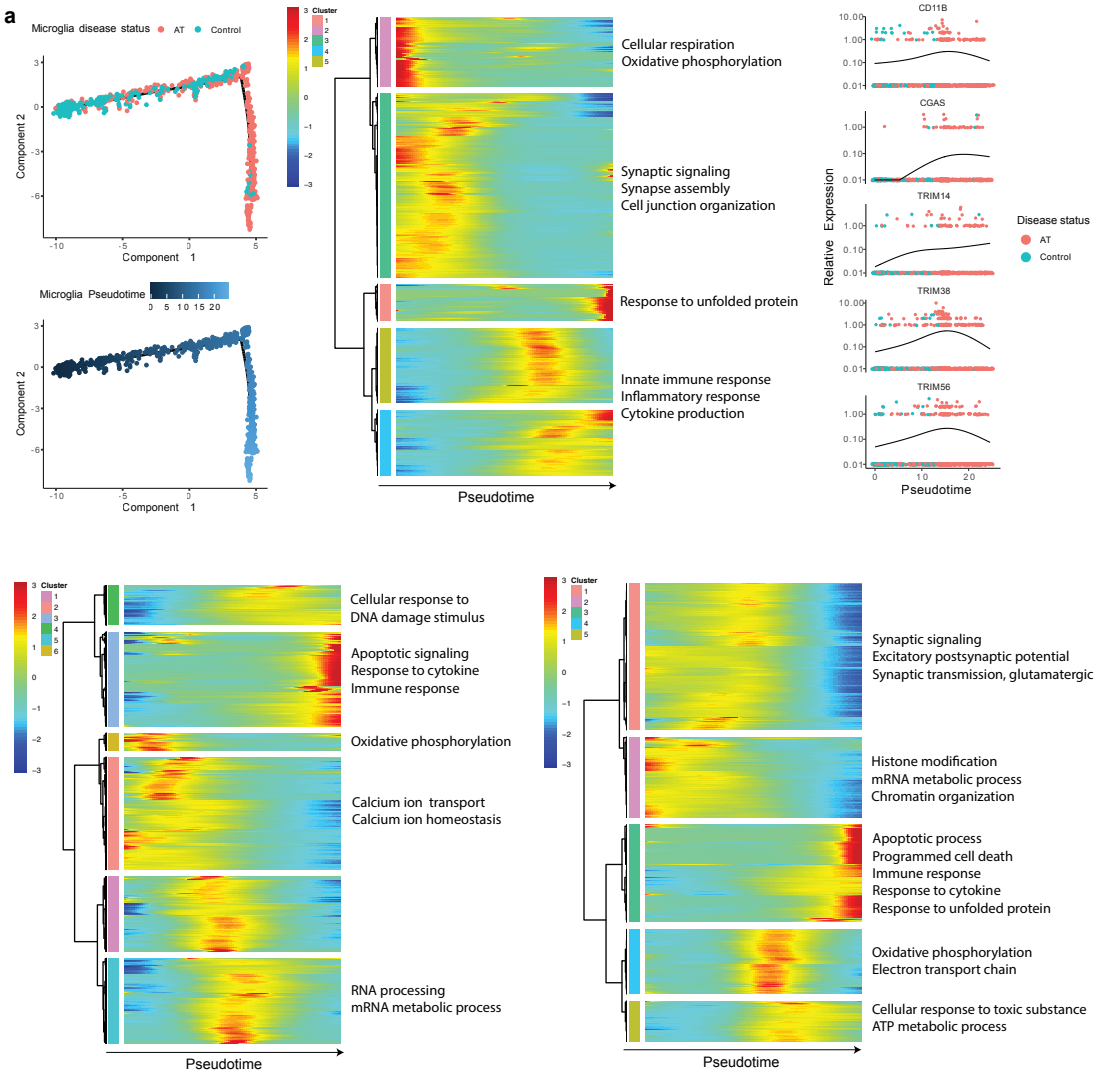


Figure A.8: Pseudotime analysis of disease progression reveals early microglia activation. A. Disease progression pseudotime trajectory of A-T cerebellar microglia, colored by disease status (red: A-T, blue: control), or pseudotime (healthy to diseased). B. Heatmap of genes that significantly change over pseudotime in microglia, clustered by pseudotemporal expression patterns. Each cluster is annotated with enriched GO terms (FDR<0.05). C. Expression of *CD11B*, *CGAS*, *TRIM14*, *TRIM38*, and *TRIM56* over pseudotime in microglia. D. Heatmap of genes that change over pseudotime in Purkinje neurons, clustered by pseudotemporal expression patterns. Each cluster is annotated with enriched GO terms (FDR<0.05). E. Heatmap of genes that change over pseudotime in granule neurons, clustered by pseudotemporal expression patterns. Each cluster is annotated with enriched GO terms (FDR<0.05). Heatmaps in B, D, E depict centered and scaled expression.

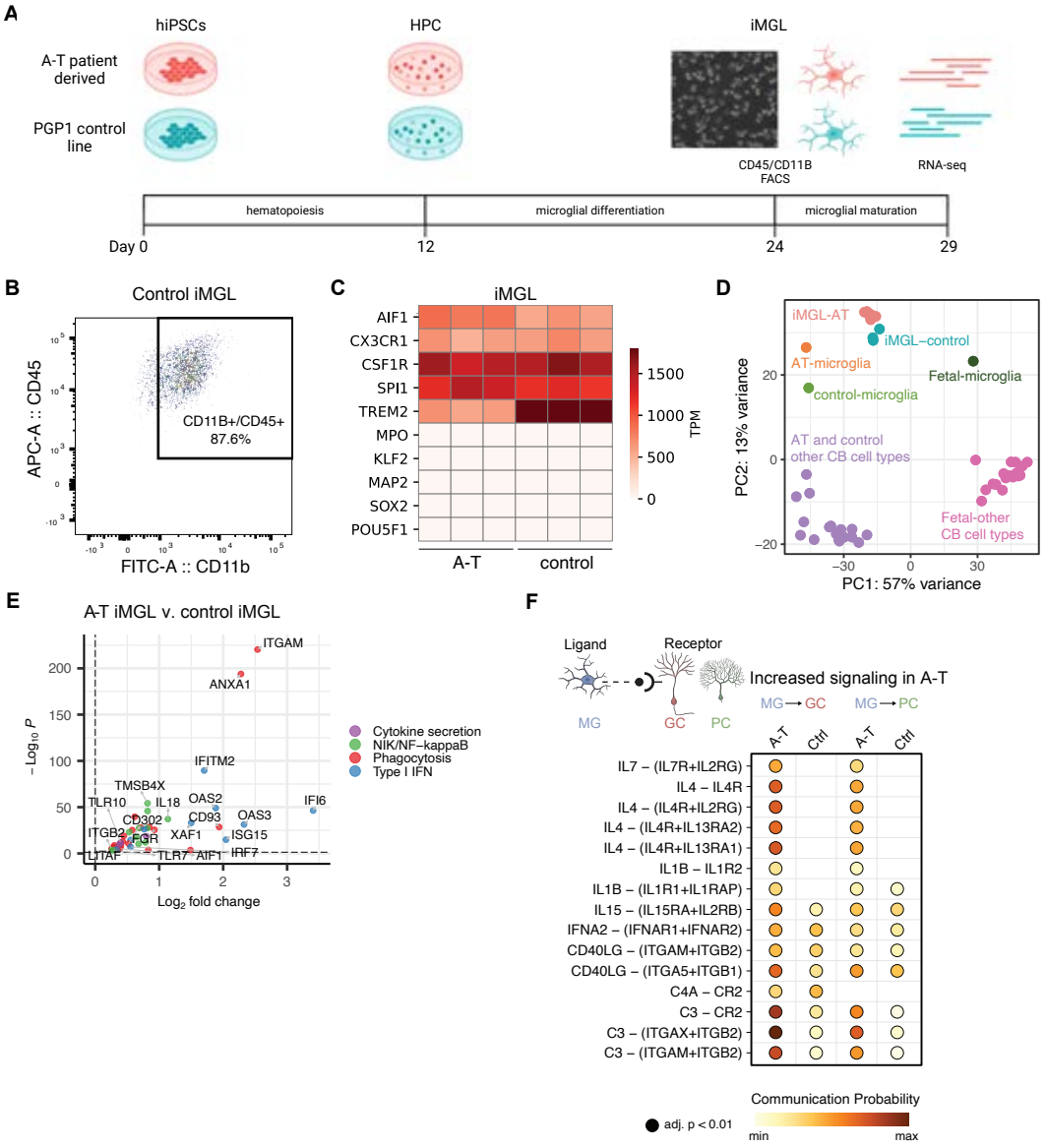


Figure A.9: A-T patient iPSC-derived microglia reveal cell-intrinsic activation of NF-kappaB and type I interferon pathways. A. Schematic for generation of iPSC-derived microglia (iMGL) from human A-T patient and control iPSCs. PGP1, Personal Genome Project 1; HPC: hematopoietic progenitor cells. B. Flow cytometry analysis of CD45 and CD11B co-expression in control iMGLs 24 days post-differentiation. C. Expression of microglia (*AIF1*, *CX3CR1*, *CSF1R*, *SPI1*, *TREM2*), myeloid lineage (*MPO*, *KLF2*), neuronal (*MAP2*, *SOX2*), and iPSC (*POU5F1*) marker genes in A-T and control iMGLs. Each column represents data from iMGL differentiated in an independent well. D. Principal component analysis plot of A-T and control iMGL (iMGL-AT/control) and adult cerebellar cell type pseudobulk transcriptomic profiles derived from snRNA-seq in this study (AT/control-microglia) and fetal cerebellar cell type pseudobulk transcriptomic profiles derived from snRNA-seq data in Aldinger et al., 2021 (Fetal-microglia).

Figure A.9 (Continued). E. Volcano plot for log2 fold change in gene expression in A-T versus control iMGL. Representative genes associated with cytokine secretion (purple), NIK/NF-kappaB (green), phagocytosis (red), and type I interferon (IFN) (blue) are highlighted. Horizontal dash line represents FDR=0.05. F. Ligand-receptor pairs with increased communication probability from microglia to granule or Purkinje neurons in A-T cerebellum versus control. Dot color represents communication probability (strength of signaling). Dot size represents significance of ligand-receptor pair and pairs with adjusted p-value <0.01 shown.

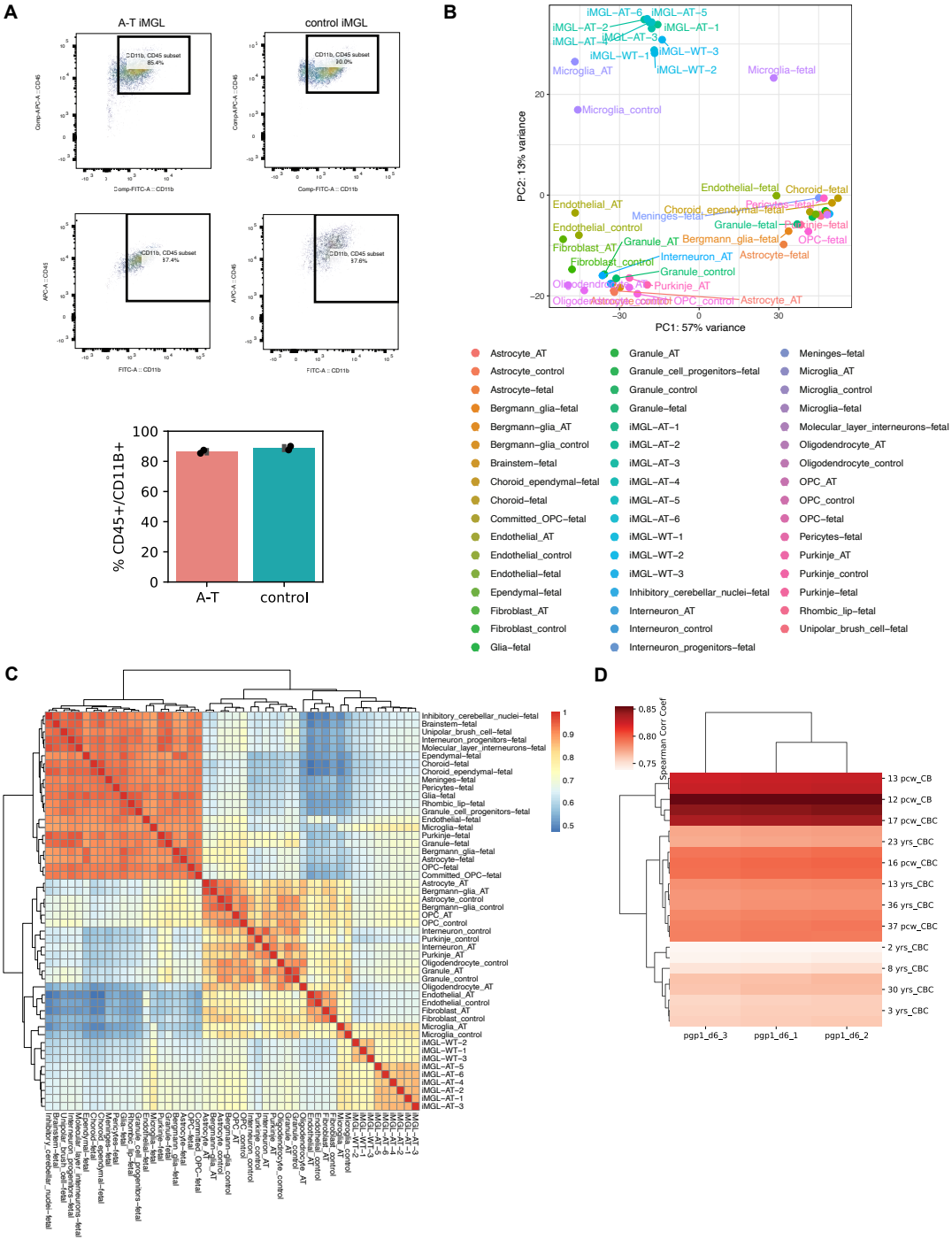


Figure A.10: Differentiation of A-T patient and control human iPSCs into microglia and neurons.

Figure A.10 (Continued). A. Flow cytometry analysis of CD45 and CD11B co-expression in A-T and control iMGLs 24 days post-differentiation from 2 independent differentiation batches. B. Principal component analysis plot of A-T and control iMGL and adult cerebellar cell type pseudobulk transcriptomic profiles derived from snRNA-seq in this study and fetal cerebellar cell type pseudobulk transcriptomic profiles derived from snRNA-seq data in Aldinger et al., 2021. C. Heatmap of Spearman Correlation Coefficients between human adult and fetal cerebellar cell type transcriptome profiles and iMGL transcriptome profiles. D. Heatmap of Spearman Correlation Coefficients between human control (PGP1) iPSC-derived neurons at day 6 of differentiation transcriptomes and human developing cerebellum (CB) and cerebellar cortex (CBC) transcriptomes from BrainSpan: Atlas of the Developing Human Brain. iMGL-AT, A-T iPSC-derived microglia. iMGL-WT, control iPSC-derived microglia.

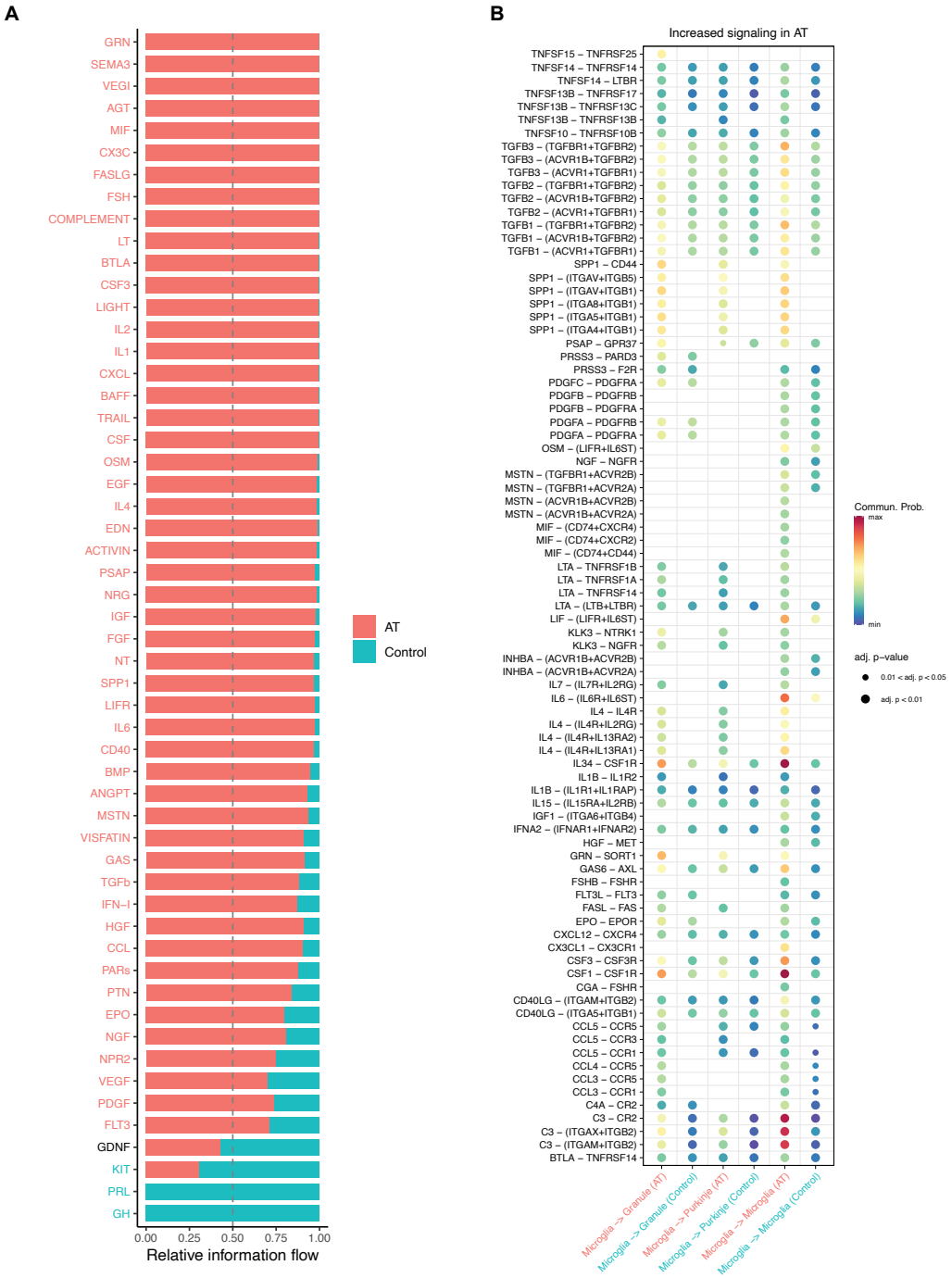


Figure A.11: Altered cell-cell communication in A-T cerebellum. A. Relative information flow for signaling pathways enriched in A-T or control cerebellum. Information flow for each pathway is calculated as the sum of the communication probability among all pairs of cell groups.

Figure A.11 (Continued). B. Ligand-receptor pairs with increased communication probability between microglia, granule, and Purkinje neurons in A-T cerebellum compared to control. Dot color represents communication probability (strength of signaling) and dot size represents significance of ligand-receptor pair (adjusted p-value). Empty space represents a communication probability of zero.

A.2 Chapter 2 Supplemental Figures

Chapter 2 contains the following supplemental figures: A.12 A.13 A.14 A.15 A.16 A.17

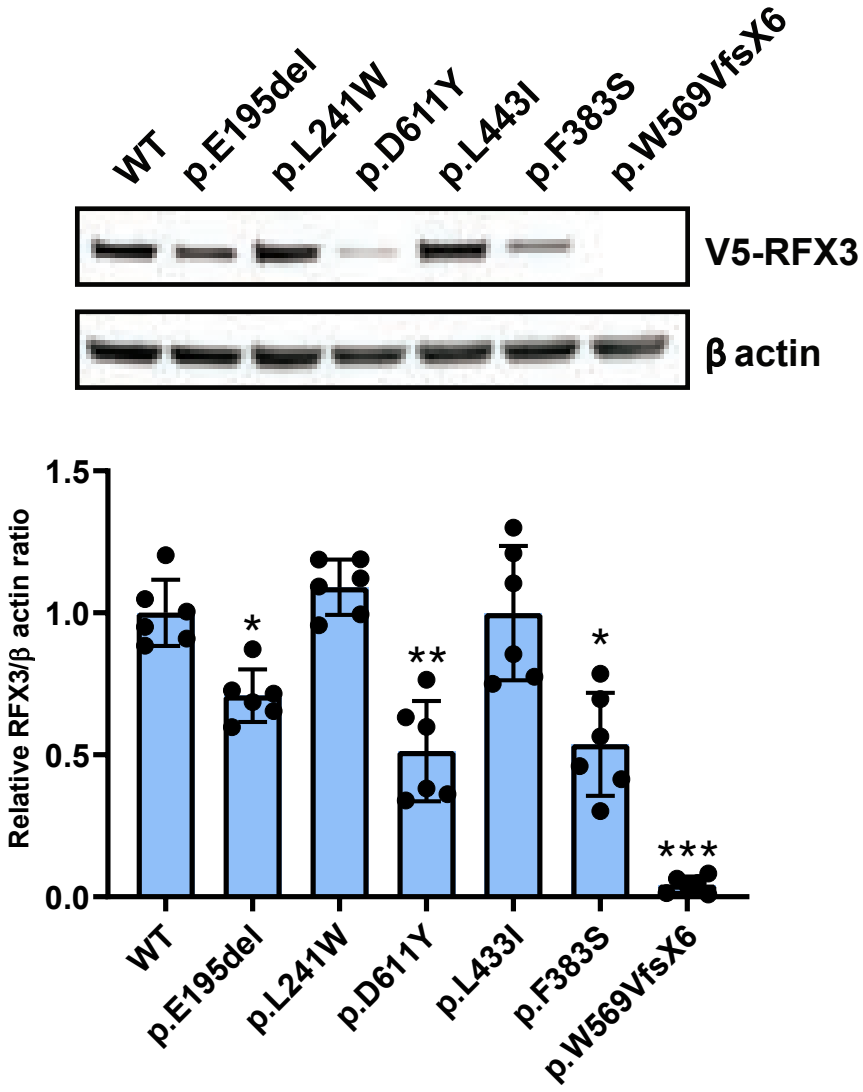


Figure A.12: Impact of observed variants on RFX3 protein expression. Immunoblot analysis of RFX3 in Hela cells expressing V5-tagged RFX3 variants. Protein levels were assessed by measuring anti-V5 immunoreactivity, normalized to beta actin immunoreactivity, and expressed in relation to cells expressing V5-tagged wild-type RFX3. Top panel: representative Western blot; bottom panel: quantitative analysis showing average of six experiments. Error bars represent standard deviation. Statistics: one-way ANOVA by repeated measures, and follow-up multiple comparisons test with correction via Dunnett test. ***p<0.001, **p<0.01, *p<0.05.

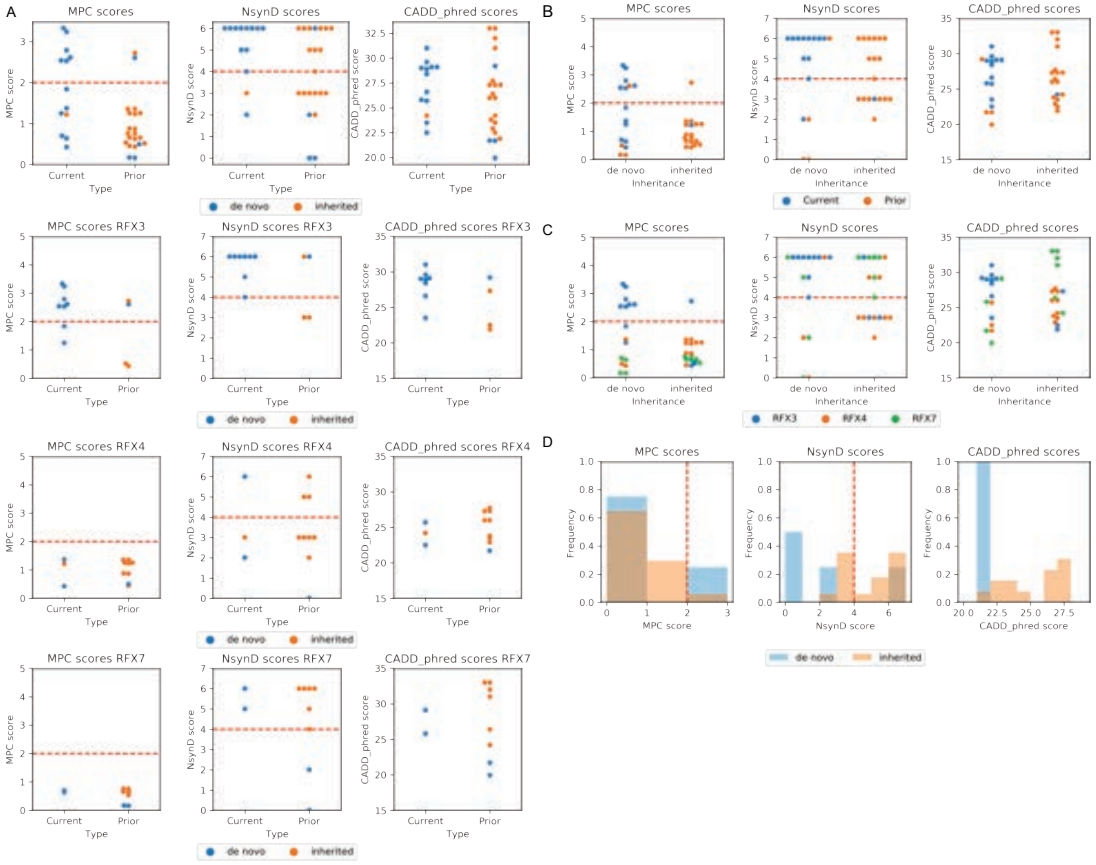


Figure A.13: Deleteriousness scores of missense variants in RFX3, RFX4, and RFX7. A. The RFX3 missense variants reported in this study have a higher frequency of damaging predictions than prior reported missense variants. B. De novo missense variants are more likely damaging than inherited missense variants. C. RFX3 missense variants are more likely damaging than RFX4 or 7 missense variants. D. Prior reported missense variants in RFX3, 4, and 7 do not show clear associations between predicted deleteriousness and inheritance. MPC, Missense badness, PolyPhen-2, and Constraint. NsynD, Nonsynonymous Damaging score. CADD, Combined Annotation Dependent Depletion.

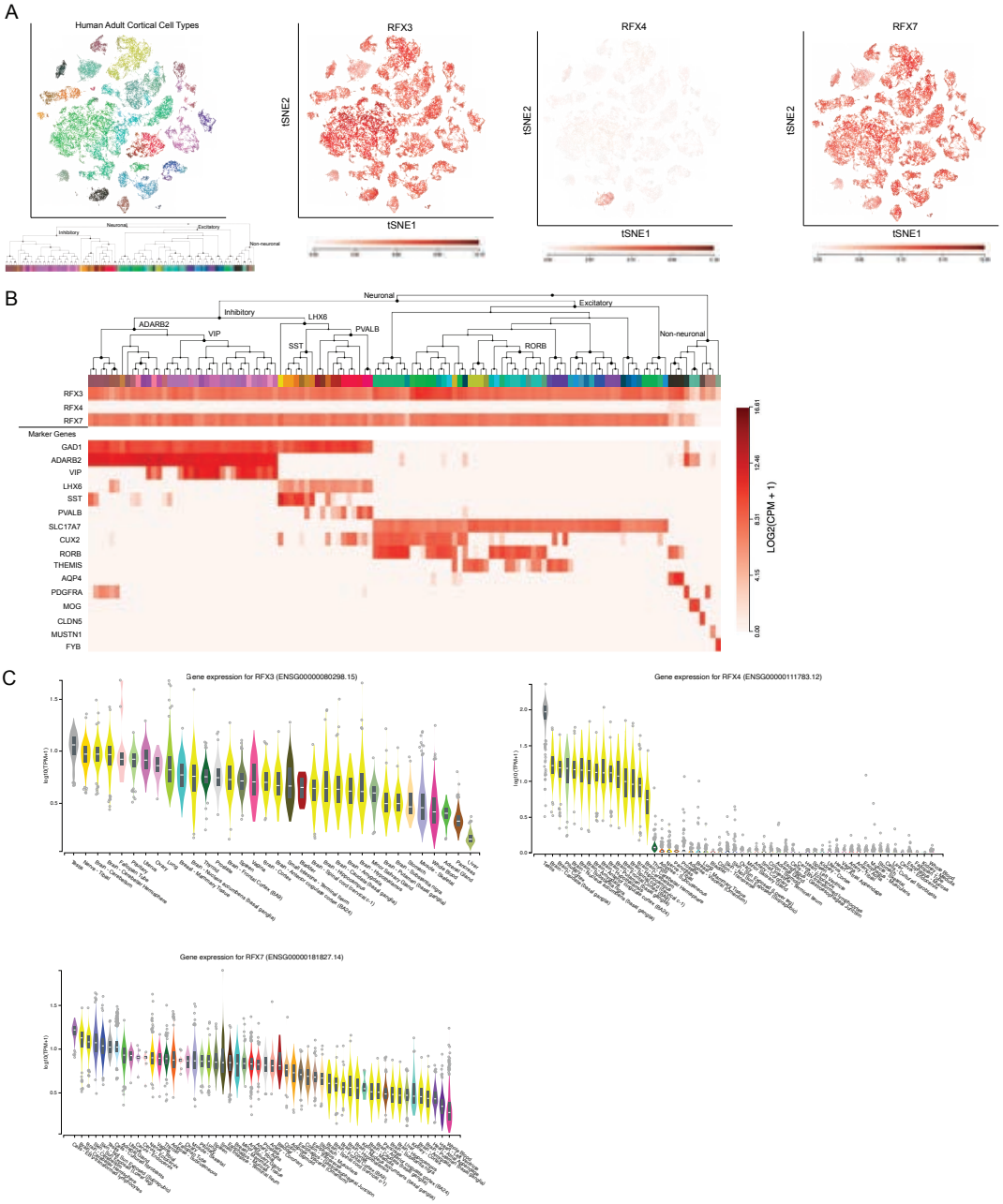


Figure A.14: RFX3, RFX4, and RFX7 expression patterns in human cortex. A. RFX3, 4, and 7 expression patterns in single cells of adult human cortex (Allen Human Brain Atlas). Cell types identified in human cortex shown in t-SNE plot with taxonomy of clusters. Expression color-scale units are Log2(CPM+1). Data from Allen Brain Map single-nucleus RNA-sequencing of human cortex (Image credit: Allen Institute). B. Heatmap visualizing RFX3, 4, and 7 expression among adult human cortical cell types, along with canonical cell type markers. (Image credit: Allen Institute). C. RFX3, RFX4, and RFX7 expression across all human tissues. (Data Source: GTEx Analysis Release V8 dbGaP Accession phs000424.v8.p2).

A

B

RFX3 Motif Occurrences

Gene Symbol	Start	Stop	p-value	q-value	Matched Sequence
AP2S1	1223	1238	2.39E-08	0.00318	GGTCGCCATGGTAACG
KDM6B	854	869	2.43E-08	0.00318	GGTTGCCATGGAAACA
ANK2	941	956	2.08E-07	0.0174	TGCTGCCATGGCAACG
MYT1L	720	735	7.63E-07	0.0361	GGTTGCCATGGCCACC
NONO	693	708	8.63E-07	0.0361	GGTTACCATGGAAATG

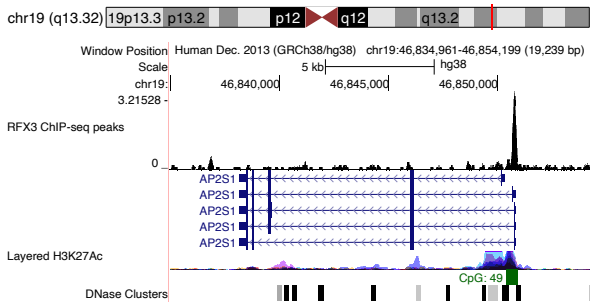
C

D

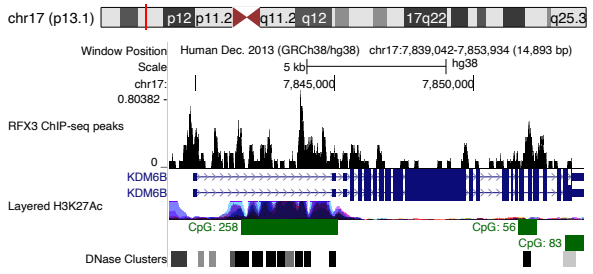
RFX4 Motif Occurrences

Gene Symbol	Start	Stop	p-value	q-value	Matched Sequence
AP2S1	1223	1238	2.94E-08	0.00736	GGTCGCCATGGTAACG
KDM6B	854	869	1.51E-07	0.0158	GGTTGCCATGGAAACA
ANK2	941	956	5.16E-07	0.0431	TGCTGCCATGGCAACG
MYT1L	720	735	1.10E-06	0.0646	GGTTGCCATGGCCACC

E



F



G

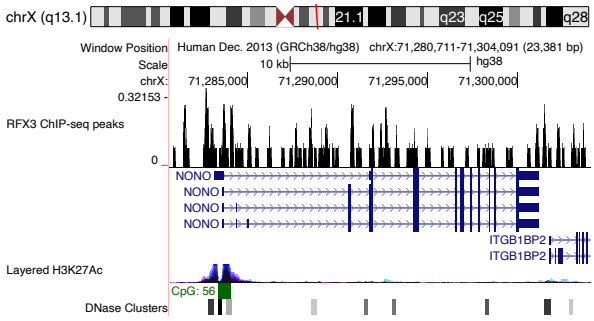


Figure A.15: Presence of RFX3 and RFX4 binding motifs in ASD-associated gene promoters. A. RFX3 binding motif MA0798.1 (JASPAR 2020). B. ASD risk genes with significant RFX3 binding motif occurrences. C. RFX4 binding motif MA0799.1 (JASPAR 2020). D. ASD risk genes with significant RFX4 binding motif occurrences. E – G. RFX3 ChIP-seq binding peaks are located in promoter region of ASD-associated genes (E) AP2S1, (F) KDM6B, (G) NONO. All motif occurrences were identified using FIMO, q-values <0.10. RFX3 in HepG2 ChIP- seq binding peaks obtained from ENCODE GSM2534235.

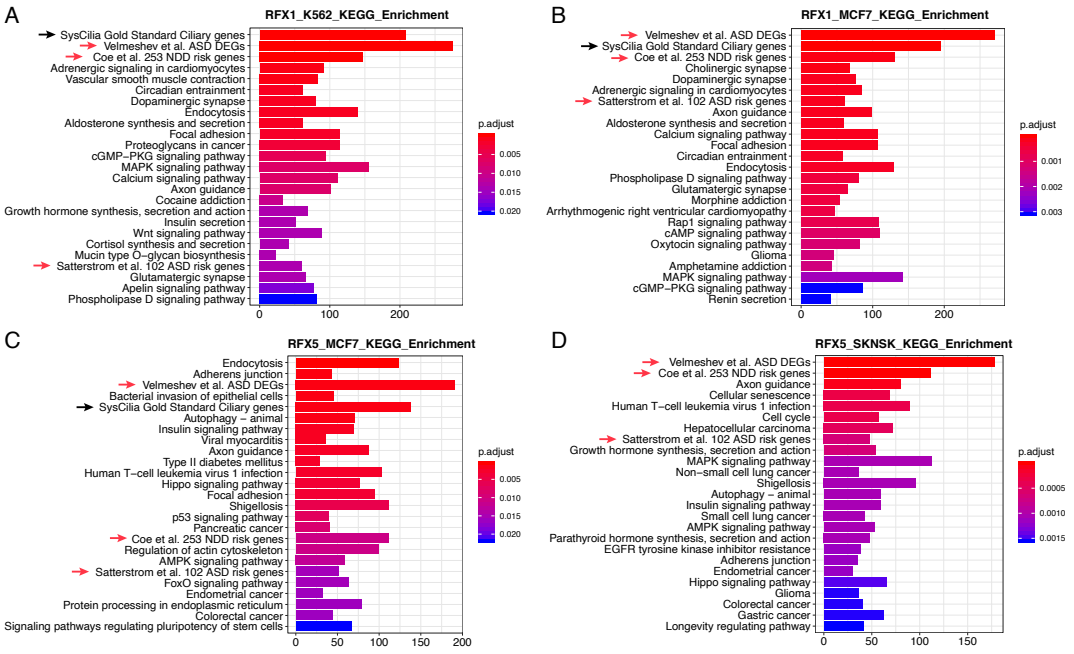


Figure A.16: Customized KEGG enrichment analysis of RFX functional binding sites. The enrichment of KEGG pathways, ciliary genes, ASD risk gene sets, and ASD differentially expressed genes (DEGs) for different RFX binding profiles. (A–D) representative barplots of the enriched customized KEGG pathways were shown using RFX1 (A-B), and RFX5 (C-D) ChIP- seq data. Pathways and ASD gene sets are ranked by their statistical significance (p.adjust values, Benjamini-Hochberg's correction). Red arrows indicate ASD risk gene sets and ASD DEGs. Black arrows indicate the gold standard ciliary genes. X-axis shows the number of genes bound by RFX in their promoter regions.

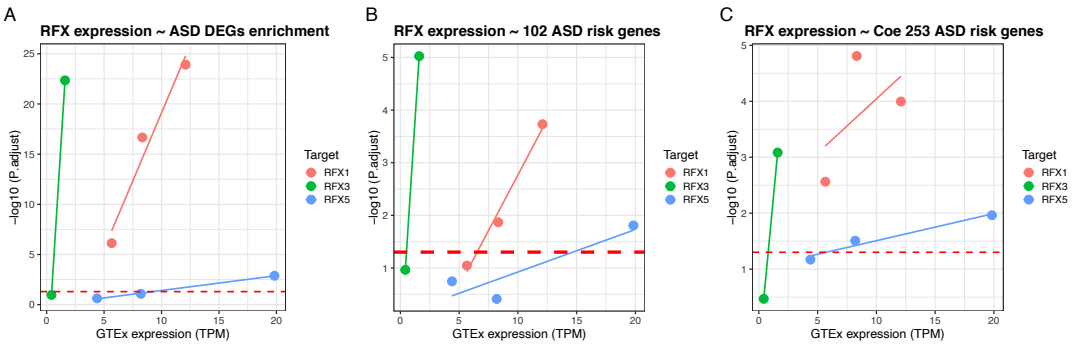


Figure A.17: Correlation of enrichment significance in ASD gene sets with RFX expression level. RFX expression levels in different cell lines were correlated with the enrichment significances [− log10(p.adjust)] for (A) Velmeshev et al. ASD differentially expressed genes (DEGs); (B) Satterstrom et al. 102 ASD risk genes; (C) Coe et al. 253 NDD risk genes. Red dotted line indicates the P-value = 0.05. X-axis shows the GTEx expression of RFX in different cell lines derived from a variety of human tissues.

A.3 Chapter 3 Supplemental Figures

Chapter 3 contains the following supplemental figures: A.18 A.19 A.20 A.21 A.22 A.23

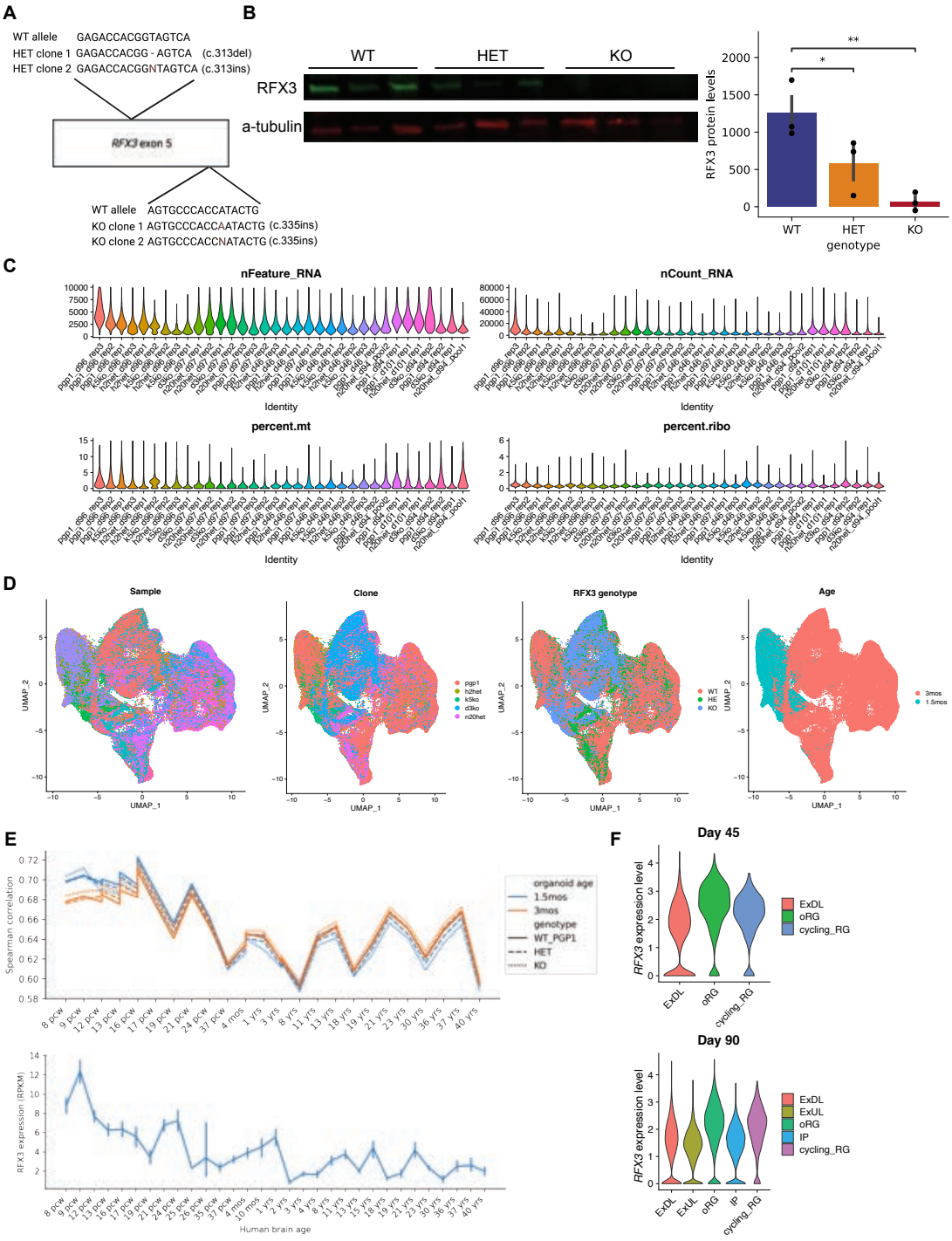


Figure A.18: scRNA-seq of RFX3 WT, HET, and KO dorsal forebrain organoids.

Figure A.18 (Continued). A. Generation of RFX3 isogenic lines (RFX3 WT, HET, and KO) using CRISPR-Cas9. The schematic shows the indels generated in RFX3 exon 5. n=2 heterozygous clones, n=2 homozygous clones. Coding nomenclature is for transcript NM_001282116.2. B. Western blot of RFX3 and alpha-tubulin protein levels in RFX3 WT, HET, and KO iPSCs. C. Violin plot of quality control metrics per sample post-filtering. D. UMAP visualization of meta data. E. Organoid transcriptomes correlate with mid-gestation fetal brain. Spearman correlation coefficient of organoid bulk transcriptome and human brain transcriptome across development (primary motor cortex, BrainSpan) (top). RFX3 expression in RPKM across human brain development (primary motor cortex, BrainSpan) (bottom). F. RFX3 expression across cell types in day 45 and day 90 organoids.

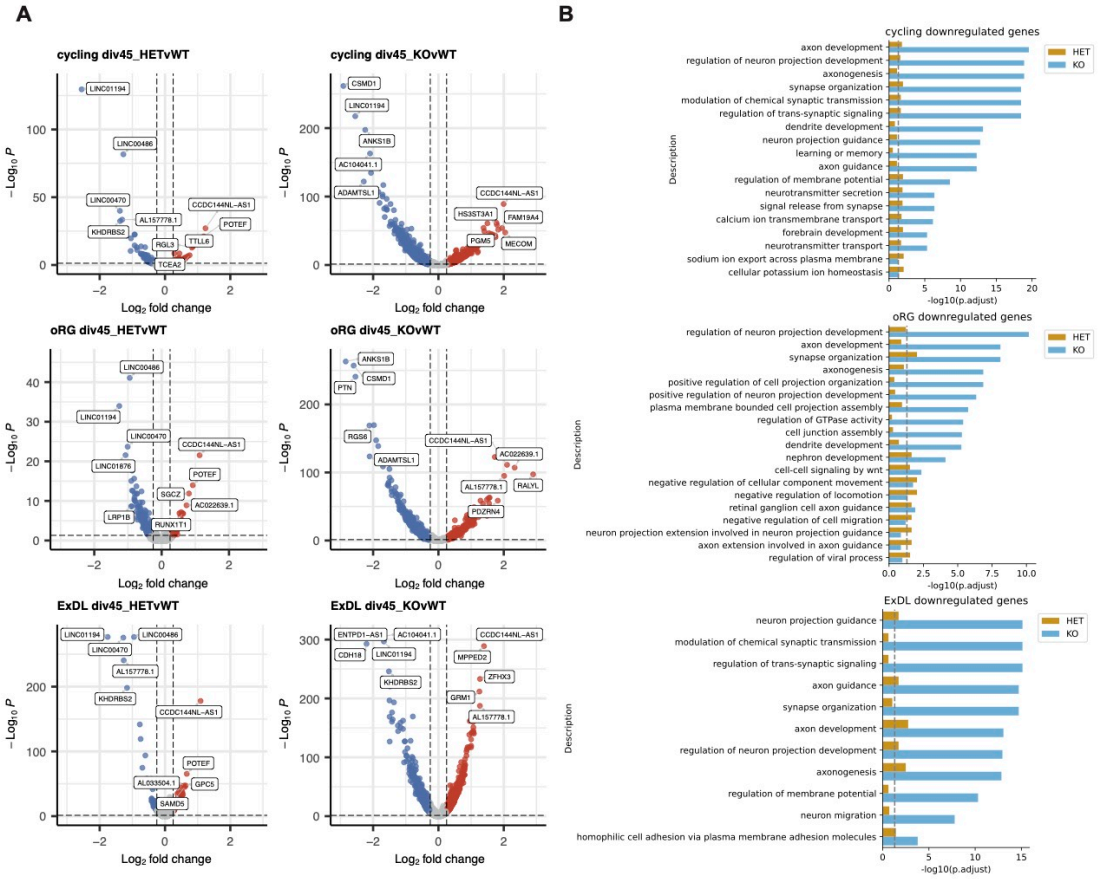


Figure A.19: Cell type specific gene expression dysregulation in day 45 RFX3 deficient dorsal forebrain organoids. A. Volcano plot showing cell type specific differentially expressed genes (DEGs) between RFX3 HET and WT organoids and RFX3 KO and WT organoids. Genes significantly downregulated are in blue (log2FC<-0.25, FDR<0.05). Genes significantly upregulated are in red (log2FC>0.25, FDR<0.05). The top DEGs ranked by significance are labeled. B. Gene Ontology (GO) term enrichment analysis of significantly downregulated genes in RFX3 HET and KO organoids.

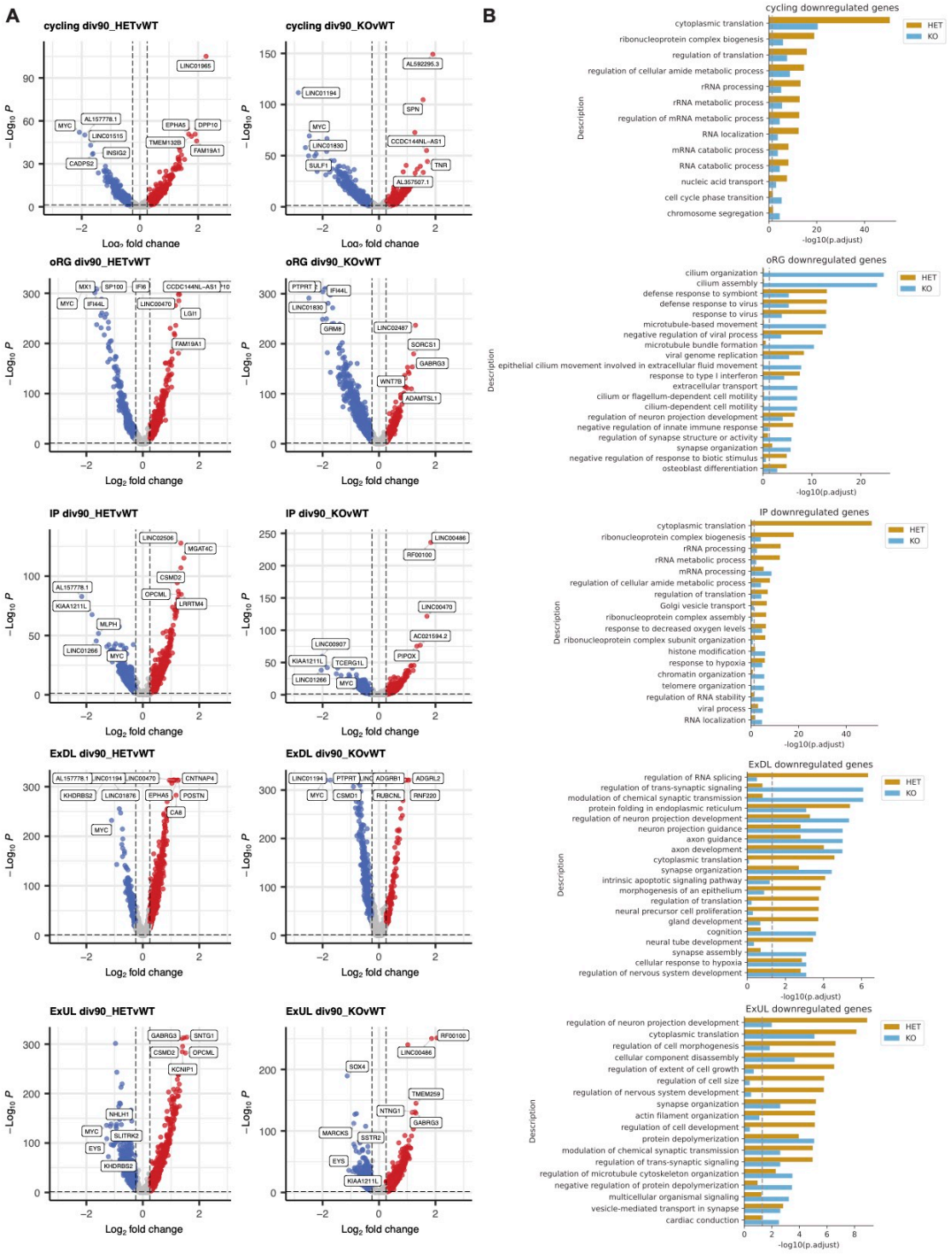


Figure A.20: Cell type specific gene expression dysregulation in day 90 RFX3 deficient dorsal forebrain organoids. A. Volcano plot showing cell type specific differentially expressed genes (DEGs) between RFX3 HET and WT organoids and RFX3 KO and WT organoids.

Figure A.20 (Continued). Genes significantly downregulated are in blue (log2FC<-0.25, FDR<0.05). Genes significantly upregulated are in red (log2FC>0.25, FDR<0.05). The top DEGs ranked by significance are labeled. B. Gene Ontology (GO) term enrichment analysis of significantly downregulated genes in RFX3 HET and KO organoids.

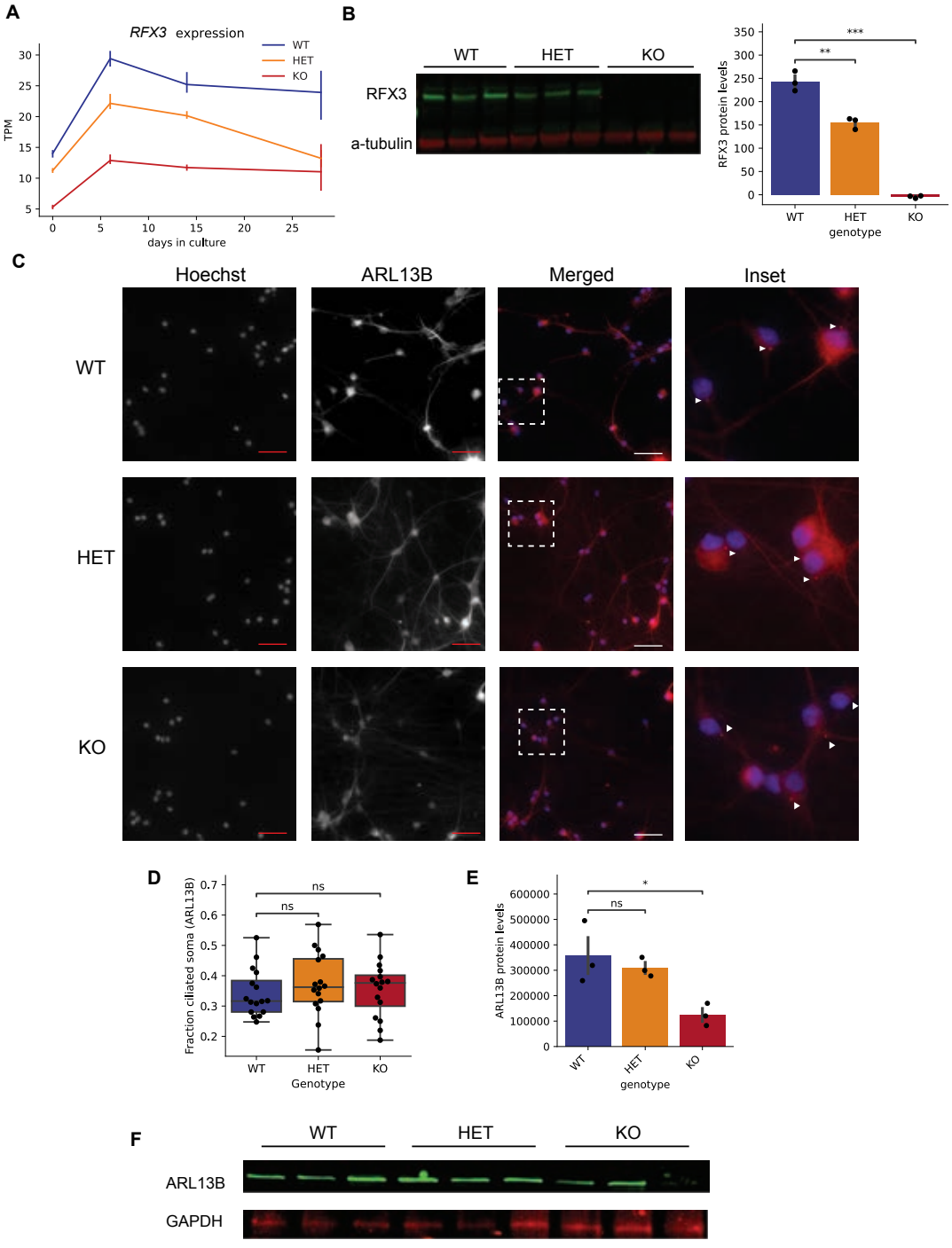


Figure A.21: Modeling RFX3 deficiency in Ngn2 neurons. A. RFX3 expression across Ngn2 iPSC-derived neuron differentiation days 0, 6, 14, and 28 measured by RNA-sequencing. TPM, transcripts per million reads.

Figure A.21 (Continued). B. Representative Western blot of RFX3 in day 14 Ngn2 neurons (left). Quantification of Western blot (right). n=3 independent differentiations per condition. **p< 0.01, ***p<0.005. one-way ANOVA with t-test with Bonferroni correction. C. Representative images of day 20 Ngn2 neurons stained with Hoechst (nuclei), or ARL13B (cilia). White arrows indicate ARL13B+ cilia. White dashed boxes indicate the inset region. Scale bar=100 um. D. Quantification of fraction of soma with ARL13B+ cilia in day 20 Ngn2 neurons. n=16 independent wells per condition. ns, not significant, one-way ANOVA with t-test with Bonferroni correction. E. Quantification of ARL13B Western blot. n=3 independent neuron differentiations per condition. Data represent mean +/- SEM. *p <0.05, ns not significant, one-way ANOVA with t-test with Bonferroni correction. F. Representative Western blot of ARL13B and GAPDH in day 14 Ngn2 neurons. n=3 independent differentiations per condition.

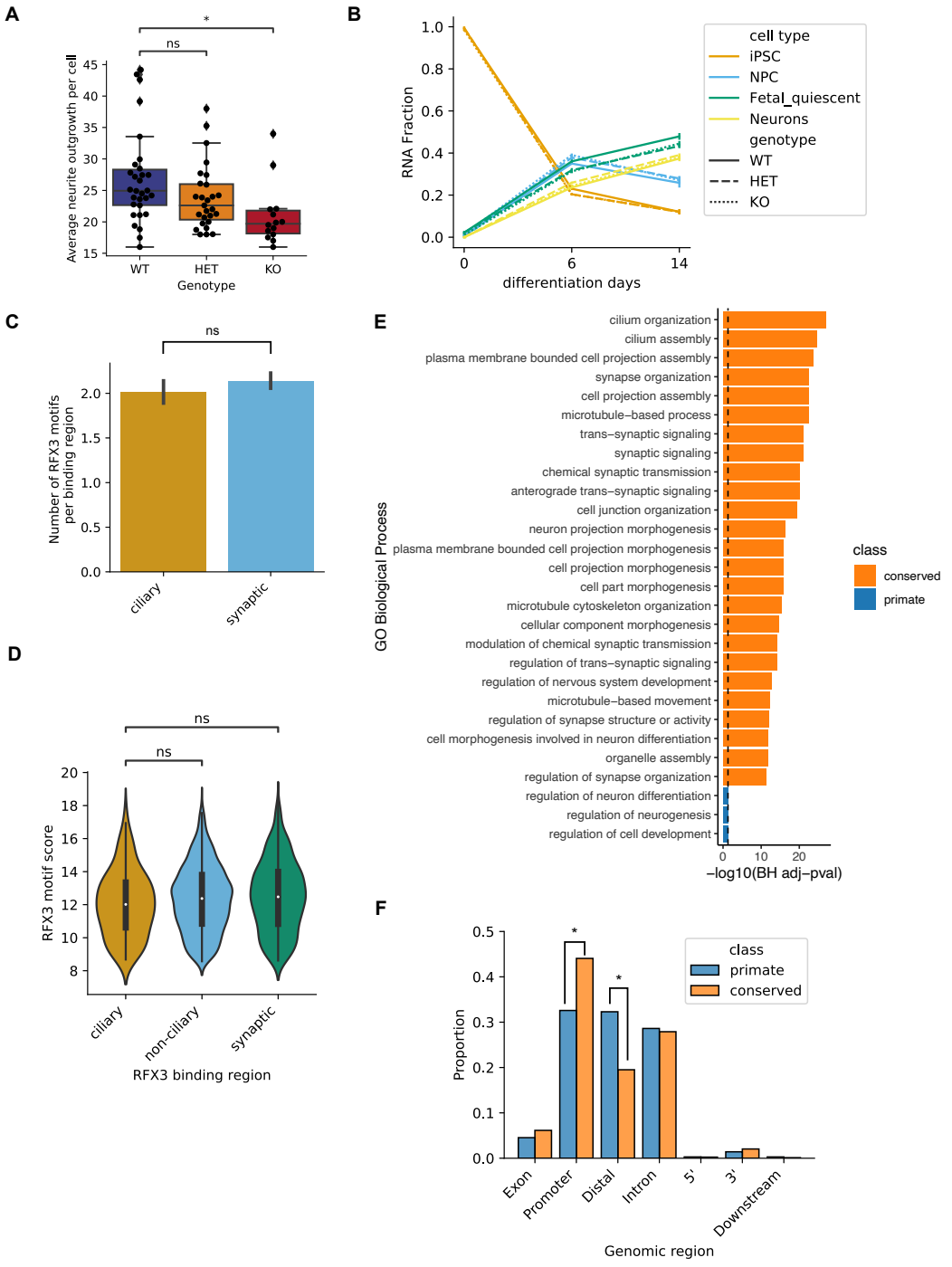


Figure A.22: Characteristics of RFX3 binding motifs in RFX3 binding sites. A. Average neurite length (um) per cell marked by TUJ1 in day 14 cultures. p <0.05, ns not significant, one-way ANOVA with t-test with Bonferroni correction.

Figure A.22 (Continued). B. RNA fraction of cell types estimated by deconvolution of bulk RNA-seq data from iPSC-derived Ngn2 neurons at days 0, 6, and 14. C. Average number of RFX3 motifs per binding region associated with ciliary genes or synaptic genes. ns, not significant, t-test. Data represent mean +/- SEM. D. Average RFX3 motif score in binding regions associated with ciliary genes, non-ciliary genes, or synaptic genes. ns, not significant, t-test. E. Gene Ontology (GO) enrichment analysis of genes associated with a primate-specific or conserved RFX3 motif. Dashed line indicates significance threshold FDR 0.05. F. Proportion of primate-specific or conserved RFX3 motifs within each genomic region. *p<0.05, proportions z-test with Bonferroni correction.

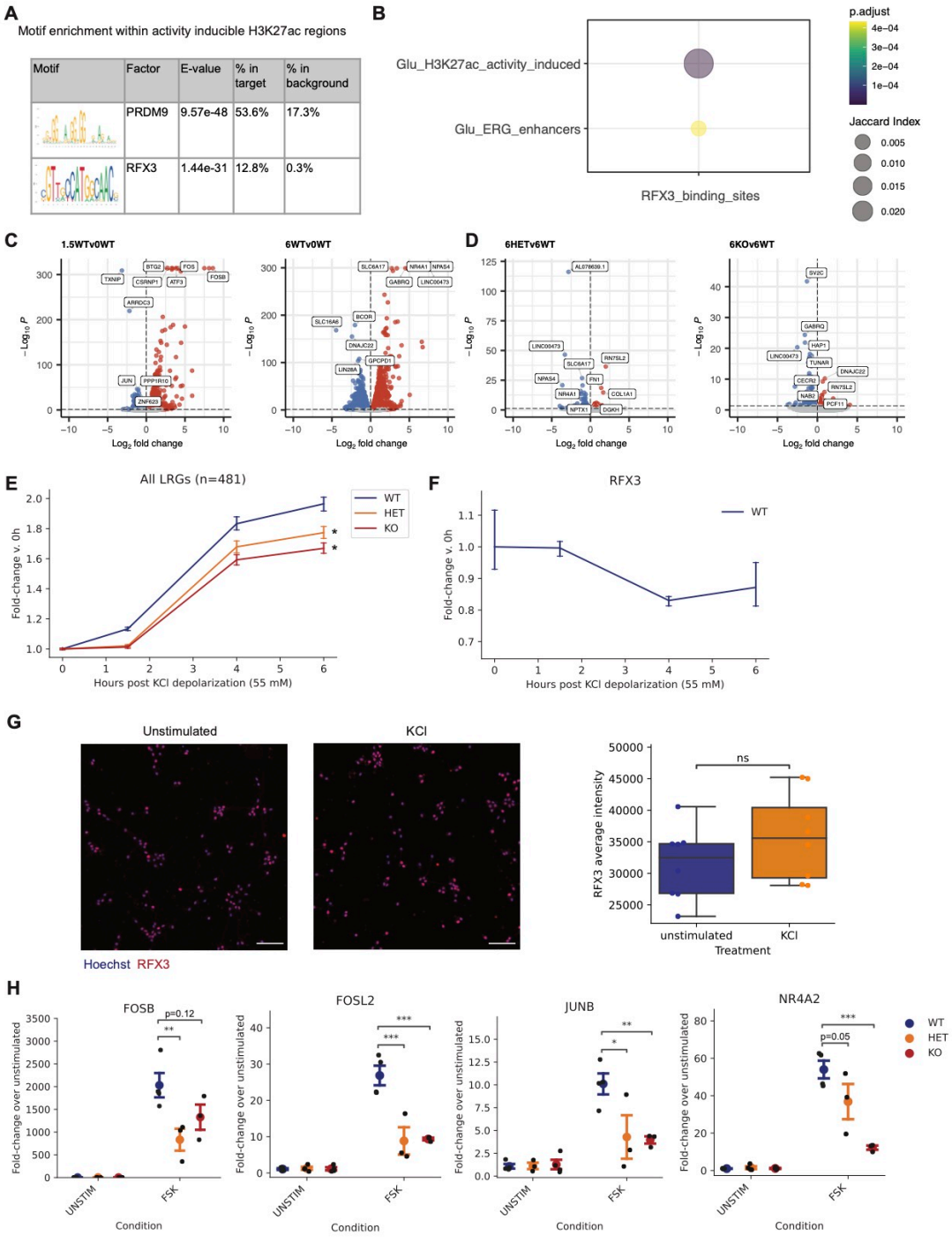


Figure A.23: RFX3 modulates CREB-dependent activity-dependent transcriptional responses. A. Top two binding motifs enriched within activity inducible H3K27ac regions in Ngn2 neurons.

Figure A.23 (Continued). B. Jaccard index of RFX3 binding sites and all activity inducible H3K27ac regions or those associated with early response genes (ERG) in glutamatergic (Glu) Ngn2 neurons. Dot size indicates Jaccard index. Color indicates significance of overlap, Fisher's exact test, two-tail adjusted p-value. C. Volcano plot of differentially expressed genes in 1.5 hour or 6 hour KCl stimulated WT v. unstimulated WT neurons at day 14 in culture. D. Volcano plot of differentially expressed genes in 6 hour KCl stimulated HET v. WT or KO v. WT neurons at day 14 in culture. Genes significantly downregulated are in blue (FDR<0.05). (C-D) Genes significantly upregulated are in red (FDR<0.05). Top 10 DEGs are labeled. E. Fold-induction of all late response genes at 1.5, 4, and 6 hours post KCl depolarization compared to unstimulated neurons. Data represented as mean and 95% CI, *p<0.05 t-test, n=3 independent neuron differentiations per genotype. F. RFX3 expression at 1.5, 4, and 6 hours post KCl depolarization compared to unstimulated WT neurons. Data represented as mean and 95% CI, n=3 independent neuron differentiations. G. Representative images and quantification of RFX3 in unstimulated and KCl stimulated (15 minutes) day 14 WT neurons. n=8 independent wells per condition, ns, not significant, t-test. Scale bar=100 um. H. Fold-induction of FOSB, FOSL2, JUNB, NR4A2 at 1.5 hours post FSK (10 uM) or H89 (10 uM) pretreatment and FSK (10 uM) assessed by RT-qPCR. Data represented as mean +/- SEM, n=3-4 independent wells per condition. *p<0.05, **p<0.01, ***p<0.005, ns not significant, one-way ANOVA with t-test with Bonferroni correction.

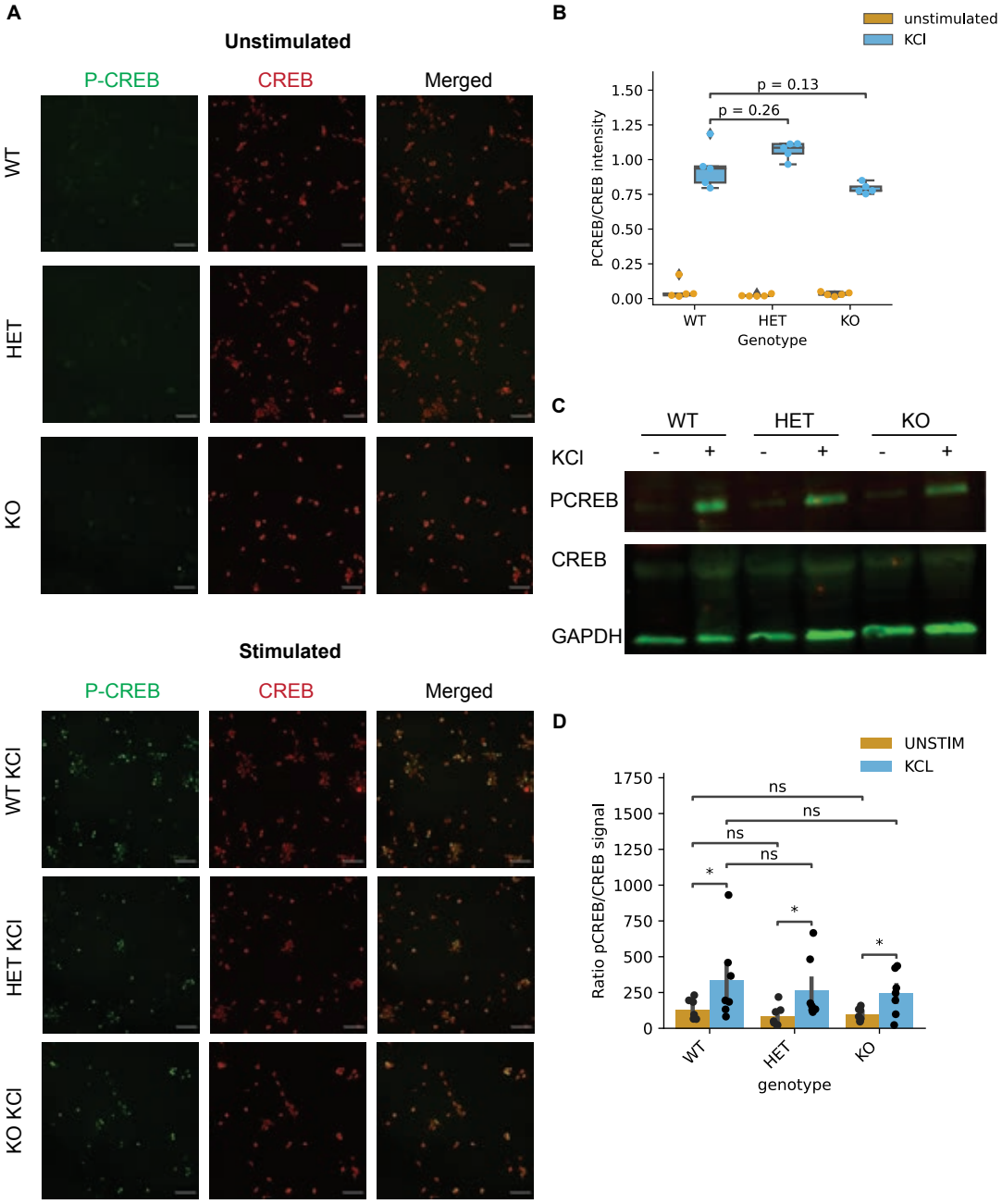


Figure A.24: Depolarization induced activation of CREB is intact in RFX3 deficient neurons. A. Representative images of PCREB and CREB in unstimulated and stimulated (1.5 hours KCl, 55 mM) WT, HET, and KO day 14 neurons. Scale bar=100 um.

Figure A.24 (Continued). B. Quantification of nuclear PCREB intensity normalized by CREB intensity in unstimulated and stimulated (1.5 hours KCl, 55 mM) WT, HET, and KO day 14 neurons. P-values correspond to one-way ANOVA with t-test with Bonferroni correction. C. Representative Western blot of PCREB, CREB, and GAPDH in unstimulated and stimulated (15 minutes KCl, 55 mM) WT, HET, and KO day 14 neurons. D. Quantification of Western blot from (C) as the ratio of PCREB/CREB signal. Data represent mean +/- SEM, n=7 replicates per condition. *p<0.05, ns, not significant, one-way ANOVA with t-test with Bonferroni correction.

References

[1] Edsel M. Abud, Ricardo N. Ramirez, Eric S. Martinez, Luke M. Healy, Cecilia H. H. Nguyen, Sean A. Newman, Andriy V. Yeromin, Vanessa M. Scarfone, Samuel E. Marsh, Cristhian Fimbres, Chad A. Caraway, Gianna M. Fote, Abdullah M. Madany, Anshu Agrawal, Rakez Kayed, Karen H. Gylys, Michael D. Cahalan, Brian J. Cummings, Jack P. Antel, Ali Mortazavi, Monica J. Carson, Wayne W. Poon, and Mathew Blurton-Jones. iPSC-Derived Human Microglia-like Cells to Study Neurological Diseases. *Neuron*, 94(2):278–293.e9, April 2017. ISSN 1097-4199. doi: 10.1016/j.neuron.2017.03.042.

[2] Geetika Aggarwal, Subhashis Banerjee, Spencer A. Jones, Yousri Benchaar, Jasmine Bélanger, Myriam Sévigny, Denise M. Smith, Michael L. Niehoff, Monica Pavlack, Ian Mitchelle S. de Vera, Terri L. Petkau, Blair R. Leavitt, Karen Ling, Paymaan Jafar-Nejad, Frank Rigo, John E. Morley, Susan A. Farr, Paul A. Dutchak, Chantelle F. Sephton, and Andrew D. Nguyen. Antisense oligonucleotides targeting the miR-29b binding site in the GRN mRNA increase progranulin translation, September 2022. URL https://www.biorxiv.org/content/10.1101/2022.01.12.476053v3. Pages: 2022.01.12.476053 Section: New Results.

[3] Aouatef Ait-Lounis, Claire Bonal, Queralt Seguín-Estévez, Christoph D. Schmid, Philipp Bucher, Pedro L. Herrera, Bénédicte Durand, Paolo Meda, and Walter Reith. The Transcription Factor Rfx3 Regulates β-Cell Differentiation, Function, and Glucokinase Expression. *Diabetes*, 59(7):1674–1685, April 2010. ISSN 0012-1797. doi: 10.2337/db09-0986. URL https://doi.org/10.2337/db09-0986.

[4] Kimberly A. Aldinger, Zachary Thomson, Ian G. Phelps, Parthiv Haldipur, Mei Deng, Andrew E. Timms, Matthew Hirano, Gabriel Santpere, Charles Roco, Alexander B. Rosenberg, Belen Lorente-Galdos, Forrest O. Gulden, Diana O'Day, Lynne M. Overman, Steven N. Lisgo, Paula Alexandre, Nenad Sestan, Dan Doherty, William B. Dobyns, Georg Seelig, Ian A. Glass, and Kathleen J. Millen. Spatial and cell type transcriptional landscape of human cerebellar development. *Nature Neuroscience*, 24(8):1163–1175, August 2021. ISSN 1546-1726. doi: 10.1038/s41593-021-00872-y. URL https://www.nature.com/articles/s41593-021-00872-y. Number: 8 Publisher: Nature Publishing Group.

[5] Sandra Almeida, Zhijun Zhang, Giovanni Coppola, Wenjie Mao, Kensuke Futai, Anna Karydas, Michael D. Geschwind, M. Carmela Tartaglia, Fuying Gao, Davide Gianni, Miguel

Sena-Esteves, Daniel H. Geschwind, Bruce L. Miller, Robert V. Farese, and Fen-Biao Gao. Induced pluripotent stem cell models of progranulin-deficient frontotemporal dementia uncover specific reversible neuronal defects. *Cell Reports*, 2(4):789–798, October 2012. ISSN 2211-1247. doi: 10.1016/j.celrep.2012.09.007.

[6] Debha N. Amatya, Sara B. Linker, Ana P.D. Mendes, Renata Santos, Galina Erikson, Maxim N. Shokhirev, Yuansheng Zhou, Tatyana Sharpee, Fred H. Gage, Maria C. Marchetto, and Yeni Kim. Dynamical Electrical Complexity Is Reduced during Neuronal Differentiation in Autism Spectrum Disorder. *Stem Cell Reports*, 13(3):474–484, August 2019. ISSN 2213-6711. doi: 10.1016/j.stemcr.2019.08.001. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6739708/.

[7] Jelle van den Ameele, Luca Tiberi, Pierre Vanderhaeghen, and Ira Espuny-Camacho. Thinking out of the dish: what to learn about cortical development using pluripotent stem cells. *Trends in Neurosciences*, 37(6):334–342, June 2014. ISSN 0166-2236, 1878-108X. doi: 10.1016/j.tins.2014.03.005. URL https://www.cell.com/trends/neurosciences/abstract/S0166-2236(14)00044-7. Publisher: Elsevier.

[8] R. E. Amir, I. B. Van den Veyver, M. Wan, C. Q. Tran, U. Francke, and H. Y. Zoghbi. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nature Genetics*, 23(2):185–188, October 1999. ISSN 1061-4036. doi: 10.1038/13810.

[9] Nickesha C. Anderson, Pin-Fang Chen, Kesavan Meganathan, Wardiya Afshar Saber, Andrew J. Petersen, Anita Bhattacharyya, Kristen L. Kroll, and Mustafa Sahin. Balancing serendipity and reproducibility: Pluripotent stem cells as experimental systems for intellectual and developmental disorders. *Stem Cell Reports*, 16(6):1446–1457, April 2021. ISSN 2213-6711. doi: 10.1016/j.stemcr.2021.03.025. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8190574/.

[10] Dvir Aran, Agnieszka P. Looney, Leqian Liu, Esther Wu, Valerie Fong, Austin Hsu, Suzanna Chak, Ram P. Naikawadi, Paul J. Wolters, Adam R. Abate, Atul J. Butte, and Mallar Bhattacharya. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nature Immunology*, 20(2):163–172, February 2019. ISSN 1529-2916. doi: 10.1038/s41590-018-0276-y. URL https://www.nature.com/articles/s41590-018-0276-y. Number: 2 Publisher: Nature Publishing Group.

[11] Bulent Ataman, Gabriella L. Boulting, David A. Harmin, Marty G. Yang, Mollie Baker-Salisbury, Ee-Lynn Yap, Athar N. Malik, Kevin Mei, Alex A. Rubin, Ivo Spiegel, Ershela Durresi, Nikhil Sharma, Linda S. Hu, Mihovil Pletikos, Eric C. Griffith, Jennifer N. Partlow, Christine R. Stevens, Mazhar Adli, Maria Chahrour, Nenad Sestan, Christopher A. Walsh, Vladimir K. Berezovskii, Margaret S. Livingstone, and Michael E. Greenberg. Evolution of Osteocrin as an activity-regulated factor in the primate brain. *Nature*, 539(7628):242–247,

November 2016. ISSN 0028-0836. doi: 10.1038/nature20111. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5499253/.

[12] D. Baas, A. Meiniel, C. Benadiba, E. Bonnafe, O. Meiniel, W. Reith, and B. Durand. A deficiency in RFX3 causes hydrocephalus associated with abnormal differentiation of ependymal cells. *The European Journal of Neuroscience*, 24(4):1020–1030, August 2006. ISSN 0953-816X. doi: 10.1111/j.1460-9568.2006.05002.x.

[13] Jose L. Badano, Norimasa Mitsuma, Phil L. Beales, and Nicholas Katsanis. The ciliopathies: an emerging class of human genetic disorders. *Annual Review of Genomics and Human Genetics*, 7:125–148, 2006. ISSN 1527-8204. doi: 10.1146/annurev.genom.7.080505.115610.

[14] Daehyun Baek and Phil Green. Sequence conservation, relative isoform frequencies, and nonsense-mediated decay in evolutionarily conserved alternative splicing. *Proceedings of the National Academy of Sciences of the United States of America*, 102(36):12813–12818, September 2005. ISSN 0027-8424. doi: 10.1073/pnas.0506139102.

[15] Joshua A. Bagley, Daniel Reumann, Shan Bian, Julie Lévi-Strauss, and Juergen A. Knoblich. Fused cerebral organoids model interactions between brain regions. *Nature Methods*, 14(7):743–751, July 2017. ISSN 1548-7105. doi: 10.1038/nmeth.4304.

[16] Timothy L. Bailey, James Johnson, Charles E. Grant, and William S. Noble. The MEME Suite. *Nucleic Acids Research*, 43(W1):W39–49, July 2015. ISSN 1362-4962. doi: 10.1093/nar/gkv416.

[17] Trygve E. Bakken, Nikolas L. Jorstad, Qiwen Hu, Blue B. Lake, Wei Tian, Brian E. Kalmbach, Megan Crow, Rebecca D. Hodge, Fenna M. Krienen, Staci A. Sorensen, Jeroen Eggermont, Zizhen Yao, Brian D. Aevermann, Andrew I. Aldridge, Anna Bartlett, Darren Bertagnolli, Tamara Casper, Rosa G. Castanon, Kirsten Crichton, Tanya L. Daigle, Rachel Dalley, Nick Dee, Nikolai Dembrow, Dinh Diep, Song-Lin Ding, Weixiu Dong, Rongxin Fang, Stephan Fischer, Melissa Goldman, Jeff Goldy, Lucas T. Graybuck, Brian R. Herb, Xiaomeng Hou, Jayaram Kancherla, Matthew Kroll, Kanan Lathia, Baldur van Lew, Yang Eric Li, Christine S. Liu, Hanqing Liu, Jacinta D. Lucero, Anup Mahurkar, Delissa McMillen, Jeremy A. Miller, Marmar Moussa, Joseph R. Nery, Philip R. Nicovich, Sheng-Yong Niu, Joshua Orvis, Julia K. Osteen, Scott Owen, Carter R. Palmer, Thanh Pham, Nongluk Plongthongkum, Olivier Poirion, Nora M. Reed, Christine Rimorin, Angeline Rivkin, William J. Romanow, Adriana E. Sedeño-Cortés, Kimberly Siletti, Saroja Somasundaram, Josef Sulc, Michael Tieu, Amy Torkelson, Herman Tung, Xinxin Wang, Fangming Xie, Anna Marie Yanny, Renee Zhang, Seth A. Ament, M. Margarita Behrens, Hector Corrada Bravo, Jerold Chun, Alexander Dobin, Jesse Gillis, Ronna Hertzano, Patrick R. Hof, Thomas Höllt, Gregory D. Horwitz, C. Dirk Keene, Peter V. Kharchenko, Andrew L. Ko, Boudewijn P. Lelieveldt, Chongyuan Luo, Eran A. Mukamel, António

Pinto-Duarte, Sebastian Preissl, Aviv Regev, Bing Ren, Richard H. Scheuermann, Kimberly Smith, William J. Spain, Owen R. White, Christof Koch, Michael Hawrylycz, Bosiljka Tasic, Evan Z. Macosko, Steven A. McCarroll, Jonathan T. Ting, Hongkui Zeng, Kun Zhang, Guoping Feng, Joseph R. Ecker, Sten Linnarsson, and Ed S. Lein. Comparative cellular analysis of motor cortex in human, marmoset and mouse. *Nature*, 598(7879): 111–119, October 2021. ISSN 1476-4687. doi: 10.1038/s41586-021-03465-8. URL https://www.nature.com/articles/s41586-021-03465-8. Number: 7879 Publisher: Nature Publishing Group.

[18] Tobias Banaschewski, Katja Becker, Susann Scherag, Barbara Franke, and David Coghill. Molecular genetics of attention-deficit/hyperactivity disorder: an overview. *European Child & Adolescent Psychiatry*, 19(3):237–257, 2010. ISSN 1018-8827. doi: 10.1007/s00787-010-0090-z. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2839490/.

[19] Nuno L. Barbosa-Morais, Manuel Irimia, Qun Pan, Hui Y. Xiong, Serge Gueroussov, Leo J. Lee, Valentina Slobodeniuc, Claudia Kutter, Stephen Watt, Recep Colak, TaeHyung Kim, Christine M. Misquitta-Ali, Michael D. Wilson, Philip M. Kim, Duncan T. Odom, Brendan J. Frey, and Benjamin J. Blencowe. The evolutionary landscape of alternative splicing in vertebrate species. *Science (New York, N.Y.)*, 338(6114):1587–1593, December 2012. ISSN 1095-9203. doi: 10.1126/science.1230612.

[20] Artem Barski, Suresh Cuddapah, Kairong Cui, Tae-Young Roh, Dustin E. Schones, Zhibin Wang, Gang Wei, Iouri Chepelev, and Keji Zhao. High-Resolution Profiling of Histone Methylations in the Human Genome. *Cell*, 129(4):823–837, May 2007. ISSN 0092-8674, 1097-4172. doi: 10.1016/j.cell.2007.05.009. URL https://www.cell.com/cell/abstract/S0092-8674(07)00600-9. Publisher: Elsevier.

[21] Marek Bartosovic, Mukund Kabbe, and Gonçalo Castelo-Branco. Single-cell CUT&Tag profiles histone modifications and transcription factors in complex tissues. *Nature biotechnology*, 39(7):825–835, July 2021. ISSN 1087-0156. doi: 10.1038/s41587-021-00869-9. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7611252/.

[22] Aziz Belkadi, Alexandre Bolze, Yuval Itan, Aurélie Cobat, Quentin B. Vincent, Alexander Antipenko, Lei Shang, Bertrand Boisson, Jean-Laurent Casanova, and Laurent Abel. Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants. *Proceedings of the National Academy of Sciences*, 112(17):5473–5478, April 2015. doi: 10.1073/pnas.1418631112. URL https://www.pnas.org/doi/abs/10.1073/pnas.1418631112. Publisher: Proceedings of the National Academy of Sciences.

[23] Carine Benadiba, Dario Magnani, Mathieu Niquille, Laurette Morlé, Delphine Valloton, Homaira Nawabi, Aouatef Ait-Lounis, Belkacem Otsmane, Walter Reith, Thomas Theil, Jean-Pierre Hornung, Cécile Lebrand, and Bénédicte Durand. The ciliogenic transcription factor RFX3 regulates early midline distribution of guidepost neurons required for corpus

callosum development. *PLoS genetics*, 8(3):e1002606, 2012. ISSN 1553-7404. doi: 10.1371/journal.pgen.1002606.

[24] Fikri Birey, Jimena Andersen, Christopher D. Makinson, Saiful Islam, Wu Wei, Nina Huber, H. Christina Fan, Kimberly R. Cordes Metzler, Georgia Panagiotakos, Nicholas Thom, Nancy A. O'Rourke, Lars M. Steinmetz, Jonathan A. Bernstein, Joachim Hallmayer, John R. Huguenard, and Sergiu P. Paşca. Assembly of functionally integrated human forebrain spheroids. *Nature*, 545(7652):54–59, May 2017. ISSN 1476-4687. doi: 10.1038/nature22330. URL https://www.nature.com/articles/nature22330. Number: 7652 Publisher: Nature Publishing Group.

[25] E. Bonnafe, M. Touka, A. AitLounis, D. Baas, E. Barras, C. Ucla, A. Moreau, F. Flamant, R. Dubruille, P. Couble, J. Collignon, B. Durand, and W. Reith. The transcription factor RFX3 directs nodal cilium development and left-right asymmetry specification. *Molecular and Cellular Biology*, 24(10):4417–4427, May 2004. ISSN 0270-7306. doi: 10.1128/MCB.24.10.4417-4427.2004.

[26] Gabriella L. Boulting, Ershela Durresi, Bulent Ataman, Maxwell A. Sherman, Kevin Mei, David A. Harmin, Ava C. Carter, Daniel R. Hochbaum, Adam J. Granger, Jesse M Engreitz, Sinisa Hrvatin, Michael R. Blanchard, Marty G. Yang, Eric C. Griffith, and Michael E. Greenberg. Activity-dependent regulome of human GABAergic neurons reveals new patterns of gene regulation and neurological disease heritability. *Nature neuroscience*, 24(3):437–448, March 2021. ISSN 1097-6256. doi: 10.1038/s41593-020-00786-1. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7933108/.

[27] Julie Bourseguin, Wen Cheng, Emily Talbot, Liana Hardy, Jenny Lai, Ailsa M Jeffries, Michael A Lodato, Eunjung Alice Lee, and Svetlana V Khoronenkova. Persistent DNA damage associated with ATM kinase deficiency promotes microglial dysfunction. *Nucleic Acids Research*, page gkac104, February 2022. ISSN 0305-1048. doi: 10.1093/nar/gkac104. URL https://doi.org/10.1093/nar/gkac104.

[28] Alan P. Boyle, Eurie L. Hong, Manoj Hariharan, Yong Cheng, Marc A. Schaub, Maya Kasowski, Konrad J. Karczewski, Julie Park, Benjamin C. Hitz, Shuai Weng, J. Michael Cherry, and Michael Snyder. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Research*, 22(9):1790–1797, September 2012. ISSN 1549-5469. doi: 10.1101/gr.137323.112.

[29] Ulrich Braunschweig, Nuno L. Barbosa-Morais, Qun Pan, Emil N. Nachman, Babak Alipanahi, Thomas Gonatopoulos-Pournatzis, Brendan Frey, Manuel Irimia, and Benjamin J. Blencowe. Widespread intron retention in mammals functionally tunes transcriptomes. *Genome Research*, 24(11):1774–1786, November 2014. ISSN 1549-5469. doi: 10.1101/gr.177790.114.

[30] Kristen Brennand, Anthony Simone, Jessica Jou, Chelsea Gelboin-Burkhart, Ngoc Tran, Sarah Sangar, Yan Li, Yangling Mu, Gong Chen, Diana Yu, Shane McCarthy, Jonathan Sebat, and Fred H. Gage. Modeling schizophrenia using hiPSC neurons. *Nature*, 473(7346): 221, May 2011. doi: 10.1038/nature09915. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3392969/. Publisher: NIH Public Access.

[31] Joshua J. Breunig, Matthew R. Sarkisian, Jon I. Arellano, Yury M. Morozov, Albert E. Ayoub, Sonal Sojitra, Baolin Wang, Richard A. Flavell, Pasko Rakic, and Terrence Town. Primary cilia regulate hippocampal neurogenesis by mediating sonic hedgehog signaling. *Proceedings of the National Academy of Sciences*, 105(35):13127–13132, September 2008. doi: 10.1073/pnas.0804558105. URL https://www.pnas.org/doi/abs/10.1073/pnas.0804558105. Publisher: Proceedings of the National Academy of Sciences.

[32] David E. Buchholz, Thomas S. Carroll, Arif Kocabas, Xiaodong Zhu, Hourinaz Behesti, Phyllis L. Faust, Lauren Stalbow, Yin Fang, and Mary E. Hatten. Novel genetic features of human and mouse Purkinje cell differentiation defined by comparative transcriptomics. *Proceedings of the National Academy of Sciences*, 117(26):15085–15095, June 2020. doi: 10.1073/pnas.2000102117. URL https://www.pnas.org/doi/10.1073/pnas.2000102117. Publisher: Proceedings of the National Academy of Sciences.

[33] Moritz O. Buchholz, Alexandra Gastone Guilabert, Benjamin Ehret, and Gregor F. P. Schuhknecht. How synaptic strength, short-term plasticity, and input synchrony contribute to neuronal spike output. *PLOS Computational Biology*, 19(4):e1011046, April 2023. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1011046. URL https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1011046. Publisher: Public Library of Science.

[34] Jason Buenrostro, Beijing Wu, Howard Chang, and William Greenleaf. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.]*, 109:21.29.1–21.29.9, January 2015. ISSN 1934-3639. doi: 10.1002/0471142727.mb2129s109. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4374986/.

[35] Emily E. Burke, Joshua G. Chenoweth, Joo Heon Shin, Leonardo Collado-Torres, Suel-Kee Kim, Nicola Micali, Yanhong Wang, Carlo Colantuoni, Richard E. Straub, Daniel J. Hoeppner, Huei-Ying Chen, Alana Sellers, Kamel Shibbani, Gregory R. Hamersky, Marcelo Diaz Bustamante, BaDoi N. Phan, William S. Ulrich, Cristian Valencia, Amritha Jaishankar, Amanda J. Price, Anandita Rajpurohit, Stephen A. Semick, Roland W. Bürli, James C. Barrow, Daniel J. Hiler, Stephanie C. Page, Keri Martinowich, Thomas M. Hyde, Joel E. Kleinman, Karen F. Berman, Jose A. Apud, Alan J. Cross, Nicholas J. Brandon, Daniel R. Weinberger, Brady J. Maher, Ronald D. G. McKay, and Andrew E. Jaffe. Dissecting transcriptomic signatures of neuronal differentiation and maturation using iPSCs. *Nature Communications*, 11(1):462, January 2020. ISSN 2041-1723. doi: 10.1038/s41467-019-14266-z.

URL https://www.nature.com/articles/s41467-019-14266-z. Number: 1 Publisher: Nature Publishing Group.

[36] Davide Cacchiarelli, Xiaojie Qiu, Sanjay Srivatsan, Anna Manfredi, Michael Ziller, Eliah Overbey, Antonio Grimaldi, Jonna Grimsby, Prapti Pokharel, Kenneth J. Livak, Shuqiang Li, Alexander Meissner, Tarjei S. Mikkelsen, John L. Rinn, and Cole Trapnell. Aligning Single-Cell Developmental and Reprogramming Trajectories Identifies Molecular Determinants of Myogenic Reprogramming Outcome. *Cell Systems*, 7(3):258–268.e3, September 2018. ISSN 24054712. doi: 10.1016/j.cels.2018.07.006. URL https://linkinghub.elsevier.com/retrieve/pii/S2405471218303144.

[37] Bilal Cakir, Yangfei Xiang, Yoshiaki Tanaka, Mehmet H. Kural, Maxime Parent, Young-Jin Kang, Kayley Chapeton, Benjamin Patterson, Yifan Yuan, Chang-Shun He, Micha Sam B. Raredon, Jake Dengelegi, Kun-Yong Kim, Pingnan Sun, Mei Zhong, Sangho Lee, Prabir Patra, Fahmeed Hyder, Laura E. Niklason, Sang-Hun Lee, Young-Sup Yoon, and In-Hyun Park. Engineering of human brain organoids with a functional vascular-like system. *Nature Methods*, 16(11):1169–1175, November 2019. ISSN 1548-7105. doi: 10.1038/s41592-019-0586-5. URL https://www.nature.com/articles/s41592-019-0586-5. Number: 11 Publisher: Nature Publishing Group.

[38] Isaac Canals, Aurélie Ginisty, Ella Quist, Raissa Timmerman, Jonas Fritze, Giedre Miskinyte, Emanuela Monni, Marita G. Hansen, Isabel Hidalgo, David Bryder, Johan Bengzon, and Henrik Ahlenius. Rapid and efficient induction of functional astrocytes from human pluripotent stem cells. *Nature Methods*, 15(9):693–696, September 2018. ISSN 1548-7105. doi: 10.1038/s41592-018-0103-2. URL https://www.nature.com/articles/s41592-018-0103-2. Number: 9 Publisher: Nature Publishing Group.

[39] Meixin Chen, Qingcai Meng, Yunfei Qin, Puping Liang, Peng Tan, Lian He, Yubin Zhou, Yongjun Chen, Junjiu Huang, Rong-Fu Wang, and Jun Cui. TRIM14 Inhibits cGAS Degradation Mediated by Selective Autophagy Receptor p62 to Promote Innate Immune Responses. *Molecular Cell*, 64(1):105–119, October 2016. ISSN 1097-4164. doi: 10.1016/j.molcel.2016.08.025.

[40] Shih-Wei Chen, Yu-Sheng Hung, Jong-Ling Fuh, Nien-Jung Chen, Yeh-Shiu Chu, Shu-Cian Chen, Ming-Ji Fann, and Yu-Hui Wong. Efficient conversion of human induced pluripotent stem cells into microglia by defined transcription factors. *Stem Cell Reports*, 16(5):1363–1380, May 2021. ISSN 2213-6711. doi: 10.1016/j.stemcr.2021.03.010.

[41] Ya Chen, Luke Bury, Fu Chen, Kimberly A Aldinger, Helen C Miranda, and Anthony Wynshaw-Boris. Generation of advanced cerebellar organoids for neurogenesis and neuronal network development. *Human Molecular Genetics*, page ddad110, June 2023. ISSN 0964-6906. doi: 10.1093/hmg/ddad110. URL https://doi.org/10.1093/hmg/ddad110.

[42] Aifang Cheng, Kai-Hei Tse, Hei-Man Chow, Yunqiao Gan, Xuan Song, Fulin Ma, Yi Xuan Yvonne Qian, Weiyi She, and Karl Herrup. ATM loss disrupts the autophagy-lysosomal pathway. *Autophagy*, 17(8):1998–2010, August 2021. ISSN 1554-8635. doi: 10.1080/15548627.2020.1805860.

[43] Semil P. Choksi, Gilbert Lauter, Peter Swoboda, and Sudipto Roy. Switching on cilia: transcriptional networks regulating ciliogenesis. *Development (Cambridge, England)*, 141(7): 1427–1441, April 2014. ISSN 1477-9129. doi: 10.1242/dev.074666.

[44] Jeffrey S. C. Chu, David L. Baillie, and Nansheng Chen. Convergent evolution of RFX transcription factors and ciliary genes predated the origin of metazoans. *BMC evolutionary biology*, 10:130, May 2010. ISSN 1471-2148. doi: 10.1186/1471-2148-10-130.

[45] Bradley P. Coe, Holly A. F. Stessman, Arvis Sulovari, Madeleine R. Geisheker, Trygve E. Bakken, Allison M. Lake, Joseph D. Dougherty, Ed S. Lein, Fereydoun Hormozdiari, Raphael A. Bernier, and Evan E. Eichler. Neurodevelopmental disease genes implicated by de novo mutation and copy number variation morbidity. *Nature Genetics*, 51(1):106–116, January 2019. ISSN 1546-1718. doi: 10.1038/s41588-018-0288-4.

[46] Auriane Cospain, Ana Rivera-Barahona, Erwan Dumontet, Blanca Gener, Isabelle Bailleul-Forestier, Isabelle Meyts, Guillaume Jouret, Bertrand Isidor, Carole Brewer, Wim Wuyts, Leen Moens, Selket Delafontaine, Wayne Wing Keung Lam, Kris Van Den Bogaert, Anneleen Boogaerts, Emmanuel Scalais, Thomas Besnard, Benjamin Cogne, Christophe Guissard, Paul Rollier, Wilfrid Carre, Regis Bouvet, Karin Tarte, Ricardo Gómez-Carmona, Pablo Lapunzina, Sylvie Odent, Marie Faoucher, Christele Dubourg, Víctor L. Ruiz-Pérez, Koen Devriendt, Laurent Pasquier, and Luis A. Pérez-Jurado. FOSL2 truncating variants in the last exon cause a neurodevelopmental disorder with scalp and enamel defects. *Genetics in Medicine: Official Journal of the American College of Medical Genetics*, 24(12):2475–2486, December 2022. ISSN 1530-0366. doi: 10.1016/j.gim.2022.09.002.

[47] Petr Danecek, James K Bonfield, Jennifer Liddle, John Marshall, Valeriu Ohan, Martin O Pollard, Andrew Whitwham, Thomas Keane, Shane A McCarthy, Robert M Davies, and Heng Li. Twelve years of SAMtools and BCFtools. *GigaScience*, 10(2):giab008, February 2021. ISSN 2047-217X. doi: 10.1093/gigascience/giab008. URL https://doi.org/10.1093/gigascience/giab008.

[48] Hélène Darville, Aurélie Poulet, Frédérique Rodet-Amsellem, Laure Chatrousse, Julie Pernelle, Claire Boissart, Delphine Héron, Caroline Nava, Anselme Perrier, Margot Jarrige, Francis Cogé, Mark J. Millan, Thomas Bourgeron, Marc Peschanski, Richard Delorme, and Alexandra Benchoua. Human Pluripotent Stem Cell-derived Cortical Neurons for High Throughput Medication Screening in Autism: A Proof of Concept Study in SHANK3 Haploinsufficiency Syndrome. *EBioMedicine*, 9:293–305, July 2016. ISSN 2352-3964. doi: 10.1016/j.ebiom.2016.05.032.

[49] Carrie A. Davis, Benjamin C. Hitz, Cricket A. Sloan, Esther T. Chan, Jean M. Davidson, Idan Gabdank, Jason A. Hilton, Kriti Jain, Ulugbek K. Baymuradov, Aditi K. Narayanan, Kathrina C. Onate, Keenan Graham, Stuart R. Miyasato, Timothy R. Dreszer, J. Seth Strattan, Otto Jolanki, Forrest Y. Tanaka, and J. Michael Cherry. The Encyclopedia of DNA elements (ENCODE): data portal update. *Nucleic Acids Research*, 46(D1):D794–D801, January 2018. ISSN 1362-4962. doi: 10.1093/nar/gkx1081.

[50] Marco De Cecco, Takahiro Ito, Anna P. Petrashen, Amy E. Elias, Nicholas J. Skvir, Steven W. Criscione, Alberto Caligiana, Greta Brocculi, Emily M. Adney, Jef D. Boeke, Oanh Le, Christian Beauséjour, Jayakrishna Ambati, Kameshwari Ambati, Matthew Simon, Andrei Seluanov, Vera Gorbunova, P. Eline Slagboom, Stephen L. Helfand, Nicola Neretti, and John M. Sedivy. L1 drives IFN in senescent cells and promotes age-associated inflammation. *Nature*, 566(7742):73–78, February 2019. ISSN 1476-4687. doi: 10.1038/s41586-018-0784-9. URL https://www.nature.com/articles/s41586-018-0784-9. Number: 7742 Publisher: Nature Publishing Group.

[51] Ditte Demontis, Francesco Lescai, Anders Børglum, Simon Glerup, Søren Dinesen Østergaard, Ole Mors, Qibin Li, Jieqin Liang, Hui Jiang, Yingrui Li, Jun Wang, Klaus-Peter Lesch, Andreas Reif, Jan K. Buitelaar, and Barbara Franke. Whole-Exome Sequencing Reveals Increased Burden of Rare Functional and Disruptive Variants in Candidate Risk Genes in Individuals With Persistent Attention-Deficit/Hyperactivity Disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, 55(6):521–523, June 2016. ISSN 0890-8567. doi: 10.1016/j.jaac.2016.03.009. URL https://www.sciencedirect.com/science/article/pii/S0890856716300971.

[52] Eric Deneault, Sean H. White, Deivid C. Rodrigues, P. Joel Ross, Muhammad Faheem, Kirill Zaslavsky, Zhuozhi Wang, Roumiana Alexandrova, Giovanna Pellecchia, Wei Wei, Alina Piekna, Gaganjot Kaur, Jennifer L. Howe, Vickie Kwan, Bhooma Thiruvahindrapuram, Susan Walker, Anath C. Lionel, Peter Pasceri, Daniele Merico, Ryan K.C. Yuen, Karun K. Singh, James Ellis, and Stephen W. Scherer. Complete Disruption of Autism-Susceptibility Genes by Gene Editing Predominantly Reduces Functional Connectivity of Isogenic Human Neurons. *Stem Cell Reports*, 11(5):1211–1225, November 2018. ISSN 2213-6711. doi: 10.1016/j.stemcr.2018.10.003. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6235011/.

[53] Brooke A. DeRosa, Jimmy El Hokayem, Elena Artimovich, Catherine Garcia-Serje, Andre W. Phillips, Derek Van Booven, Jonathan E. Nestor, Lily Wang, Michael L. Cuccaro, Jeffery M. Vance, Margaret A. Pericak-Vance, Holly N. Cukier, Michael W. Nestor, and Derek M. Dykxhoorn. Convergent Pathways in Idiopathic Autism Revealed by Time Course Transcriptomic Analysis of Patient-Derived Neurons. *Scientific Reports*, 8:8423, May 2018. ISSN 2045-2322. doi: 10.1038/s41598-018-26495-1. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5976773/.

[54] Ashish Dhir, Somdutta Dhir, Lukasz S. Borowski, Laura Jimenez, Michael Teitell, Agnès Rötig, Yanick J. Crow, Gillian I. Rice, Darragh Duffy, Christelle Tamby, Takayuki Nojima, Arnold Munnich, Manuel Schiff, Claudia Ribeiro de Almeida, Jan Rehwinkel, Andrzej Dziembowski, Roman J. Szczesny, and Nicholas J. Proudfoot. Mitochondrial double-stranded RNA triggers antiviral signalling in humans. *Nature*, 560(7717): 238–242, August 2018. ISSN 1476-4687. doi: 10.1038/s41586-018-0363-0. URL https://www.nature.com/articles/s41586-018-0363-0. Number: 7717 Publisher: Nature Publishing Group.

[55] Ilan Dinstein, Karen Pierce, Lisa Eyler, Stephanie Solso, Rafael Malach, Marlene Behrmann, and Eric Courchesne. Disrupted neural synchronization in toddlers with autism. *Neuron*, 70(6):1218–1225, June 2011. ISSN 0896-6273. doi: 10.1016/j.neuron.2011.04.018. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3119852/.

[56] Tracy Dixon-Salazar, Jennifer L. Silhavy, Sarah E. Marsh, Carrie M. Louie, Lesley C. Scott, Aithala Gururaj, Lihadh Al-Gazali, Asma A. Al-Tawari, Hulya Kayserili, László Sztriha, and Joseph G. Gleeson. Mutations in the AHI1 gene, encoding jouberin, cause Joubert syndrome with cortical polymicrogyria. *American Journal of Human Genetics*, 75(6):979–987, December 2004. ISSN 0002-9297. doi: 10.1086/425985.

[57] Ryan N. Doan, Elaine T. Lim, Silvia De Rubeis, Catalina Betancur, David J. Cutler, Andreas G. Chiocchetti, Lynne M. Overman, Aubrie Soucy, Susanne Goetze, Autism Sequencing Consortium, Christine M. Freitag, Mark J. Daly, Christopher A. Walsh, Joseph D. Buxbaum, and Timothy W. Yu. Recessive gene disruptions in autism spectrum disorder. *Nature Genetics*, 51(7):1092–1098, July 2019. ISSN 1546-1718. doi: 10.1038/s41588-019-0433-8.

[58] Alexander Dobin, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, and Thomas R. Gingeras. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1):15–21, January 2013. ISSN 1367-4803. doi: 10.1093/bioinformatics/bts635. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3530905/.

[59] Zhixun Dou, Kanad Ghosh, Maria Grazia Vizioli, Jiajun Zhu, Payel Sen, Kirk J. Wangensteen, Johayra Simithy, Yemin Lan, Yanping Lin, Zhuo Zhou, Brian C. Capell, Caiyue Xu, Mingang Xu, Julia E. Kieckhaefer, Tianying Jiang, Michal Shoshkes-Carmel, K. M. Ahasan Al Tanim, Glen N. Barber, John T. Seykora, Sarah E. Millar, Klaus H. Kaestner, Benjamin A. Garcia, Peter D. Adams, and Shelley L. Berger. Cytoplasmic chromatin triggers inflammation in senescence and cancer. *Nature*, 550(7676):402–406, October 2017. ISSN 1476-4687. doi: 10.1038/nature24050. URL https://www.nature.com/articles/nature24050. Number: 7676 Publisher: Nature Publishing Group.

[60] Allison D. Ebert, Junying Yu, Ferrill F. Rose, Virginia B. Mattis, Christian L. Lorson, James A. Thomson, and Clive N. Svendsen. Induced pluripotent stem cells from a spinal muscular atrophy patient. *Nature*, 457(7227):277–280, January 2009. ISSN 0028-0836. doi: 10.1038/nature07677. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2659408/.

[61] Daniel H. Ebert and Michael E. Greenberg. Activity-dependent neuronal signalling and autism spectrum disorder. *Nature*, 493(7432):327–337, January 2013. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature11860. URL http://www.nature.com/articles/nature11860.

[62] Evgeni Efimenko, Kerry Bubb, Ho Yi Mak, Ted Holzman, Michel R. Leroux, Gary Ruvkun, James H. Thomas, and Peter Swoboda. Analysis of xbx genes in C. elegans. *Development (Cambridge, England)*, 132(8):1923–1934, April 2005. ISSN 0950-1991. doi: 10.1242/dev.01775.

[63] Marc Ehrlich, Sabah Mozafari, Michael Glatza, Laura Starost, Sergiy Velychko, Anna-Lena Hallmann, Qiao-Ling Cui, Axel Schambach, Kee-Pyo Kim, Corinne Bachelin, Antoine Marteyn, Gunnar Hargus, Radia Marie Johnson, Jack Antel, Jared Sterneckert, Holm Zaehres, Hans R. Schöler, Anne Baron-Van Evercooren, and Tanja Kuhlmann. Rapid and efficient generation of oligodendrocytes from human induced pluripotent stem cells using transcription factors. *Proceedings of the National Academy of Sciences of the United States of America*, 114(11):E2243, March 2017. doi: 10.1073/pnas.1614412114. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5358375/. Publisher: National Academy of Sciences.

[64] Loubna El Zein, Aouatef Ait-Lounis, Laurette Morlé, Joëlle Thomas, Brigitte Chhin, Nathalie Spassky, Walter Reith, and Bénédicte Durand. RFX3 governs growth and beating efficiency of motile cilia in mouse and controls the expression of genes involved in human ciliopathies. *Journal of Cell Science*, 122(Pt 17):3180–3189, September 2009. ISSN 1477-9137. doi: 10.1242/jcs.048348.

[65] Ran Elkon, Beatrice Milon, Laura Morrison, Manan Shah, Sarath Vijayakumar, Manoj Racherla, Carmen C. Leitch, Lorna Silipino, Shadan Hadi, Michèle Weiss-Gayet, Emmanuèle Barras, Christoph D. Schmid, Aouatef Ait-Lounis, Ashley Barnes, Yang Song, David J. Eisenman, Efrat Eliyahu, Gregory I. Frolenkov, Scott E. Strome, Bénédicte Durand, Norann A. Zaghloul, Sherri M. Jones, Walter Reith, and Ronna Hertzano. RFX transcription factors are essential for hearing in mice. *Nature Communications*, 6:8549, October 2015. ISSN 2041-1723. doi: 10.1038/ncomms9549.

[66] Muhammad Farooq, Louise Lindbæk, Nicolai Krogh, Canan Doganli, Cecilie Keller, Maren Mönnich, André Brás Gonçalves, Srinivasan Sakthivel, Yuan Mang, Ambrin Fatima, Vivi Søgaard Andersen, Muhammad S. Hussain, Hans Eiberg, Lars Hansen, Klaus Wilbrandt Kjaer, Jay Gopalakrishnan, Lotte Bang Pedersen, Kjeld Møllgård, Henrik Nielsen, Shahid M. Baig,

Niels Tommerup, Søren Tvorup Christensen, and Lars Allan Larsen. RRP7A links primary microcephaly to dysfunction of ribosome biogenesis, resorption of primary cilia, and neurogenesis. *Nature Communications*, 11(1):5816, November 2020. ISSN 2041-1723. doi: 10.1038/s41467-020-19658-0. URL https://www.nature.com/articles/s41467-020-19658-0. Number: 1 Publisher: Nature Publishing Group.

[67] Christof Fellmann, Benjamin G. Gowen, Pei-Chun Lin, Jennifer A. Doudna, and Jacob E. Corn. Cornerstones of CRISPR–Cas in drug discovery and therapy. *Nature Reviews Drug Discovery*, 16(2):89–100, February 2017. ISSN 1474-1784. doi: 10.1038/nrd.2016.238. URL https://www.nature.com/articles/nrd.2016.238. Number: 2 Publisher: Nature Publishing Group.

[68] Alejandra Fernandez, Daniel W. Meechan, Beverly A. Karpinski, Elizabeth M. Paronett, Corey A. Bryan, Hanna L. Rutz, Eric A. Radin, Noah Lubin, Erin R. Bonner, Anastas Popratiloff, Lawrence A. Rothblat, Thomas M. Maynard, and Anthony-Samuel LaMantia. Mitochondrial Dysfunction Leads to Cortical Under-Connectivity and Cognitive Impairment. *Neuron*, 102(6):1127–1142.e3, June 2019. ISSN 1097-4199. doi: 10.1016/j.neuron.2019.04.013.

[69] Richard S. Finkel, Eugenio Mercuri, Basil T. Darras, Anne M. Connolly, Nancy L. Kuntz, Janbernd Kirschner, Claudia A. Chiriboga, Kayoko Saito, Laurent Servais, Eduardo Tizzano, Haluk Topaloglu, Már Tulinius, Jacqueline Montes, Allan M. Glanzman, Kathie Bishop, Z. John Zhong, Sarah Gheuens, C. Frank Bennett, Eugene Schneider, Wildon Farwell, Darryl C. De Vivo, and ENDEAR Study Group. Nusinersen versus Sham Control in Infantile-Onset Spinal Muscular Atrophy. *The New England Journal of Medicine*, 377(18):1723–1732, November 2017. ISSN 1533-4406. doi: 10.1056/NEJMoa1702752.

[70] Stephen J. Fleming, Mark D. Chaffin, Alessandro Arduini, Amer-Denis Akkad, Eric Banks, John C. Marioni, Anthony A. Philippakis, Patrick T. Ellinor, and Mehrtash Babadi. Unsupervised removal of systematic background noise from droplet-based single-cell experiments using CellBender, July 2022. URL https://www.biorxiv.org/content/10.1101/791699v2. Pages: 791699 Section: New Results.

[71] Oriol Fornes, Jaime A. Castro-Mondragon, Aziz Khan, Robin van der Lee, Xi Zhang, Phillip A. Richmond, Bhavi P. Modi, Solenne Correard, Marius Gheorghe, Damir Baranašić, Walter Santana-Garcia, Ge Tan, Jeanne Chèneby, Benoit Ballester, François Parcy, Albin Sandelin, Boris Lenhard, Wyeth W. Wasserman, and Anthony Mathelier. JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Research*, 48(D1):D87–D92, January 2020. ISSN 1362-4962. doi: 10.1093/nar/gkz1001.

[72] Michael J. Gandal, Jillian R. Haney, Brie Wamsley, Chloe X. Yap, Sepideh Parhami, Prashant S. Emani, Nathan Chang, George T. Chen, Gil D. Hoftman, Diego de Alba, Gokul Ramaswami, Christopher L. Hartl, Arjun Bhattacharya, Chongyuan Luo, Ting Jin,

Daifeng Wang, Riki Kawaguchi, Diana Quintero, Jing Ou, Ye Emily Wu, Neelroop N. Parikshak, Vivek Swarup, T. Grant Belgard, Mark Gerstein, Bogdan Pasaniuc, and Daniel H. Geschwind. Broad transcriptomic dysregulation occurs across the cerebral cortex in ASD. *Nature*, 611(7936):532–539, November 2022. ISSN 1476-4687. doi: 10.1038/s41586-022-05377-7. URL https://www.nature.com/articles/s41586-022-05377-7. Number: 7936 Publisher: Nature Publishing Group.

[73] Tianshun Gao and Jiang Qian. EnhancerAtlas 2.0: an updated resource with enhancer annotation in 586 tissue/cell types across nine species. *Nucleic Acids Research*, 48(D1):D58–D64, January 2020. ISSN 0305-1048. doi: 10.1093/nar/gkz980. URL https://doi.org/10.1093/nar/gkz980.

[74] Mary Gazea, Evangelia Tasouri, Marianna Tolve, Viktoria Bosch, Anna Kabanova, Christian Gojak, Bahtiyar Kurtulmus, Orna Novikov, Joachim Spatz, Gislene Pereira, Wolfgang Hübner, Claude Brodski, Kerry L. Tucker, and Sandra Blaess. Primary cilia are critical for Sonic hedgehog-mediated dopaminergic neurogenesis in the embryonic midbrain. *Developmental Biology*, 409(1):55–71, January 2016. ISSN 0012-1606. doi: 10.1016/j.ydbio.2015.10.033. URL https://www.sciencedirect.com/science/article/pii/S0012160615302633.

[75] L. Giordano, A. Vignoli, L. Pinelli, F. Brancati, P. Accorsi, F. Faravelli, R. Gasparotti, T. Granata, G. Giaccone, F. Inverardi, C. Frassoni, B. Dallapiccola, E. M. Valente, and R. Spreafico. Joubert syndrome with bilateral polymicrogyria: clinical and neuropathological findings in two brothers. *American Journal of Medical Genetics. Part A*, 149A(7):1511–1515, July 2009. ISSN 1552-4833. doi: 10.1002/ajmg.a.32936.

[76] David Goertsen, Nicholas C. Flytzanis, Nick Goeden, Miguel R. Chuapoco, Alexander Cummins, Yijing Chen, Yingying Fan, Qiangge Zhang, Jitendra Sharma, Yangyang Duan, Liping Wang, Guoping Feng, Yu Chen, Nancy Y. Ip, James Pickel, and Viviana Gradinaru. AAV capsid variants with brain-wide transgene expression and decreased liver targeting after intravenous delivery in mouse and marmoset. *Nature Neuroscience*, 25(1):106–115, January 2022. ISSN 1546-1726. doi: 10.1038/s41593-021-00969-4. URL https://www.nature.com/articles/s41593-021-00969-4. Number: 1 Publisher: Nature Publishing Group.

[77] Kathleen Grabert, Tom Michoel, Michail H. Karavolos, Sara Clohisey, J. Kenneth Baillie, Mark P. Stevens, Tom C. Freeman, Kim M. Summers, and Barry W. McColl. Microglial brain region-dependent diversity and selective regional sensitivities to aging. *Nature Neuroscience*, 19(3):504–516, March 2016. ISSN 1546-1726. doi: 10.1038/nn.4222.

[78] Charles E. Grant, Timothy L. Bailey, and William Stafford Noble. FIMO: scanning for occurrences of a given motif. *Bioinformatics*, 27(7):1017–1018, April 2011. ISSN 1367-4803. doi: 10.1093/bioinformatics/btr064. URL https://doi.org/10.1093/bioinformatics/btr064.

[79] Melissa Haendel, Nicole Vasilevsky, Deepak Unni, Cristian Bologa, Nomi Harris, Heidi Rehm, Ada Hamosh, Gareth Baynam, Tudor Groza, Julie McMurry, Hugh Dawkins, Ana Rath, Courtney Thaxon, Giovanni Bocci, Marcin P. Joachimiak, Sebastian Köhler, Peter N. Robinson, Chris Mungall, and Tudor I. Oprea. How many rare diseases are there? *Nature reviews. Drug discovery*, 19(2):77–78, February 2020. ISSN 1474-1776. doi: 10.1038/d41573-019-00180-y. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7771654/.

[80] Christoph Hafemeister and Rahul Satija. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biology*, 20(1):296, December 2019. ISSN 1474-760X. doi: 10.1186/s13059-019-1874-1. URL https://doi.org/10.1186/s13059-019-1874-1.

[81] Zhou Han, Chunling Chen, Anne Christiansen, Sophina Ji, Qian Lin, Charles Anumonwo, Chante Liu, Steven C. Leiser, null Meena, Isabel Aznarez, Gene Liau, and Lori L. Isom. Antisense oligonucleotides increase Scn1a expression and reduce seizures and SUDEP incidence in a mouse model of Dravet syndrome. *Science Translational Medicine*, 12(558):eaaz6100, August 2020. ISSN 1946-6242. doi: 10.1126/scitranslmed.aaz6100.

[82] Thomas V. O. Hansen, Jens F. Rehfeld, and Finn C. Nielsen. KCl and forskolin synergistically up-regulate cholecystokinin gene expression via coordinate activation of CREB and the co-activator CBP. *Journal of Neurochemistry*, 89(1):15–23, 2004. ISSN 1471-4159. doi: 10.1046/j.1471-4159.2003.02252.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1471-4159.2003.02252.x. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1046/j.1471-4159.2003.02252.x.

[83] Shane M. Harding, Joseph L. Benci, Jerome Irianto, Dennis E. Discher, Andy J. Minn, and Roger A. Greenberg. Mitotic progression following DNA damage enables pattern recognition within micronuclei. *Nature*, 548(7668):466–470, August 2017. ISSN 1476-4687. doi: 10.1038/nature23470. URL https://www.nature.com/articles/nature23470. Number: 7668 Publisher: Nature Publishing Group.

[84] Tamar Harel, Gozde Yesil, Yavuz Bayram, Zeynep Coban-Akdemir, Wu-Lin Charng, Ender Karaca, Ali Al Asmari, Mohammad K. Eldomery, Jill V. Hunter, Shalini N. Jhangiani, Jill A. Rosenfeld, Davut Pehlivan, Ayman W. El-Hattab, Mohammed A. Saleh, Charles A. LeDuc, Donna Muzny, Eric Boerwinkle, Baylor-Hopkins Center for Mendelian Genomics, Richard A. Gibbs, Wendy K. Chung, Yaping Yang, John W. Belmont, and James R. Lupski. Monoallelic and Biallelic Variants in EMC1 Identified in Individuals with Global Developmental Delay, Hypotonia, Scoliosis, and Cerebellar Atrophy. *American Journal of Human Genetics*, 98(3):562–570, March 2016. ISSN 1537-6605. doi: 10.1016/j.ajhg.2016.01.011.

[85] Holly K. Harris, Tojo Nakayama, Jenny Lai, Boxun Zhao, Nikoleta Argyrou, Cynthia S. Gubbels, Aubrie Soucy, Casie A. Genetti, Victoria Suslovitch, Lance H. Rodan, George E.

Tiller, Gaetan Lesca, Karen W. Gripp, Reza Asadollahi, Ada Hamosh, Carolyn D. Applegate, Peter D. Turnpenny, Marleen E. H. Simon, Catharina M. L. Volker-Touw, Koen L. I. van Gassen, Ellen van Binsbergen, Rolph Pfundt, Thatjana Gardeitchik, Bert B. A. de Vries, LaDonna L. Immken, Catherine Buchanan, Marcia Willing, Tomi L. Toler, Emily Fassi, Laura Baker, Fleur Vansenne, Xiadong Wang, Julian L. Ambrus, Madeleine Fannemel, Jennifer E. Posey, Emanuele Agolini, Antonio Novelli, Anita Rauch, Paranchai Boonsawat, Christina R. Fagerberg, Martin J. Larsen, Maria Kibaek, Audrey Labalme, Alice Poisson, Katelyn K. Payne, Laurence E. Walsh, Kimberly A. Aldinger, Jorune Balciuniene, Cara Skraban, Christopher Gray, Jill Murrell, Caleb P. Bupp, Giulia Pascolini, Paola Grammatico, Martin Broly, Sébastien Küry, Mathilde Nizon, Iqra Ghulam Rasool, Muhammad Yasir Zahoor, Cornelia Kraus, André Reis, Muhammad Iqbal, Kevin Uguen, Severine Audebert-Bellanger, Claude Ferec, Sylvia Redon, Janice Baker, Yunhong Wu, Guiseppe Zampino, Steffan Syrbe, Ines Brosse, Rami Abou Jamra, William B. Dobyns, Lilian L. Cohen, Anne Blomhoff, Cyril Mignot, Boris Keren, Thomas Courtin, Pankaj B. Agrawal, Alan H. Beggs, and Timothy W. Yu. Disruption of RFX family transcription factors causes autism, attention-deficit/hyperactivity disorder, intellectual disability, and dysregulated behavior. *Genetics in Medicine: Official Journal of the American College of Medical Genetics*, 23(6): 1028–1040, June 2021. ISSN 1530-0366. doi: 10.1038/s41436-021-01114-z.

[86] Joshua G. Harrison, W. John Calder, Vivaswat Shastry, and C. Alex Buerkle. Dirichlet-multinomial modelling outperforms alternatives for analysis of microbiome and other ecological count data. *Molecular Ecology Resources*, 20(2):481–497, 2020. ISSN 1755-0998. doi: 10.1111/1755-0998.13128. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/1755-0998.13128. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1755-0998.13128.

[87] Gaiti Hasan and Anamika Sharma. Regulation of neuronal physiology by Ca2+ release through the IP3R. *Current Opinion in Physiology*, 17:1–8, October 2020. ISSN 2468-8673. doi: 10.1016/j.cophys.2020.06.001. URL https://www.sciencedirect.com/science/article/pii/S246886732030050X.

[88] Kerstin Hasenpusch-Theil, Christine Laclef, Matt Colligan, Eamon Fitzgerald, Katherine Howe, Emily Carroll, Shaun R Abrams, Jeremy F Reiter, Sylvie Schneider-Maunoury, and Thomas Theil. A transient role of the ciliary gene Inpp5e in controlling direct versus indirect neurogenesis in cortical development. *eLife*, 9:e58162, August 2020. ISSN 2050-084X. doi: 10.7554/eLife.58162. URL https://doi.org/10.7554/eLife.58162. Publisher: eLife Sciences Publications, Ltd.

[89] Sven Heinz, Christopher Benner, Nathanael Spann, Eric Bertolino, Yin C. Lin, Peter Laslo, Jason X. Cheng, Cornelis Murre, Harinder Singh, and Christopher K. Glass. Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Molecular Cell*, 38(4):576–

589, May 2010. ISSN 1097-2765. doi: 10.1016/j.molcel.2010.05.004. URL https://www.sciencedirect.com/science/article/pii/S1097276510003667.

[90] A. Hemmati-Brivanlou and D. A. Melton. A truncated activin receptor inhibits mesoderm induction and formation of axial structures in Xenopus embryos. *Nature*, 359(6396):609–614, October 1992. ISSN 0028-0836. doi: 10.1038/359609a0.

[91] A. Hemmati-Brivanlou and D. A. Melton. Inhibition of activin receptor signaling promotes neuralization in Xenopus. *Cell*, 77(2):273–281, April 1994. ISSN 0092-8674. doi: 10.1016/0092-8674(94)90319-0.

[92] A. Hemmati-Brivanlou, O. G. Kelly, and D. A. Melton. Follistatin, an antagonist of activin, is expressed in the Spemann organizer and displays direct neuralizing activity. *Cell*, 77(2):283–295, April 1994. ISSN 0092-8674. doi: 10.1016/0092-8674(94)90320-4.

[93] Richard A. Hickman, Sarah A. O'Shea, Mark F. Mehler, and Wendy K. Chung. Neurogenetic disorders across the lifespan: from aberrant development to degeneration. *Nature Reviews Neurology*, 18(2):117–124, February 2022. ISSN 1759-4766. doi: 10.1038/s41582-021-00595-5. URL https://www.nature.com/articles/s41582-021-00595-5. Number: 2 Publisher: Nature Publishing Group.

[94] Suzanne Hickman, Saef Izzy, Pritha Sen, Liza Morsett, and Joseph El Khoury. Microglia in neurodegeneration. *Nature Neuroscience*, 21(10):1359–1369, October 2018. ISSN 1546-1726. doi: 10.1038/s41593-018-0242-x. URL https://www.nature.com/articles/s41593-018-0242-x. Number: 10 Publisher: Nature Publishing Group.

[95] Philip Hieter and Mark Boguski. Functional Genomics: It's All How You Read It. *Science*, 278(5338):601–602, October 1997. doi: 10.1126/science.278.5338.601. URL https://www.science.org/doi/abs/10.1126/science.278.5338.601. Publisher: American Association for the Advancement of Science.

[96] Chihiro Hisatsune, Hiroyuki Miyamoto, Moritoshi Hirono, Naohide Yamaguchi, Takeyuki Sugawara, Naoko Ogawa, Etsuko Ebisui, Toshio Ohshima, Masahisa Yamada, Takao Hensch, Mitsuharu Hattori, and Katsuhiko Mikoshiba. IP3R1 deficiency in the cerebellum/brainstem causes basal ganglia-independent dystonia by triggering tonic Purkinje cell firings in mice. *Frontiers in Neural Circuits*, 7, 2013. ISSN 1662-5110. URL https://www.frontiersin.org/articles/10.3389/fncir.2013.00156.

[97] Christy Hong, Andrea E. Tijhuis, and Floris Foijer. The cGAS Paradox: Contrasting Roles for cGAS-STING Pathway in Chromosomal Instability. *Cells*, 8(10):1228, October 2019. ISSN 2073-4409. doi: 10.3390/cells8101228. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6830079/.

[98] Soyon Hong, Victoria F. Beja-Glasser, Bianca M. Nfonoyim, Arnaud Frouin, Shaomin Li, Saranya Ramakrishnan, Katherine M. Merry, Qiaoqiao Shi, Arnon Rosenthal, Ben A. Barres, Cynthia A. Lemere, Dennis J. Selkoe, and Beth Stevens. Complement and microglia mediate early synapse loss in Alzheimer mouse models. *Science (New York, N.Y.)*, 352(6286):712–716, May 2016. ISSN 1095-9203. doi: 10.1126/science.aad8373.

[99] Sinisa Hrvatin, Daniel R. Hochbaum, M. Aurel Nagy, Marcelo Cicconet, Keiramarie Robertson, Lucas Cheadle, Rapolas Zilionis, Alex Ratner, Rebeca Borges-Monroy, Allon M. Klein, Bernardo L. Sabatini, and Michael E. Greenberg. Single-Cell Analysis of Experience-Dependent Transcriptomic States in Mouse Visual Cortex. *Nature neuroscience*, 21(1):120–129, January 2018. ISSN 1097-6256. doi: 10.1038/s41593-017-0029-5. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5742025/.

[100] Bao-Yang Hu and Su-Chun Zhang. Differentiation of spinal motor neurons from pluripotent human stem cells. *Nature Protocols*, 4(9):1295–1304, September 2009. ISSN 1750-2799. doi: 10.1038/nprot.2009.127. URL https://www.nature.com/articles/nprot.2009.127. Number: 9 Publisher: Nature Publishing Group.

[101] Ming-Ming Hu and Hong-Bing Shu. Multifaceted roles of TRIM38 in innate immune and inflammatory responses. *Cellular and Molecular Immunology*, 14(4):331–338, April 2017. ISSN 1672-7681. doi: 10.1038/cmi.2016.66. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5380946/.

[102] Yimin Hua, Timothy A. Vickers, Hazeem L. Okunola, C. Frank Bennett, and Adrian R. Krainer. Antisense Masking of an hnRNP A1/A2 Intronic Splicing Silencer Corrects SMN2 Splicing in Transgenic Mice. *American Journal of Human Genetics*, 82(4):834–848, April 2008. ISSN 0002-9297. doi: 10.1016/j.ajhg.2008.01.014. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2427210/.

[103] Yimin Hua, Kentaro Sahashi, Gene Hung, Frank Rigo, Marco A. Passini, C. Frank Bennett, and Adrian R. Krainer. Antisense correction of SMN2 splicing in the CNS rescues necrosis in a type III SMA mouse model. *Genes & Development*, 24(15):1634–1644, August 2010. ISSN 1549-5477. doi: 10.1101/gad.1941310.

[104] Ching Ying Huang, Martin W. Nicholson, Jyun Yuan Wang, Chien Yu Ting, Ming Heng Tsai, Yu Che Cheng, Chun Lin Liu, Darien Z. H. Chan, Yi Chan Lee, Ching Chuan Hsu, Yu Hung Hsu, Chiou Fong Yang, Cindy M. C. Chang, Shu Chian Ruan, Po Ju Lin, Jen Hao Lin, Li Lun Chen, Marvin L. Hsieh, Yuan Yuan Cheng, Wan Tseng Hsu, Yi Ling Lin, Chien Hsiun Chen, Yu Hsiang Hsu, Ying Ta Wu, Timothy A. Hacker, Joseph C. Wu, Timothy J. Kamp, and Patrick C. H. Hsieh. Population-based high-throughput toxicity screen of human iPSC-derived cardiomyocytes and neurons. *Cell Reports*, 39(1):110643, April 2022. ISSN 2211-1247. doi: 10.1016/j.celrep.2022.110643.

[105] Lijia Huang, Jodi Warman Chardon, Melissa T. Carter, Kathie L. Friend, Tracy E. Dudding, Jeremy Schwartzentruber, Ruobing Zou, Peter W. Schofield, Stuart Douglas, Dennis E. Bulman, and Kym M. Boycott. Missense mutations in ITPR1 cause autosomal dominant congenital nonprogressive spinocerebellar ataxia. *Orphanet Journal of Rare Diseases*, 7:67, September 2012. ISSN 1750-1172. doi: 10.1186/1750-1172-7-67.

[106] Wei-Kai Huang, Samuel Zheng Hao Wong, Sarshan R. Pather, Phuong T. T. Nguyen, Feng Zhang, Daniel Y. Zhang, Zhijian Zhang, Lu Lu, Wanqi Fang, Luyun Chen, Analiese Fernandes, Yijing Su, Hongjun Song, and Guo-Li Ming. Generation of hypothalamic arcuate organoids from human induced pluripotent stem cells. *Cell Stem Cell*, 28(9):1657–1670.e10, September 2021. ISSN 1875-9777. doi: 10.1016/j.stem.2021.04.006.

[107] Chin Wai Hui, Xuan Song, Fulin Ma, Xuting Shen, and Karl Herrup. Ibuprofen prevents progression of ataxia telangiectasia symptoms in ATM-deficient mice. *Journal of Neuroinflammation*, 15(1):308, November 2018. ISSN 1742-2094. doi: 10.1186/s12974-018-1338-7. URL https://doi.org/10.1186/s12974-018-1338-7.

[108] Jeffrey J. Hutsler and Hong Zhang. Increased dendritic spine densities on cortical projection neurons in autism spectrum disorders. *Brain Research*, 1309:83–94, January 2010. ISSN 1872-6240. doi: 10.1016/j.brainres.2009.09.120.

[109] Anetta Härtlova, Saskia F. Erttmann, Faizal Am Raffi, Anja M. Schmalz, Ulrike Resch, Sharath Anugula, Stefan Lienenklaus, Lisa M. Nilsson, Andrea Kröger, Jonas A. Nilsson, Torben Ek, Siegfried Weiss, and Nelson O. Gekara. DNA damage primes the type I interferon system via the cytosolic DNA sensor STING to promote anti-microbial innate immunity. *Immunity*, 42(2):332–343, February 2015. ISSN 1097-4180. doi: 10.1016/j.immuni.2015.01.012.

[110] Mason A. Israel, Shauna H. Yuan, Cedric Bardy, Sol M. Reyna, Yangling Mu, Cheryl Herrera, Michael P. Hefferan, Sebastiaan Van Gorp, Kristopher L. Nazor, Francesca S. Boscolo, Christian T. Carson, Louise C. Laurent, Martin Marsala, Fred H. Gage, Anne M. Remes, Edward H. Koo, and Lawrence S. B. Goldstein. Probing sporadic and familial Alzheimer's disease using induced pluripotent stem cells. *Nature*, 482(7384):216–220, January 2012. ISSN 1476-4687. doi: 10.1038/nature10821.

[111] Paymaan Jafar-Nejad, Berit Powers, Armand Soriano, Hien Zhao, Daniel A. Norris, John Matson, Beatrice DeBrosse-Serra, Jamie Watson, Padmakumar Narayanan, Seung J. Chun, Curt Mazur, Holly Kordasiewicz, Eric E. Swayze, and Frank Rigo. The atlas of RNase H antisense oligonucleotide distribution and activity in the CNS of rodents and non-human primates following central administration. *Nucleic Acids Research*, 49(2):657–673, January 2021. ISSN 1362-4962. doi: 10.1093/nar/gkaa1235.

[112] Mehrnoosh Jafari, Adrian-Minh Schumacher, Nicolas Snaidero, Emily M. Ullrich Gavilanes, Tradite Neziraj, Virág Kocsis-Jutka, Daniel Engels, Tanja Jürgens, Ingrid Wagner,

Juan Daniel Flórez Weidinger, Stephanie S. Schmidt, Eduardo Beltrán, Nellwyn Hagan, Lisa Woodworth, Dimitry Ofengeim, Joseph Gans, Fred Wolf, Mario Kreutzfeldt, Ruben Portugues, Doron Merkler, Thomas Misgeld, and Martin Kerschensteiner. Phagocyte-mediated synapse removal in cortical neuroinflammation is promoted by local calcium accumulation. *Nature Neuroscience*, 24(3):355–367, March 2021. ISSN 1546-1726. doi: 10.1038/s41593-020-00780-7.

[113] Dewei Jiang, Ying Zhang, Ronald P. Hart, Jianmin Chen, Karl Herrup, and Jiali Li. Alteration in 5-hydroxymethylcytosine-mediated epigenetic regulation leads to Purkinje cell vulnerability in ATM deficiency. *Brain*, 138(12):3520–3536, December 2015. ISSN 0006-8950. doi: 10.1093/brain/awv284. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4668921/.

[114] Suoqin Jin, Christian F. Guerrero-Juarez, Lihua Zhang, Ivan Chang, Raul Ramos, Chen-Hsiang Kuan, Peggy Myung, Maksim V. Plikus, and Qing Nie. Inference and analysis of cell-cell communication using CellChat. *Nature Communications*, 12(1):1088, February 2021. ISSN 2041-1723. doi: 10.1038/s41467-021-21246-9. URL https://www.nature.com/articles/s41467-021-21246-9. Number: 1 Publisher: Nature Publishing Group.

[115] Xuemei Jin and Toshihide Yamashita. Microglia in central nervous system repair after injury. *The Journal of Biochemistry*, 159(5):491–496, May 2016. ISSN 0021-924X. doi: 10.1093/jb/mvw009. URL https://doi.org/10.1093/jb/mvw009.

[116] David S. Johnson, Ali Mortazavi, Richard M. Myers, and Barbara Wold. Genome-wide mapping of in vivo protein-DNA interactions. *Science (New York, N.Y.)*, 316(5830):1497–1502, June 2007. ISSN 1095-9203. doi: 10.1126/science.1141319.

[117] Mathias Jucker. The benefits and limitations of animal models for translational research in neurodegenerative diseases. *Nature Medicine*, 16(11):1210–1214, November 2010. ISSN 1546-170X. doi: 10.1038/nm.2224. URL https://www.nature.com/articles/nm.2224. Number: 11 Publisher: Nature Publishing Group.

[118] Sarah Jäkel, Eneritz Agirre, Ana Mendanha Falcão, David van Bruggen, Ka Wai Lee, Irene Knuesel, Dheeraj Malhotra, Charles Ffrench-Constant, Anna Williams, and Gonçalo Castelo-Branco. Altered human oligodendrocyte heterogeneity in multiple sclerosis. *Nature*, 566(7745):543–547, February 2019. ISSN 1476-4687. doi: 10.1038/s41586-019-0903-2.

[119] Hyo Jung Kang, Yuka Imamura Kawasawa, Feng Cheng, Ying Zhu, Xuming Xu, Mingfeng Li, André M. M. Sousa, Mihovil Pletikos, Kyle A. Meyer, Goran Sedmak, Tobias Guennel, Yurae Shin, Matthew B. Johnson, Zeljka Krsnik, Simone Mayer, Sofia Fertuzinhos, Sheila Umlauf, Steven N. Lisgo, Alexander Vortmeyer, Daniel R. Weinberger, Shrikant Mane, Thomas M. Hyde, Anita Huttner, Mark Reimers, Joel E. Kleinman, and Nenad Sestan. Spatio-temporal transcriptome of the human brain. *Nature*, 478(7370):483–489, October 2011. ISSN 1476-4687. doi: 10.1038/nature10523.

[120] Konrad J. Karczewski, Laurent C. Francioli, Grace Tiao, Beryl B. Cummings, Jessica Alföldi, Qingbo Wang, Ryan L. Collins, Kristen M. Laricchia, Andrea Ganna, Daniel P. Birnbaum, Laura D. Gauthier, Harrison Brand, Matthew Solomonson, Nicholas A. Watts, Daniel Rhodes, Moriel Singer-Berk, Eleina M. England, Eleanor G. Seaby, Jack A. Kosmicki, Raymond K. Walters, Katherine Tashman, Yossi Farjoun, Eric Banks, Timothy Poterba, Arcturus Wang, Cotton Seed, Nicola Whiffin, Jessica X. Chong, Kaitlin E. Samocha, Emma Pierce-Hoffman, Zachary Zappala, Anne H. O'Donnell-Luria, Eric Vallabh Minikel, Ben Weisburd, Monkol Lek, James S. Ware, Christopher Vittal, Irina M. Armean, Louis Bergelson, Kristian Cibulskis, Kristen M. Connolly, Miguel Covarrubias, Stacey Donnelly, Steven Ferriera, Stacey Gabriel, Jeff Gentry, Namrata Gupta, Thibault Jeandet, Diane Kaplan, Christopher Llanwarne, Ruchi Munshi, Sam Novod, Nikelle Petrillo, David Roazen, Valentin Ruano-Rubio, Andrea Saltzman, Molly Schleicher, Jose Soto, Kathleen Tibbetts, Charlotte Tolonen, Gordon Wade, Michael E. Talkowski, Benjamin M. Neale, Mark J. Daly, and Daniel G. MacArthur. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*, 581(7809):434–443, May 2020. ISSN 1476-4687. doi: 10.1038/s41586-020-2308-7. URL https://www.nature.com/articles/s41586-020-2308-7. Number: 7809 Publisher: Nature Publishing Group.

[121] Joyce Keifer and Cliff H. Summers. Putting the "Biology" Back into "Neurobiology": The Strength of Diversity in Animal Model Systems for Neuroscience Research. *Frontiers in Systems Neuroscience*, 10, 2016. ISSN 1662-5137. URL https://www.frontiersin.org/articles/10.3389/fnsys.2016.00069.

[122] W. James Kent, Charles W. Sugnet, Terrence S. Furey, Krishna M. Roskin, Tom H. Pringle, Alan M. Zahler, and and David Haussler. The Human Genome Browser at UCSC. *Genome Research*, 12(6):996–1006, June 2002. ISSN 1088-9051, 1549-5469. doi: 10.1101/gr.229102. URL https://genome.cshlp.org/content/12/6/996. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab.

[123] Hadas Keren-Shaul, Amit Spinrad, Assaf Weiner, Orit Matcovitch-Natan, Raz Dvir-Szternfeld, Tyler K. Ulland, Eyal David, Kuti Baruch, David Lara-Astaiso, Beata Toth, Shalev Itzkovitz, Marco Colonna, Michal Schwartz, and Ido Amit. A Unique Microglia Type Associated with Restricting Development of Alzheimer's Disease. *Cell*, 169(7):1276–1290.e17, June 2017. ISSN 1097-4172. doi: 10.1016/j.cell.2017.05.018.

[124] Cari F. Kessing and William R. Tyor. Interferon-α Induces Neurotoxicity Through Activation of the Type I Receptor and the GluN2A Subunit of the NMDA Receptor. *Journal of Interferon & Cytokine Research*, 35(4):317–324, April 2015. ISSN 1079-9907. doi: 10.1089/jir.2014.0105. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4389917/.

[125] Rustem Khazipov and Heiko J. Luhmann. Early patterns of electrical activity in the developing cerebral cortex of humans and rodents. *Trends in Neurosciences*, 29(7):414–418, July 2006. ISSN 0166-2236. doi: 10.1016/j.tins.2006.05.007.

[126] Daniel Seung Kim, Amber A. Burt, Jane E. Ranchalis, Beth Wilmot, Joshua D. Smith, Karynne E. Patterson, Bradley P. Coe, Yatong K. Li, Michael J. Bamshad, Molly Nikolas, Evan E. Eichler, James M. Swanson, Joel T. Nigg, Deborah A. Nickerson, Gail P. Jarvik, and University of Washington Center for Mendelian Genomics. Sequencing of sporadic Attention-Deficit Hyperactivity Disorder (ADHD) identifies novel and potentially pathogenic de novo variants and excludes overlap with genes associated with autism spectrum disorder. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics: The Official Publication of the International Society of Psychiatric Genetics*, 174(4):381–389, June 2017. ISSN 1552-485X. doi: 10.1002/ajmg.b.32527.

[127] Jinkuk Kim, Chunguang Hu, Christelle Moufawad El Achkar, Lauren E. Black, Julie Douville, Austin Larson, Mary K. Pendergast, Sara F. Goldkind, Eunjung A. Lee, Ashley Kuniholm, Aubrie Soucy, Jai Vaze, Nandkishore R. Belur, Kristina Fredriksen, Iva Stojkovska, Alla Tsytsykova, Myriam Armant, Renata L. DiDonato, Jaejoon Choi, Laura Cornelissen, Luis M. Pereira, Erika F. Augustine, Casie A. Genetti, Kira Dies, Brenda Barton, Lucinda Williams, Benjamin D. Goodlett, Bobbie L. Riley, Amy Pasternak, Emily R. Berry, Kelly A. Pflock, Stephen Chu, Chantal Reed, Kimberly Tyndall, Pankaj B. Agrawal, Alan H. Beggs, P. Ellen Grant, David K. Urion, Richard O. Snyder, Susan E. Waisbren, Annapurna Poduri, Peter J. Park, Al Patterson, Alessandra Biffi, Joseph R. Mazzulli, Olaf Bodamer, Charles B. Berde, and Timothy W. Yu. Patient-Customized Oligonucleotide Therapy for a Rare Genetic Disease. *New England Journal of Medicine*, 381(17): 1644–1652, October 2019. ISSN 0028-4793. doi: 10.1056/NEJMoa1813279. URL https://doi.org/10.1056/NEJMoa1813279. Publisher: Massachusetts Medical Society _eprint: https://doi.org/10.1056/NEJMoa1813279.

[128] Jinkuk Kim, Sijae Woo, Claudio M. de Gusmao, Boxun Zhao, Diana H. Chin, Renata L. DiDonato, Minh A. Nguyen, Tojo Nakayama, Chunguang April Hu, Aubrie Soucy, Ashley Kuniholm, Jennifer Karlin Thornton, Olivia Riccardi, Danielle A. Friedman, Christelle Moufawad El Achkar, Zane Dash, Laura Cornelissen, Carolina Donado, Kamli N. W. Faour, Lynn W. Bush, Victoria Suslovitch, Claudia Lentucci, Peter J. Park, Eunjung Alice Lee, Al Patterson, Anthony A. Philippakis, Brad Margus, Charles B. Berde, and Timothy W. Yu. A framework for individualized splice-switching oligonucleotide therapy. *Nature*, 619 (7971):828–836, July 2023. ISSN 1476-4687. doi: 10.1038/s41586-023-06277-0. URL https://www.nature.com/articles/s41586-023-06277-0. Number: 7971 Publisher: Nature Publishing Group.

[129] Kun-Yong Kim, Eriona Hysolli, and In-Hyun Park. Neuronal maturation defect in induced pluripotent stem cells from patients with Rett syndrome. *Proceedings of the National*

Academy of Sciences of the United States of America, 108(34):14169–14174, August 2011. ISSN 0027-8424. doi: 10.1073/pnas.1018979108. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3161557/.

[130] Jason C. Klein, Aidan Keith, Vikram Agarwal, Timothy Durham, and Jay Shendure. Functional characterization of enhancer evolution in the primate lineage. *Genome Biology*, 19(1):99, July 2018. ISSN 1474-760X. doi: 10.1186/s13059-018-1473-6. URL https://doi.org/10.1186/s13059-018-1473-6.

[131] Frank Koopmans, Pim van Nierop, Maria Andres-Alonso, Andrea Byrnes, Tony Cijsouw, Marcelo P. Coba, L. Niels Cornelisse, Ryan J. Farrell, Hana L. Goldschmidt, Daniel P. Howrigan, Natasha K. Hussain, Cordelia Imig, Arthur P. H. de Jong, Hwajin Jung, Mahdokht Kohansalnodehi, Barbara Kramarz, Noa Lipstein, Ruth C. Lovering, Harold MacGillavry, Vittoria Mariano, Huaiyu Mi, Momchil Ninov, David Osumi-Sutherland, Rainer Pielot, Karl-Heinz Smalla, Haiming Tang, Katherine Tashman, Ruud F. G. Toonen, Chiara Verpelli, Rita Reig-Viader, Kyoko Watanabe, Jan van Weering, Tilmann Achsel, Ghazaleh Ashrafi, Nimra Asi, Tyler C. Brown, Pietro De Camilli, Marc Feuermann, Rebecca E. Foulger, Pascale Gaudet, Anoushka Joglekar, Alexandros Kanellopoulos, Robert Malenka, Roger A. Nicoll, Camila Pulido, Jaime de Juan-Sanz, Morgan Sheng, Thomas C. Südhof, Hagen U. Tilgner, Claudia Bagni, Àlex Bayés, Thomas Biederer, Nils Brose, John Jia En Chua, Daniela C. Dieterich, Eckart D. Gundelfinger, Casper Hoogenraad, Richard L. Huganir, Reinhard Jahn, Pascal S. Kaeser, Eunjoon Kim, Michael R. Kreutz, Peter S. McPherson, Ben M. Neale, Vincent O'Connor, Danielle Posthuma, Timothy A. Ryan, Carlo Sala, Guoping Feng, Steven E. Hyman, Paul D. Thomas, August B. Smit, and Matthijs Verhage. SynGO: An Evidence-Based, Expert-Curated Knowledge Base for the Synapse. *Neuron*, 103(2):217–234.e4, July 2019. ISSN 0896-6273. doi: 10.1016/j.neuron.2019.05.002. URL https://www.sciencedirect.com/science/article/pii/S0896627319304271.

[132] L. Korbo and B. B. Andersen. The distributions of Purkinje cell perikaryon and nuclear volume in human and rat cerebellum with the nucleator method. *Neuroscience*, 69(1):151–158, November 1995. ISSN 0306-4522. doi: 10.1016/0306-4522(95)00223-6. URL https://www.sciencedirect.com/science/article/pii/0306452295002236.

[133] Vladislav A. Korobeynikov, Alexander K. Lyashchenko, Beatriz Blanco-Redondo, Paymaan Jafar-Nejad, and Neil A. Shneider. Antisense oligonucleotide silencing of FUS expression as a therapeutic approach in amyotrophic lateral sclerosis. *Nature Medicine*, 28(1):104–116, January 2022. ISSN 1546-170X. doi: 10.1038/s41591-021-01615-z. URL https://www.nature.com/articles/s41591-021-01615-z. Number: 1 Publisher: Nature Publishing Group.

[134] Velina Kozareva, Caroline Martin, Tomas Osorno, Stephanie Rudolph, Chong Guo, Charles Vanderburg, Naeem Nadaf, Aviv Regev, Wade G. Regehr, and Evan Macosko. A transcriptomic atlas of mouse cerebellar cortex comprehensively defines cell types. *Nature*, 598(7879):

214–219, October 2021. ISSN 1476-4687. doi: 10.1038/s41586-021-03220-z. URL https://www.nature.com/articles/s41586-021-03220-z. Number: 7879 Publisher: Nature Publishing Group.

[135] Sonja Kriks, Jae-Won Shim, Jinghua Piao, Yosif M. Ganat, Dustin R. Wakeman, Zhong Xie, Luis Carrillo-Reid, Gordon Auyeung, Chris Antonacci, Amanda Buch, Lichuan Yang, M. Flint Beal, D. James Surmeier, Jeffrey H. Kordower, Viviane Tabar, and Lorenz Studer. Dopamine neurons derived from human ES cells efficiently engraft in animal models of Parkinson's disease. *Nature*, 480(7378):547–551, November 2011. ISSN 1476-4687. doi: 10.1038/nature10648.

[136] Niklas Krumm, Tychele N. Turner, Carl Baker, Laura Vives, Kiana Mohajeri, Kali Witherspoon, Archana Raja, Bradley P. Coe, Holly A. Stessman, Zong-Xiao He, Suzanne M. Leal, Raphael Bernier, and Evan E. Eichler. Excess of rare, inherited truncating mutations in autism. *Nature Genetics*, 47(6):582–588, June 2015. ISSN 1546-1718. doi: 10.1038/ng.3303.

[137] Blue B. Lake, Rizi Ai, Gwendolyn E. Kaeser, Neeraj S. Salathia, Yun C. Yung, Rui Liu, Andre Wildberg, Derek Gao, Ho-Lim Fung, Song Chen, Raakhee Vijayaraghavan, Julian Wong, Allison Chen, Xiaoyan Sheng, Fiona Kaper, Richard Shen, Mostafa Ronaghi, Jian-Bing Fan, Wei Wang, Jerold Chun, and Kun Zhang. Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science (New York, N.Y.)*, 352(6293):1586–1590, June 2016. ISSN 1095-9203. doi: 10.1126/science.aaf1204.

[138] Blue B. Lake, Simone Codeluppi, Yun C. Yung, Derek Gao, Jerold Chun, Peter V. Kharchenko, Sten Linnarsson, and Kun Zhang. A comparative strategy for single-nucleus and single-cell transcriptomes confirms accuracy in predicted cell-type expression from nuclear RNA. *Scientific Reports*, 7(1):6031, July 2017. ISSN 2045-2322. doi: 10.1038/s41598-017-04426-w.

[139] Blue B. Lake, Song Chen, Brandon C. Sos, Jean Fan, Gwendolyn E. Kaeser, Yun C. Yung, Thu E. Duong, Derek Gao, Jerold Chun, Peter V. Kharchenko, and Kun Zhang. Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain. *Nature Biotechnology*, 36(1):70–80, January 2018. ISSN 1546-1696. doi: 10.1038/nbt.4038. URL https://www.nature.com/articles/nbt.4038. Number: 1 Publisher: Nature Publishing Group.

[140] Madeline A. Lancaster, Magdalena Renner, Carol-Anne Martin, Daniel Wenzel, Louise S. Bicknell, Matthew E. Hurles, Tessa Homfray, Josef M. Penninger, Andrew P. Jackson, and Juergen A. Knoblich. Cerebral organoids model human brain development and microcephaly. *Nature*, 501(7467):373–379, September 2013. ISSN 1476-4687. doi: 10.1038/nature12517. URL https://www.nature.com/articles/nature12517. Number: 7467 Publisher: Nature Publishing Group.

[141] Melissa J. Landrum, Jennifer M. Lee, Mark Benson, Garth R. Brown, Chen Chao, Shanmuga Chitipiralla, Baoshan Gu, Jennifer Hart, Douglas Hoffman, Wonhee Jang, Karen Karapetyan, Kenneth Katz, Chunlei Liu, Zenith Maddipatla, Adriana Malheiro, Kurt McDaniel, Michael Ovetsky, George Riley, George Zhou, J. Bradley Holmes, Brandi L. Kattman, and Donna R. Maglott. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Research*, 46(D1):D1062–D1067, January 2018. ISSN 1362-4962. doi: 10.1093/nar/gkx1153.

[142] Ben Langmead and Steven L. Salzberg. Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9(4):357–359, April 2012. ISSN 1548-7105. doi: 10.1038/nmeth.1923. URL https://www.nature.com/articles/nmeth.1923. Number: 4 Publisher: Nature Publishing Group.

[143] M. F. Lavin and Y. Shiloh. The genetic defect in ataxia-telangiectasia. *Annual Review of Immunology*, 15:177–202, 1997. ISSN 0732-0582. doi: 10.1146/annurev.immunol.15.1.177.

[144] Martin F. Lavin. The appropriateness of the mouse model for ataxia-telangiectasia: neurological defects but no neurodegeneration. *DNA repair*, 12(8):612–619, August 2013. ISSN 1568-7856. doi: 10.1016/j.dnarep.2013.04.014.

[145] Hyeseung Lee, Robert J. Fenster, S. Sebastian Pineda, Whitney S. Gibbs, Shahin Mohammadi, Jose Davila-Velderrain, Francisco J. Garcia, Martine Therrien, Hailey S. Novis, Fan Gao, Hilary Wilkinson, Thomas Vogt, Manolis Kellis, Matthew J. LaVoie, and Myriam Heiman. Cell Type-Specific Transcriptomics Reveals that Mutant Huntingtin Leads to Mitochondrial RNA Release and Neuronal Innate Immune Activation. *Neuron*, 107(5): 891–908.e8, September 2020. ISSN 1097-4199. doi: 10.1016/j.neuron.2020.06.021.

[146] Sylvain Lemeille, Marie Paschaki, Dominique Baas, Laurette Morlé, Jean-Luc Duteyrat, Aouatef Ait-Lounis, Emmanuèle Barras, Fabien Soulavie, Julie Jerber, Joëlle Thomas, Yong Zhang, Michael J Holtzman, W Stephen Kistler, Walter Reith, and Bénédicte Durand. Interplay of RFX transcription factors 1, 2 and 3 in motile ciliogenesis. *Nucleic Acids Research*, 48(16):9019–9036, July 2020. ISSN 0305-1048. doi: 10.1093/nar/gkaa625. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7498320/.

[147] Anastasia Levchenko, Alexander Kanapin, Anastasia Samsonova, and Raul R Gainetdinov. Human Accelerated Regions and Other Human-Specific Sequence Variations in the Context of Evolution and Their Relevance for Brain Development. *Genome Biology and Evolution*, 10(1):166–188, January 2018. ISSN 1759-6653. doi: 10.1093/gbe/evx240. URL https://doi.org/10.1093/gbe/evx240.

[148] Hadar Levi, Ela Bar, Stav Cohen-Adiv, Suzan Sweitat, Sivan Kanner, Ronit Galron, Yulia Mitiagin, and Ari Barzilai. Dysfunction of cerebellar microglia in Ataxia-telangiectasia. *Glia*, 70(3):536–557, March 2022. ISSN 1098-1136. doi: 10.1002/glia.24122.

[149] Jiali Li, Jianmin Chen, Christopher L. Ricupero, Ronald P. Hart, Melanie S. Schwartz, Alexander Kusnecov, and Karl Herrup. Nuclear accumulation of HDAC4 in ATM deficiency promotes neurodegeneration in ataxia telangiectasia. *Nature Medicine*, 18(5):783–790, May 2012. ISSN 1546-170X. doi: 10.1038/nm.2709.

[150] Jinchen Li, Lin Wang, Hui Guo, Leisheng Shi, Kun Zhang, Meina Tang, Shanshan Hu, Shanshan Dong, Yanling Liu, Tianyun Wang, Ping Yu, Xin He, Zhengmao Hu, Jinping Zhao, Chunyu Liu, Zhong Sheng Sun, and Kun Xia. Targeted sequencing and functional analysis reveal brain-size-related genes and their networks in autism spectrum disorders. *Molecular Psychiatry*, 22(9):1282–1290, September 2017. ISSN 1476-5578. doi: 10.1038/mp.2017.140.

[151] Jun Li, Ting Zhang, Aarthi Ramakrishnan, Bernd Fritzsch, Jinshu Xu, Elaine Y M Wong, Yong-Hwee Eddie Loh, Jianqiang Ding, Li Shen, and Pin-Xian Xu. Dynamic changes in cis-regulatory occupancy by Six1 and its cooperative interactions with distinct cofactors drive lineage-specific gene expression programs during progressive differentiation of the auditory sensory epithelium. *Nucleic Acids Research*, 48(6):2880–2896, April 2020. ISSN 0305-1048. doi: 10.1093/nar/gkaa012. URL https://doi.org/10.1093/nar/gkaa012.

[152] Qunhua Li, James B. Brown, Haiyan Huang, and Peter J. Bickel. Measuring Reproducibility of High-Throughput Experiments. *The Annals of Applied Statistics*, 5(3):1752–1779, 2011. ISSN 1932-6157. URL https://www.jstor.org/stable/23069353. Publisher: Institute of Mathematical Statistics.

[153] Xue-Jun Li, Xiaoqing Zhang, M. Austin Johnson, Zhi-Bo Wang, Timothy Lavaute, and Su-Chun Zhang. Coordination of sonic hedgehog and Wnt signaling determines ventral and dorsal telencephalic neuron types from human embryonic stem cells. *Development (Cambridge, England)*, 136(23):4055–4063, December 2009. ISSN 1477-9129. doi: 10.1242/dev.036624.

[154] Y. I. Li, D. A. Knowles, J. Humphrey, A. N. Barbeira, S. P. Dickinson, H. K. Im, and J. K. Pritchard. Annotation-free quantification of RNA splicing using LeafCutter. *Nat Genet*, 50(1):151–158, January 2018. ISSN 1546-1718 (Electronic) 1061-4036 (Linking). doi: 10.1038/s41588-017-0004-9. URL https://www.ncbi.nlm.nih.gov/pubmed/29229983. Edition: 2017/12/13.

[155] Yanjie Li, Jixue Li, Jun Wang, David R Lynch, Xiulong Shen, David R Corey, Darshan Parekh, Balkrishen Bhat, Caroline Woo, Jonathan J Cherry, Jill S Napierala, and Marek Napierala. Targeting 3′ and 5′ untranslated regions with antisense oligonucleotides to stabilize frataxin mRNA and increase protein expression. *Nucleic Acids Research*, 49(20):11560–11574, November 2021. ISSN 0305-1048. doi: 10.1093/nar/gkab954. URL https://doi.org/10.1093/nar/gkab954.

[156] Yang Liao, Gordon K. Smyth, and Wei Shi. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics (Oxford, England)*, 30(7):923–930, April 2014. ISSN 1367-4811. doi: 10.1093/bioinformatics/btt656.

[157] Arthur Liberzon, Aravind Subramanian, Reid Pinchback, Helga Thorvaldsdóttir, Pablo Tamayo, and Jill P. Mesirov. Molecular signatures database (MSigDB) 3.0. *Bioinformatics*, 27 (12):1739–1740, June 2011. ISSN 1367-4803. doi: 10.1093/bioinformatics/btr260. URL https://doi.org/10.1093/bioinformatics/btr260.

[158] Kian Huat Lim, Zhou Han, Hyun Yong Jeon, Jacob Kach, Enxuan Jing, Sebastien Weyn-Vanhentenryck, Mikaela Downs, Anna Corrionero, Raymond Oh, Juergen Scharner, Aditya Venkatesh, Sophina Ji, Gene Liau, Barry Ticho, Huw Nash, and Isabel Aznarez. Antisense oligonucleotide modulation of non-productive alternative splicing upregulates gene expression. *Nature Communications*, 11(1):3501, July 2020. ISSN 2041-1723. doi: 10.1038/s41467-020-17093-9. URL https://www.nature.com/articles/s41467-020-17093-9. Number: 1 Publisher: Nature Publishing Group.

[159] Dongjing Liu, Dara Meyer, Brian Fennessy, Claudia Feng, Esther Cheng, Jessica S. Johnson, You Jeong Park, Marysia-Kolbe Rieder, Steven Ascolillo, Agathe de Pins, Amanda Dobbyn, Dannielle Lebovitch, Emily Moya, Tan-Hoang Nguyen, Lillian Wilkins, Arsalan Hassan, Katherine E. Burdick, Joseph D. Buxbaum, Enrico Domenici, Sophia Frangou, Annette M. Hartmann, Claudine Laurent-Levinson, Dheeraj Malhotra, Carlos N. Pato, Michele T. Pato, Kerry Ressler, Panos Roussos, Dan Rujescu, Celso Arango, Alessandro Bertolino, Giuseppe Blasi, Luisella Bocchio-Chiavetto, Dominique Campion, Vaughan Carr, Janice M. Fullerton, Massimo Gennarelli, Javier González-Peñas, Douglas F. Levinson, Bryan Mowry, Vishwajit L. Nimgaokar, Giulio Pergola, Antonio Rampino, Jorge A. Cervilla, Margarita Rivera, Sibylle G. Schwab, Dieter B. Wildenauer, Mark Daly, Benjamin Neale, Tarjinder Singh, Michael C. O'Donovan, Michael J. Owen, James T. Walters, Muhammad Ayub, Anil K. Malhotra, Todd Lencz, Patrick F. Sullivan, Pamela Sklar, Eli A. Stahl, Laura M. Huckins, and Alexander W. Charney. Schizophrenia risk conferred by rare protein-truncating variants is conserved across diverse human populations. *Nature Genetics*, 55(3):369–376, 2023. ISSN 1061-4036. doi: 10.1038/s41588-023-01305-1. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10011128/.

[160] Yan Liu, Huisheng Liu, Conall Sauvey, Lin Yao, Ewa D. Zarnowska, and Su-Chun Zhang. Directed differentiation of forebrain GABA interneurons from human pluripotent stem cells. *Nature Protocols*, 8(9):1670–1679, September 2013. ISSN 1750-2799. doi: 10.1038/nprot.2013.106.

[161] Claudio Lorenzi, Sylvain Barriere, Katharina Arnold, Reini F. Luco, Andrew J. Oldfield, and William Ritchie. IRFinder-S: a comprehensive suite to discover and explore intron retention. *Genome Biology*, 22(1):307, November 2021. ISSN 1474-760X. doi: 10.1186/s13059-021-02515-8. URL https://doi.org/10.1186/s13059-021-02515-8.

[162] Michael I. Love, Wolfgang Huber, and Simon Anders. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12):550, December 2014. ISSN 1474-760X. doi: 10.1186/s13059-014-0550-8. URL https://doi.org/10.1186/s13059-014-0550-8.

[163] J. Lévy, S. Grotto, C. Mignot, A. Maruani, A. Delahaye-Duriez, B. Benzacken, B. Keren, D. Haye, J. Xavier, M. Heulin, E. Charles, A. Verloes, C. Dupont, E. Pipiras, and A.-C. Tabet. NR4A2 haploinsufficiency is associated with intellectual disability and autism spectrum disorder. *Clinical Genetics*, 94(2):264–268, 2018. ISSN 1399-0004. doi: 10.1111/cge.13383. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/cge.13383. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/cge.13383.

[164] Karen J. Mackenzie, Paula Carroll, Carol-Anne Martin, Olga Murina, Adeline Fluteau, Daniel J. Simpson, Nelly Olova, Hannah Sutcliffe, Jacqueline K. Rainger, Andrea Leitch, Ruby T. Osborn, Ann P. Wheeler, Marcin Nowotny, Nick Gilbert, Tamir Chandra, Martin A. M. Reijns, and Andrew P. Jackson. cGAS surveillance of micronuclei links genome instability to innate immunity. *Nature*, 548(7668):461–465, August 2017. ISSN 1476-4687. doi: 10.1038/nature23449. URL https://www.nature.com/articles/nature23449. Number: 7668 Publisher: Nature Publishing Group.

[165] Dario Magnani, Laurette Morlé, Kerstin Hasenpusch-Theil, Marie Paschaki, Monique Jacoby, Stéphane Schurmans, Bénédicte Durand, and Thomas Theil. The ciliogenic transcription factor Rfx3 is required for the formation of the thalamocortical tract by regulating the patterning of prethalamus and ventral telencephalon. *Human Molecular Genetics*, 24(9):2578–2593, May 2015. ISSN 0964-6906. doi: 10.1093/hmg/ddv021. URL https://doi.org/10.1093/hmg/ddv021.

[166] Dominique Makowski, Mattan S. Ben-Shachar, and Daniel Lüdecke. bayestestR: Describing Effects and their Uncertainty, Existence and Significance within the Bayesian Framework. *Journal of Open Source Software*, 4(40):1541, August 2019. ISSN 2475-9066. doi: 10.21105/joss.01541. URL https://joss.theoj.org/papers/10.21105/joss.01541.

[167] Abed AlFatah Mansour, J. Tiago Gonçalves, Cooper W. Bloyd, Hao Li, Sarah Fernandes, Daphne Quang, Stephen Johnston, Sarah L. Parylak, Xin Jin, and Fred H. Gage. An in vivo model of functional and vascularized human brain organoids. *Nature Biotechnology*, 36(5):432–441, May 2018. ISSN 1546-1696. doi: 10.1038/nbt.4127. URL https://www.nature.com/articles/nbt.4127. Number: 5 Publisher: Nature Publishing Group.

[168] Alberto Mantovani, Charles A. Dinarello, Martina Molgora, and Cecilia Garlanda. Interleukin-1 and Related Cytokines in the Regulation of Inflammation and Immunity. *Immunity*, 50(4):778–795, April 2019. ISSN 1074-7613. doi: 10.1016/j.immuni.2019.03.012. URL https://www.sciencedirect.com/science/article/pii/S1074761319301293.

[169] Maria C. Marchetto, Haim Belinson, Yuan Tian, Beatriz C. Freitas, Chen Fu, Krishna Vadodaria, Patricia Beltrao-Braga, Cleber A. Trujillo, Ana P. D. Mendes, Krishnan Padmanabhan, Yanelli Nunez, Jing Ou, Himanish Ghosh, Rebecca Wright, Kristen Brennand, Karen Pierce, Lawrence Eichenfield, Tiziano Pramparo, Lisa Eyler, Cynthia C. Barnes, Eric Courchesne, Daniel H. Geschwind, Fred H. Gage, Anthony Wynshaw-Boris, and Alysson R. Muotri. Altered proliferation and networks in neural cells derived from idiopathic autistic individuals. *Molecular Psychiatry*, 22(6):820–835, June 2017. ISSN 1476-5578. doi: 10.1038/mp.2016.95.

[170] Maria C. N. Marchetto, Cassiano Carromeu, Allan Acab, Diana Yu, Gene W. Yeo, Yangling Mu, Gong Chen, Fred H. Gage, and Alysson R. Muotri. A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. *Cell*, 143(4): 527–539, November 2010. ISSN 1097-4172. doi: 10.1016/j.cell.2010.10.016.

[171] Jessica Mariani, Gianfilippo Coppola, Ping Zhang, Alexej Abyzov, Lauren Provini, Livia Tomasini, Mariangela Amenduni, Anna Szekely, Dean Palejev, Michael Wilson, Mark Gerstein, Elena Grigorenko, Katarzyna Chawarska, Kevin Pelphrey, James Howe, and Flora M. Vaccarino. FOXG1-dependent dysregulation of GABA/glutamate neuron differentiation in autism spectrum disorders. *Cell*, 162(2):375–390, July 2015. ISSN 0092-8674. doi: 10.1016/j.cell.2015.06.034. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4519016/.

[172] Marcel Martin. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, 17(1):10–12, May 2011. ISSN 2226-6089. doi: 10.14806/ej.17.1.200. URL https://journal.embnet.org/index.php/embnetjournal/article/view/200. Number: 1.

[173] José L. Marín-Teva, Miguel A. Cuadros, David Martín-Oliva, and Julio Navascués. Microglia and neuronal cell death. *Neuron Glia Biology*, 7(1):25–40, February 2011. ISSN 1741-0533. doi: 10.1017/S1740925X12000014.

[174] Anouch Matevossian and Schahram Akbarian. Neuronal nuclei isolation from human postmortem brain tissue. *Journal of Visualized Experiments: JoVE*, (20):914, October 2008. ISSN 1940-087X. doi: 10.3791/914.

[175] Hansruedi Mathys, Jose Davila-Velderrain, Zhuyu Peng, Fan Gao, Shahin Mohammadi, Jennie Z. Young, Madhvi Menon, Liang He, Fatema Abdurrob, Xueqiao Jiang, Anthony J. Martorell, Richard M. Ransohoff, Brian P. Hafler, David A. Bennett, Manolis Kellis, and Li-Huei Tsai. Single-cell transcriptomic analysis of Alzheimer's disease. *Nature*, 570(7761): 332–337, June 2019. ISSN 1476-4687. doi: 10.1038/s41586-019-1195-2. URL https://www.nature.com/articles/s41586-019-1195-2. Number: 7761 Publisher: Nature Publishing Group.

[176] Davis J. McCarthy, Yunshun Chen, and Gordon K. Smyth. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Research*, 40(10):4288–4297, May 2012. ISSN 1362-4962. doi: 10.1093/nar/gks042.

[177] Madelyn E. McCauley and C. Frank Bennett. Antisense drugs for rare and ultra-rare genetic neurological diseases. *Neuron*, 0(0), June 2023. ISSN 0896-6273. doi: 10.1016/j.neuron.2023.05.027. URL https://www.cell.com/neuron/abstract/S0896-6273(23)00401-4. Publisher: Elsevier.

[178] Christopher S. McGinnis, Lyndsay M. Murrow, and Zev J. Gartner. DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Systems*, 8(4):329–337.e4, April 2019. ISSN 2405-4720. doi: 10.1016/j.cels.2019.03.003.

[179] Michele Menotta, Sara Biagiotti, Marzia Bianchi, Luciana Chessa, and Mauro Magnani. Dexamethasone Partially Rescues Ataxia Telangiectasia-mutated (ATM) Deficiency in Ataxia Telangiectasia by Promoting a Shortened Protein Variant Retaining Kinase Activity. *The Journal of Biological Chemistry*, 287(49):41352–41363, November 2012. ISSN 0021-9258. doi: 10.1074/jbc.M112.344473. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3510833/.

[180] Eugenio Mercuri, Basil T. Darras, Claudia A. Chiriboga, John W. Day, Craig Campbell, Anne M. Connolly, Susan T. Iannaccone, Janbernd Kirschner, Nancy L. Kuntz, Kayoko Saito, Perry B. Shieh, Már Tulinius, Elena S. Mazzone, Jacqueline Montes, Kathie M. Bishop, Qingqing Yang, Richard Foster, Sarah Gheuens, C. Frank Bennett, Wildon Farwell, Eugene Schneider, Darryl C. De Vivo, and Richard S. Finkel. Nusinersen versus Sham Control in Later-Onset Spinal Muscular Atrophy. *New England Journal of Medicine*, 378(7):625–635, February 2018. ISSN 0028-4793, 1533-4406. doi: 10.1056/NEJMoa1710504. URL http://www.nejm.org/doi/10.1056/NEJMoa1710504.

[181] Jason Merkin, Caitlin Russell, Ping Chen, and Christopher B. Burge. Evolutionary dynamics of gene and isoform regulation in Mammalian tissues. *Science (New York, N.Y.)*, 338(6114):1593–1599, December 2012. ISSN 1095-9203. doi: 10.1126/science.1228186.

[182] Natalie-Ya Mevises, Carmen Falcone, Tiffany Hong, Brett Dufour, Xiaohui Chen, Stephen C. Noctor, and Verónica Martínez Cerdeño. NEURONAL AND GLIAL CELL NUMBER IS ALTERED IN A CORTICAL LAYER-SPECIFIC MANNER IN AUTISM. *Autism : the international journal of research and practice*, 25(8):2238, November 2021. doi: 10.1177/13623613211014408. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9762515/. Publisher: NIH Public Access.

[183] Robert Middleton, Dadi Gao, Aubin Thomas, Babita Singh, Amy Au, Justin J-L Wong, Alexandra Bomane, Bertrand Cosson, Eduardo Eyras, John E. J. Rasko, and William Ritchie.

IRFinder: assessing the impact of intron retention on mammalian gene expression. *Genome Biology*, 18(1):51, March 2017. ISSN 1474-760X. doi: 10.1186/s13059-017-1184-4. URL https://doi.org/10.1186/s13059-017-1184-4.

[184] Borbala Mifsud, Filipe Tavares-Cadete, Alice N. Young, Robert Sugar, Stefan Schoenfelder, Lauren Ferreira, Steven W. Wingett, Simon Andrews, William Grey, Philip A. Ewels, Bram Herman, Scott Happe, Andy Higgs, Emily LeProust, George A. Follows, Peter Fraser, Nicholas M. Luscombe, and Cameron S. Osborne. Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. *Nature Genetics*, 47(6):598–606, June 2015. ISSN 1546-1718. doi: 10.1038/ng.3286. URL https://www.nature.com/articles/ng.3286. Number: 6 Publisher: Nature Publishing Group.

[185] Yuki Miura, Min-Yin Li, Fikri Birey, Kazuya Ikeda, Omer Revah, Mayuri Vijay Thete, Jin-Young Park, Alyssa Puno, Samuel H. Lee, Matthew H. Porteus, and Sergiu P. Pașca. Generation of human striatal organoids and cortico-striatal assembloids from human pluripotent stem cells. *Nature Biotechnology*, 38(12):1421–1430, December 2020. ISSN 1546-1696. doi: 10.1038/s41587-020-00763-w.

[186] Kiana Mohajeri, Rachita Yadav, Eva D'haene, Philip M. Boone, Serkan Erdin, Dadi Gao, Mariana Moyses-Oliveira, Riya Bhavsar, Benjamin B. Currall, Kathryn O'Keefe, Nicholas D. Burt, Chelsea Lowther, Diane Lucente, Monica Salani, Mathew Larson, Claire Redin, Olga Dudchenko, Erez Lieberman Aiden, Björn Menten, Derek J. C. Tai, James F. Gusella, Sarah Vergult, and Michael E. Talkowski. Transcriptional and functional consequences of alterations to MEF2C and its topological organization in neuronal models. *The American Journal of Human Genetics*, 109(11):2049–2067, November 2022. ISSN 0002-9297. doi: 10.1016/j.ajhg.2022.09.015. URL https://www.sciencedirect.com/science/article/pii/S0002929722004487.

[187] C. S. Moreno, G. W. Beresford, P. Louis-Plence, A. C. Morris, and J. M. Boss. CREB regulates MHC class II expression in a CIITA-dependent manner. *Immunity*, 10(2):143–151, February 1999. ISSN 1074-7613. doi: 10.1016/s1074-7613(00)80015-1.

[188] Meredith A. Mortberg, Juliana E. Gentile, Naeem M. Nadaf, Charles Vanderburg, Sean Simmons, Dan Dubinsky, Adam Slamin, Salome Maldonado, Caroline L. Petersen, Nichole Jones, Holly B. Kordasiewicz, Hien T. Zhao, Sonia M. Vallabh, and Eric Vallabh Minikel. A single-cell map of antisense oligonucleotide activity in the brain. *Nucleic Acids Research*, page gkad371, May 2023. ISSN 1362-4962. doi: 10.1093/nar/gkad371.

[189] Britt Mossink, Anouk H.A. Verboven, Eline J.H. van Hugte, Teun M. Klein Gunnewiek, Giulia Parodi, Katrin Linda, Chantal Schoenmaker, Tjitske Kleefstra, Tamas Kozicz, Hans van Bokhoven, Dirk Schubert, Nael Nadif Kasri, and Monica Frega. Human neuronal networks on micro-electrode arrays are a highly robust tool to study disease-specific genotype-phenotype correlations in vitro. *Stem Cell Reports*, 16(9):2182–2196, September 2021. ISSN

22136711. doi: 10.1016/j.stemcr.2021.07.001. URL https://linkinghub.elsevier.com/retrieve/pii/S221367112100326X.

[190] Corina Nagy, Malosree Maitra, Arnaud Tanti, Matthew Suderman, Jean-Francois Théroux, Maria Antonietta Davoli, Kelly Perlman, Volodymyr Yerko, Yu Chang Wang, Shreejoy J. Tripathy, Paul Pavlidis, Naguib Mechawar, Jiannis Ragoussis, and Gustavo Turecki. Single-nucleus transcriptomics of the prefrontal cortex in major depressive disorder implicates oligodendrocyte precursor cells and excitatory neurons. *Nature Neuroscience*, 23(6):771–781, June 2020. ISSN 1546-1726. doi: 10.1038/s41593-020-0621-y.

[191] Gergely Nagy and Laszlo Nagy. Motif grammar: The basis of the language of gene expression. *Computational and Structural Biotechnology Journal*, 18:2026–2032, July 2020. ISSN 2001-0370. doi: 10.1016/j.csbj.2020.07.007. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7406977/.

[192] Hisako Nakayama, Manabu Abe, Chie Morimoto, Tadatsune Iida, Shigeo Okabe, Kenji Sakimura, and Kouichi Hashimoto. Microglia permit climbing fiber elimination by promoting GABAergic inhibition in the developing cerebellum. *Nature Communications*, 9(1):2830, July 2018. ISSN 2041-1723. doi: 10.1038/s41467-018-05100-z. URL https://www.nature.com/articles/s41467-018-05100-z. Number: 1 Publisher: Nature Publishing Group.

[193] Sam P. Nayler, Joseph E. Powell, Darya P. Vanichkina, Othmar Korn, Christine A. Wells, Refik Kanjhan, Jian Sun, Ryan J. Taft, Martin F. Lavin, and Ernst J. Wolvetang. Human iPSC-Derived Cerebellar Neurons from a Patient with Ataxia-Telangiectasia Reveal Disrupted Gene Regulatory Networks. *Frontiers in Cellular Neuroscience*, 11, 2017. ISSN 1662-5102. URL https://www.frontiersin.org/articles/10.3389/fncel.2017.00321.

[194] Eric J. Nestler and Steven E. Hyman. Animal models of neuropsychiatric disorders. *Nature Neuroscience*, 13(10):1161–1169, October 2010. ISSN 1546-1726. doi: 10.1038/nn.2647. URL https://www.nature.com/articles/nn.2647. Number: 10 Publisher: Nature Publishing Group.

[195] The Lancet Neurology. Rare diseases: maintaining momentum. *The Lancet Neurology*, 21(3):203, March 2022. ISSN 1474-4422, 1474-4465. doi: 10.1016/S1474-4422(22)00046-1. URL https://www.thelancet.com/journals/laneur/article/PIIS1474-4422(22)00046-1/fulltext. Publisher: Elsevier.

[196] Rhys Newell, Richard Pienaar, Brad Balderson, Michael Piper, Alexandra Essebier, and Mikael Bodén. ChIP-R: Assembling reproducible sets of ChIP-seq and ATAC-seq peaks from multiple replicates. *Genomics*, 113(4):1855–1866, July 2021. ISSN 0888-7543. doi: 10.1016/j.ygeno.2021.04.026. URL https://www.sciencedirect.com/science/article/pii/S0888754321001531.

[197] Muchun Niu, Wenjian Cao, Yongcheng Wang, Qiangyuan Zhu, Jiayi Luo, Baiping Wang, Hui Zheng, David A. Weitz, and Chenghang Zong. Droplet-based transcriptome profiling of individual synapses. *Nature Biotechnology*, pages 1–13, January 2023. ISSN 1546-1696. doi: 10.1038/s41587-022-01635-1. URL https://www.nature.com/articles/s41587-022-01635-1. Publisher: Nature Publishing Group.

[198] Ramil N. Nurtdinov, Irena I. Artamonova, Andrei A. Mironov, and Mikhail S. Gelfand. Low conservation of alternative splicing patterns in the human and mouse genomes. *Human Molecular Genetics*, 12(11):1313–1320, June 2003. ISSN 0964-6906. doi: 10.1093/hmg/ddg137. URL https://doi.org/10.1093/hmg/ddg137.

[199] Marta Olah, Ellis Patrick, Alexandra-Chloe Villani, Jishu Xu, Charles C. White, Katie J. Ryan, Paul Piehowski, Alifiya Kapasi, Parham Nejad, Maria Cimpean, Sarah Connor, Christina J. Yung, Michael Frangieh, Allison McHenry, Wassim Elyaman, Vlad Petyuk, Julie A. Schneider, David A. Bennett, Philip L. De Jager, and Elizabeth M. Bradshaw. A transcriptomic atlas of aged human microglia. *Nature Communications*, 9(1):539, February 2018. ISSN 2041-1723. doi: 10.1038/s41467-018-02926-5.

[200] Qun Pan, Ofer Shai, Leo J. Lee, Brendan J. Frey, and Benjamin J. Blencowe. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nature Genetics*, 40(12):1413–1415, December 2008. ISSN 1546-1718. doi: 10.1038/ng.259. URL https://www.nature.com/articles/ng.259. Number: 12 Publisher: Nature Publishing Group.

[201] Neelroop N. Parikshak, Rui Luo, Alice Zhang, Hyejung Won, Jennifer K. Lowe, Vijayendran Chandran, Steve Horvath, and Daniel H. Geschwind. Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell*, 155(5):1008–1021, November 2013. ISSN 1097-4172. doi: 10.1016/j.cell.2013.10.031.

[202] Austin P. Passaro, Onur Aydin, M. Taher A. Saif, and Steven L. Stice. Development of an objective index, neural activity score (NAS), reveals neural network ontogeny and treatment effects on microelectrode arrays. *Scientific Reports*, 11(1):9110, April 2021. ISSN 2045-2322. doi: 10.1038/s41598-021-88675-w. URL https://www.nature.com/articles/s41598-021-88675-w. Number: 1 Publisher: Nature Publishing Group.

[203] Bruna Paulsen, Silvia Velasco, Amanda J. Kedaigle, Martina Pigoni, Giorgia Quadrato, Anthony J. Deo, Xian Adiconis, Ana Uzquiano, Rafaela Sartore, Sung Min Yang, Sean K. Simmons, Panagiotis Symvoulidis, Kwanho Kim, Kalliopi Tsafou, Archana Podury, Catherine Abbate, Ashley Tucewicz, Samantha N. Smith, Alexandre Albanese, Lindy Barrett, Neville E. Sanjana, Xi Shi, Kwanghun Chung, Kasper Lage, Edward S. Boyden, Aviv Regev, Joshua Z. Levin, and Paola Arlotta. Autism genes converge on asynchronous development of shared neuron classes. *Nature*, 602(7896):268–273, February 2022. ISSN 1476-

4687. doi: 10.1038/s41586-021-04358-6. URL https://www.nature.com/articles/s41586-021-04358-6. Number: 7896 Publisher: Nature Publishing Group.

[204] Anca M. Paşca, Steven A. Sloan, Laura E. Clarke, Yuan Tian, Christopher D. Makinson, Nina Huber, Chul Hoon Kim, Jin-Young Park, Nancy A. O'Rourke, Khoa D. Nguyen, Stephen J. Smith, John R. Huguenard, Daniel H. Geschwind, Ben A. Barres, and Sergiu P. Paşca. Functional cortical neurons and astrocytes from human pluripotent stem cells in 3D culture. *Nature Methods*, 12(7):671–678, July 2015. ISSN 1548-7105. doi: 10.1038/nmeth.3415. URL http://www.nature.com/articles/nmeth.3415. Number: 7 Publisher: Nature Publishing Group.

[205] Sergiu P. Paşca, Paola Arlotta, Helen S. Bateup, J. Gray Camp, Silvia Cappello, Fred H. Gage, Jürgen A. Knoblich, Arnold R. Kriegstein, Madeline A. Lancaster, Guo-Li Ming, Alysson R. Muotri, In-Hyun Park, Orly Reiner, Hongjun Song, Lorenz Studer, Sally Temple, Giuseppe Testa, Barbara Treutlein, and Flora M. Vaccarino. A nomenclature consensus for nervous system organoids and assembloids. *Nature*, 609(7929):907–910, September 2022. ISSN 1476-4687. doi: 10.1038/s41586-022-05219-6. URL https://www.nature.com/articles/s41586-022-05219-6. Number: 7929 Publisher: Nature Publishing Group.

[206] Emily Petley, Alexander Yule, Shaun Alexander, Shalini Ojha, and William P. Whitehouse. The natural history of ataxia-telangiectasia (A-T): A systematic review. *PLoS ONE*, 17(3): e0264177, March 2022. ISSN 1932-6203. doi: 10.1371/journal.pone.0264177. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9049793/.

[207] Ulrich Pfisterer, Viktor Petukhov, Samuel Demharter, Johanna Meichsner, Jonatan J. Thompson, Mykhailo Y. Batiuk, Andrea Asenjo-Martinez, Navneet A. Vasistha, Ashish Thakur, Jens Mikkelsen, Istvan Adorjan, Lars H. Pinborg, Tune H. Pers, Jakob von Engelhardt, Peter V. Kharchenko, and Konstantin Khodosevich. Identification of epilepsy-associated neuronal subtypes and gene expression underlying epileptogenesis. *Nature Communications*, 11(1):5038, October 2020. ISSN 2041-1723. doi: 10.1038/s41467-020-18752-7. URL https://www.nature.com/articles/s41467-020-18752-7. Number: 1 Publisher: Nature Publishing Group.

[208] Damon Polioudakis, Luis de la Torre-Ubieta, Justin Langerman, Andrew G. Elkins, Xu Shi, Jason L. Stein, Celine K. Vuong, Susanne Nichterwitz, Melinda Gevorgian, Carli K. Opland, Daning Lu, William Connell, Elizabeth K. Ruzzo, Jennifer K. Lowe, Tarik Hadzic, Flora I. Hinz, Shan Sabri, William E. Lowry, Mark B. Gerstein, Kathrin Plath, and Daniel H. Geschwind. A Single-Cell Transcriptomic Atlas of Human Neocortical Development during Mid-gestation. *Neuron*, 103(5):785–801.e8, September 2019. ISSN 0896-6273. doi: 10.1016/j.neuron.2019.06.011. URL https://www.sciencedirect.com/science/article/pii/S0896627319305616.

[209] Katherine S. Pollard, Sofie R. Salama, Bryan King, Andrew D. Kern, Tim Dreszer, Sol Katzman, Adam Siepel, Jakob S. Pedersen, Gill Bejerano, Robert Baertsch, Kate R. Rosenbloom, Jim Kent, and David Haussler. Forces Shaping the Fastest Evolving Regions in the Human Genome. *PLOS Genetics*, 2(10):e168, October 2006. ISSN 1553-7404. doi: 10.1371/journal.pgen.0020168. URL https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.0020168. Publisher: Public Library of Science.

[210] Alex A. Pollen, Tomasz J. Nowakowski, Jiadong Chen, Hanna Retallack, Carmen Sandoval-Espinosa, Cory R. Nicholas, Joe Shuga, Siyuan J. Liu, Michael C. Oldham, Aaron Diaz, Daniel A. Lim, Anne A. Leyrat, Jay A. West, and Arnold R. Kriegstein. Molecular Identity of Human Outer Radial Glia During Cortical Development. *Cell*, 163(1):55–67, September 2015. ISSN 0092-8674. doi: 10.1016/j.cell.2015.09.004. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4583716/.

[211] Andrew Polyak, Jill A. Rosenfeld, and Santhosh Girirajan. An assessment of sex bias in neurodevelopmental disorders. *Genome Medicine*, 7(1):94, August 2015. ISSN 1756-994X. doi: 10.1186/s13073-015-0216-5. URL https://doi.org/10.1186/s13073-015-0216-5.

[212] Galina Popova, Sarah S. Soliman, Chang N. Kim, Matthew G. Keefe, Kelsey M. Hennick, Samhita Jain, Tao Li, Dario Tejera, David Shin, Bryant B. Chhun, Christopher S. McGinnis, Matthew Speir, Zev J. Gartner, Shalin B. Mehta, Maximilian Haeussler, Keith B. Hengen, Richard R. Ransohoff, Xianhua Piao, and Tomasz J. Nowakowski. Human microglia states are conserved across experimental models and regulate neural stem cell responses in chimeric organoids. *Cell Stem Cell*, 28(12):2153–2166.e6, December 2021. ISSN 1875-9777. doi: 10.1016/j.stem.2021.08.015.

[213] Jing Qiu, Jamie McQueen, Bilada Bilican, Owen Dando, Dario Magnani, Karolina Punovuori, Bhuvaneish T Selvaraj, Matthew Livesey, Ghazal Haghi, Samuel Heron, Karen Burr, Rickie Patani, Rinku Rajan, Olivia Sheppard, Peter C Kind, T Ian Simpson, Victor LJ Tybulewicz, David JA Wyllie, Elizabeth MC Fisher, Sally Lowell, Siddharthan Chandran, and Giles E Hardingham. Evidence for evolutionary divergence of activity-dependent gene expression in developing neurons. *eLife*, 5:e20337, October 2016. ISSN 2050-084X. doi: 10.7554/eLife.20337. URL https://doi.org/10.7554/eLife.20337. Publisher: eLife Sciences Publications, Ltd.

[214] Xiaojie Qiu, Qi Mao, Ying Tang, Li Wang, Raghav Chawla, Hannah A. Pliner, and Cole Trapnell. Reversed graph embedding resolves complex single-cell trajectories. *Nature Methods*, 14(10):979–982, October 2017. ISSN 1548-7105. doi: 10.1038/nmeth.4402.

[215] Hazel Quek, John Luff, KaGeen Cheung, Sergei Kozlov, Magtouf Gatei, C. Soon Lee, Mark C. Bellingham, Peter G. Noakes, Yi Chieh Lim, Nigel L. Barnett, Steven Dingwall, Ernst Wolvetang, Tomoji Mashimo, Tara L. Roberts, and Martin F. Lavin. A rat model of

ataxia-telangiectasia: evidence for a neurodegenerative phenotype. *Human Molecular Genetics*, 26(1):109–123, January 2017. ISSN 1460-2083. doi: 10.1093/hmg/ddw371.

[216] Fidel Ramírez, Devon P. Ryan, Björn Grüning, Vivek Bhardwaj, Fabian Kilpert, Andreas S. Richter, Steffen Heyne, Friederike Dündar, and Thomas Manke. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Research*, 44(W1):W160–165, July 2016. ISSN 1362-4962. doi: 10.1093/nar/gkw257.

[217] Timothy Ravasi, Harukazu Suzuki, Carlo Vittorio Cannistraci, Shintaro Katayama, Vladimir B. Bajic, Kai Tan, Altuna Akalin, Sebastian Schmeier, Mutsumi Kanamori-Katayama, Nicolas Bertin, Piero Carninci, Carsten O. Daub, Alistair R. R. Forrest, Julian Gough, Sean Grimmond, Jung-Hoon Han, Takehiro Hashimoto, Winston Hide, Oliver Hofmann, Atanas Kamburov, Mandeep Kaur, Hideya Kawaji, Atsutaka Kubosaki, Timo Lassmann, Erik van Nimwegen, Cameron Ross MacPherson, Chihiro Ogawa, Aleksandar Radovanovic, Ariel Schwartz, Rohan D. Teasdale, Jesper Tegnér, Boris Lenhard, Sarah A. Teichmann, Takahiro Arakawa, Noriko Ninomiya, Kayoko Murakami, Michihira Tagami, Shiro Fukuda, Kengo Imamura, Chikatoshi Kai, Ryoko Ishihara, Yayoi Kitazume, Jun Kawai, David A. Hume, Trey Ideker, and Yoshihide Hayashizaki. An Atlas of Combinatorial Transcriptional Regulation in Mouse and Man. *Cell*, 140(5):744–752, March 2010. ISSN 0092-8674. doi: 10.1016/j.cell.2010.01.044. URL https://www.sciencedirect.com/science/article/pii/S0092867410000796.

[218] Omer Revah, Felicity Gore, Kevin W. Kelley, Jimena Andersen, Noriaki Sakai, Xiaoyu Chen, Min-Yin Li, Fikri Birey, Xiao Yang, Nay L. Saw, Samuel W. Baker, Neal D. Amin, Shravanti Kulkarni, Rachana Mudipalli, Bianxiao Cui, Seiji Nishino, Gerald A. Grant, Juliet K. Knowles, Mehrdad Shamloo, John R. Huguenard, Karl Deisseroth, and Sergiu P. Pașca. Maturation and circuit integration of transplanted human cortical organoids. *Nature*, 610 (7931):319–326, October 2022. ISSN 1476-4687. doi: 10.1038/s41586-022-05277-w. URL https://www.nature.com/articles/s41586-022-05277-w. Number: 7931 Publisher: Nature Publishing Group.

[219] Carlo Rinaldi and Matthew J. A. Wood. Antisense oligonucleotides: the next frontier for treatment of neurological disorders. *Nature Reviews Neurology*, 14(1):9–21, January 2018. ISSN 1759-4766. doi: 10.1038/nrneurol.2017.148. URL https://www.nature.com/articles/nrneurol.2017.148. Number: 1 Publisher: Nature Publishing Group.

[220] Mark D. Robinson, Davis J. McCarthy, and Gordon K. Smyth. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26 (1):139–140, January 2010. ISSN 1367-4803. doi: 10.1093/bioinformatics/btp616. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2796818/.

[221] Megan E. Rowland, Jana M. Jajarmi, Tess S. M. Osborne, and Annie Vogel Ciernia. Insights Into the Emerging Role of Baf53b in Autism Spectrum Disorder. *Frontiers in Molecular*

Neuroscience, 15:805158, February 2022. ISSN 1662-5099. doi: 10.3389/fnmol.2022.805158. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8852769/.

[222] Avik Roy, Yiu K. Fung, Xiaojuan Liu, and Kalipada Pahan. Up-regulation of Microglial CD11b Expression by Nitric Oxide. *The Journal of biological chemistry*, 281(21):14971–14980, May 2006. ISSN 0021-9258. doi: 10.1074/jbc.M600236200. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1963414/.

[223] Hiroaki Sacai, Kazuto Sakoori, Kohtarou Konno, Kenichiro Nagahama, Honoka Suzuki, Takaki Watanabe, Masahiko Watanabe, Naofumi Uesaka, and Masanobu Kano. Autism spectrum disorder-like behavior caused by reduced excitatory synaptic transmission in pyramidal neurons of mouse prefrontal cortex. *Nature Communications*, 11(1):5140, October 2020. ISSN 2041-1723. doi: 10.1038/s41467-020-18861-3. URL https://www.nature.com/articles/s41467-020-18861-3. Number: 1 Publisher: Nature Publishing Group.

[224] Ishani Sahama, Kate Sinclair, Simona Fiori, Kerstin Pannek, Martin Lavin, and Stephen Rose. Altered corticomotor-cerebellar integrity in young ataxia telangiectasia patients. *Movement Disorders: Official Journal of the Movement Disorder Society*, 29(10):1289–1298, September 2014. ISSN 1531-8257. doi: 10.1002/mds.25970.

[225] Trilochan Sahoo, Aaron Theisen, Jill A. Rosenfeld, Allen N. Lamb, J. Britt Ravnan, Roger A. Schultz, Beth S. Torchia, Nicholas Neill, Ian Casci, Bassem A. Bejjani, and Lisa G. Shaffer. Copy number variants of schizophrenia susceptibility loci are associated with a spectrum of speech and developmental delays and behavior problems. *Genetics in Medicine: Official Journal of the American College of Medical Genetics*, 13(10):868–880, October 2011. ISSN 1530-0366. doi: 10.1097/GIM.0b013e3182217a06.

[226] Antoine-Emmanuel Saliba, Alexander J. Westermann, Stanislaw A. Gorski, and Jörg Vogel. Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Research*, 42(14):8845–8860, August 2014. ISSN 0305-1048. doi: 10.1093/nar/gku555. URL https://doi.org/10.1093/nar/gku555.

[227] Eleanor I. Sams, Jeffrey K. Ng, Victoria Tate, Ying-Chen Claire Hou, Yang Cao, Lucinda Antonacci-Fulton, Khadija Belhassan, Julie Neidich, Robi D. Mitra, F. Sessions Cole, Patricia Dickson, Jeffrey Milbrandt, and Tychele N. Turner. From karyotypes to precision genomics in 9p deletion and duplication syndromes. *Human Genetics and Genomics Advances*, 3(1):100081, December 2021. ISSN 2666-2477. doi: 10.1016/j.xhgg.2021.100081. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8756500/.

[228] Carlos Sanchez-Priego, Ruiqi Hu, Linda L. Boshans, Matthew Lalli, Justyna A. Janas, Sarah E. Williams, Zhiqiang Dong, and Nan Yang. Mapping cis-regulatory elements in human neurons links psychiatric disease heritability and activity-regulated transcriptional programs. *Cell Reports*, 39(9), May 2022. ISSN 2211-1247. doi: 10.1016/j.celrep.2022.110877.

URL https://www.cell.com/cell-reports/abstract/S2211-1247(22)00652-0. Publisher: Elsevier.

[229] Richard Sando, Natalia Gounko, Simon Pieraut, Lujian Liao, John Yates, and Anton Maximov. HDAC4 governs a transcriptional program essential for synaptic plasticity and memory. *Cell*, 151(4):821–834, November 2012. ISSN 1097-4172. doi: 10.1016/j.cell.2012.09.037.

[230] Dhruv Sareen, Jacqueline G. O'Rourke, Pratap Meera, A. K. M. G. Muhammad, Sharday Grant, Megan Simpkinson, Shaughn Bell, Sharon Carmona, Loren Ornelas, Anais Sahabian, Tania Gendron, Leonard Petrucelli, Michael Baughn, John Ravits, Matthew B. Harms, Frank Rigo, C. Frank Bennett, Thomas S. Otis, Clive N. Svendsen, and Robert H. Baloh. Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a C9ORF72 repeat expansion. *Science Translational Medicine*, 5(208):208ra149, October 2013. ISSN 1946-6242. doi: 10.1126/scitranslmed.3007529.

[231] Shruti Sasaki, Rachel Sun, Huynh-Hoa Bui, Jeff R. Crosby, Brett P. Monia, and Shuling Guo. Steric Inhibition of 5′ UTR Regulatory Elements Results in Upregulation of Human CFTR. *Molecular Therapy*, 27(10):1749–1757, October 2019. ISSN 1525-0016, 1525-0024. doi: 10.1016/j.ymthe.2019.06.016. URL https://www.cell.com/molecular-therapy-family/molecular-therapy/abstract/S1525-0016(19)30312-0. Publisher: Elsevier.

[232] F. Kyle Satterstrom, Jack A. Kosmicki, Jiebiao Wang, Michael S. Breen, Silvia De Rubeis, Joon-Yong An, Minshi Peng, Ryan Collins, Jakob Grove, Lambertus Klei, Christine Stevens, Jennifer Reichert, Maureen S. Mulhern, Mykyta Artomov, Sherif Gerges, Brooke Sheppard, Xinyi Xu, Aparna Bhaduri, Utku Norman, Harrison Brand, Grace Schwartz, Rachel Nguyen, Elizabeth E. Guerrero, Caroline Dias, Catalina Betancur, Edwin H. Cook, Louise Gallagher, Michael Gill, James S. Sutcliffe, Audrey Thurm, Michael E. Zwick, Anders D. Børglum, Matthew W. State, A. Ercument Cicek, Michael E. Talkowski, David J. Cutler, Bernie Devlin, Stephan J. Sanders, Kathryn Roeder, Mark J. Daly, and Joseph D. Buxbaum. Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell*, 180(3):568–584.e23, February 2020. ISSN 0092-8674. doi: 10.1016/j.cell.2019.12.036. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7250485/.

[233] Arpiar Saunders, Evan Z. Macosko, Alec Wysoker, Melissa Goldman, Fenna M. Krienen, Heather de Rivera, Elizabeth Bien, Matthew Baum, Laura Bortolin, Shuyu Wang, Aleksandrina Goeva, James Nemesh, Nolan Kamitaki, Sara Brumbaugh, David Kulp, and Steven A. McCarroll. Molecular Diversity and Specializations among the Cells of the Adult Mouse Brain. *Cell*, 174(4):1015–1030.e16, August 2018. ISSN 1097-4172. doi: 10.1016/j.cell.2018.07.028.

[234] K. Savitsky, A. Bar-Shira, S. Gilad, G. Rotman, Y. Ziv, L. Vanagaite, D. A. Tagle, S. Smith, T. Uziel, S. Sfez, M. Ashkenazi, I. Pecker, M. Frydman, R. Harnik, S. R. Patanjali, A. Simmons, G. A. Clines, A. Sartiel, R. A. Gatti, L. Chessa, O. Sanal, M. F. Lavin, N. G. Jaspers, A. M. Taylor, C. F. Arlett, T. Miki, S. M. Weissman, M. Lovett, F. S. Collins, and Y. Shiloh. A single ataxia telangiectasia gene with a product similar to PI-3 kinase. *Science (New York, N.Y.)*, 268(5218):1749–1753, June 1995. ISSN 0036-8075. doi: 10.1126/science.7792600.

[235] Lucas Schirmer, Dmitry Velmeshev, Staffan Holmqvist, Max Kaufmann, Sebastian Werneburg, Diane Jung, Stephanie Vistnes, John H. Stockley, Adam Young, Maike Steindel, Brian Tung, Nitasha Goyal, Aparna Bhaduri, Simone Mayer, Jan Broder Engler, Omer A. Bayraktar, Robin J. M. Franklin, Maximilian Haeussler, Richard Reynolds, Dorothy P. Schafer, Manuel A. Friese, Lawrence R. Shiow, Arnold R. Kriegstein, and David H. Rowitch. Neuronal vulnerability and multilineage diversity in multiple sclerosis. *Nature*, 573(7772): 75–82, September 2019. ISSN 1476-4687. doi: 10.1038/s41586-019-1404-z. URL https://www.nature.com/articles/s41586-019-1404-z. Number: 7772 Publisher: Nature Publishing Group.

[236] Gil Ju Seo, Charlotte Kim, Woo-Jin Shin, Ella H. Sklan, Hyungjin Eoh, and Jae U. Jung. TRIM56-mediated monoubiquitination of cGAS for cytosolic DNA sensing. *Nature Communications*, 9(1):613, February 2018. ISSN 2041-1723. doi: 10.1038/s41467-018-02936-3. URL https://www.nature.com/articles/s41467-018-02936-3. Number: 1 Publisher: Nature Publishing Group.

[237] M. Sheng, G. McFadden, and M. E. Greenberg. Membrane depolarization and calcium induce c-fos transcription via phosphorylation of transcription factor CREB. *Neuron*, 4(4): 571–582, April 1990. ISSN 0896-6273. doi: 10.1016/0896-6273(90)90115-v.

[238] Steven D. Sheridan, Kraig M. Theriault, Surya A. Reis, Fen Zhou, Jon M. Madison, Laurence Daheron, Jeanne F. Loring, and Stephen J. Haggarty. Epigenetic Characterization of the FMR1 Gene and Aberrant Neurodevelopment in Human Induced Pluripotent Stem Cell Models of Fragile X Syndrome. *PLOS ONE*, 6(10):e26203, October 2011. ISSN 1932-6203. doi: 10.1371/journal.pone.0026203. URL https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0026203. Publisher: Public Library of Science.

[239] Yichen Shi, Peter Kirwan, and Frederick J. Livesey. Directed differentiation of human pluripotent stem cells to cerebral cortex neurons and neural networks. *Nature Protocols*, 7 (10):1836–1846, October 2012. ISSN 1750-2799. doi: 10.1038/nprot.2012.116.

[240] Sangwoo Shim, Raman Goyal, Alexios A. Panoutsopoulos, Olga A. Balashova, David Lee, and Laura N. Borodinsky. Calcium dynamics at the neural cell primary cilium regulate Hedgehog signaling–dependent neurogenesis in the embryonic neural tube. *Proceedings of the National Academy of Sciences*, 120(23):e2220037120, June 2023. doi: 10.1073/pnas.

2220037120. URL https://www.pnas.org/doi/abs/10.1073/pnas.2220037120. Publisher: Proceedings of the National Academy of Sciences.

[241] Jamie L. Shirley, Ype P. de Jong, Cox Terhorst, and Roland W. Herzog. Immune Responses to Viral Gene Therapy Vectors. *Molecular Therapy*, 28(3):709–722, March 2020. ISSN 1525-0016, 1525-0024. doi: 10.1016/j.ymthe.2020.01.001. URL https://www.cell.com/molecular-therapy-family/molecular-therapy/abstract/S1525-0016(20)30002-2. Publisher: Elsevier.

[242] Wolf Singer. Development and Plasticity of Cortical Processing Architectures. *Science*, 270(5237):758–764, November 1995. doi: 10.1126/science.270.5237.758. URL https://www-science-org.ezp-prod1.hul.harvard.edu/doi/10.1126/science.270.5237.758. Publisher: American Association for the Advancement of Science.

[243] Peter J Skene and Steven Henikoff. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *eLife*, 6:e21856, January 2017. ISSN 2050-084X. doi: 10.7554/eLife.21856. URL https://doi.org/10.7554/eLife.21856. Publisher: eLife Sciences Publications, Ltd.

[244] Nara Sobreira, François Schiettecatte, David Valle, and Ada Hamosh. GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. *Human Mutation*, 36(10):928–930, October 2015. ISSN 1098-1004. doi: 10.1002/humu.22844.

[245] Wilbur M. Song and Marco Colonna. The identity and function of microglia in neurodegeneration. *Nature Immunology*, 19(10):1048–1058, October 2018. ISSN 1529-2916. doi: 10.1038/s41590-018-0212-1. URL https://www.nature.com/articles/s41590-018-0212-1. Number: 10 Publisher: Nature Publishing Group.

[246] Xuan Song, Fulin Ma, and Karl Herrup. Accumulation of Cytoplasmic DNA Due to ATM Deficiency Activates the Microglial Viral Response System with Neurotoxic Consequences. *Journal of Neuroscience*, 39(32):6378–6394, August 2019. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.0774-19.2019. URL https://www.jneurosci.org/content/39/32/6378. Publisher: Society for Neuroscience Section: Research Articles.

[247] Staci A. Sorensen, Amy Bernard, Vilas Menon, Joshua J. Royall, Katie J. Glattfelder, Tsega Desta, Karla Hirokawa, Marty Mortrud, Jeremy A. Miller, Hongkui Zeng, John G. Hohmann, Allan R. Jones, and Ed S. Lein. Correlated gene expression and target specificity demonstrate excitatory projection neuron diversity. *Cerebral Cortex (New York, N.Y.: 1991)*, 25(2):433–449, February 2015. ISSN 1460-2199. doi: 10.1093/cercor/bht243.

[248] Kirsty L. Spalding, Ratan D. Bhardwaj, Bruce A. Buchholz, Henrik Druid, and Jonas Frisén. Retrospective birth dating of cells in humans. *Cell*, 122(1):133–143, July 2005. ISSN 0092-8674. doi: 10.1016/j.cell.2005.04.028.

[249] Hans-Georg Sprenger, Thomas MacVicar, Amir Bahat, Kai Uwe Fiedler, Steffen Hermans, Denise Ehrentraut, Katharina Ried, Dusanka Milenkovic, Nina Bonekamp, Nils-Göran Larsson, Hendrik Nolte, Patrick Giavalisco, and Thomas Langer. Cellular pyrimidine imbalance triggers mitochondrial DNA–dependent innate immunity. *Nature Metabolism*, 3(5):636–650, May 2021. ISSN 2522-5812. doi: 10.1038/s42255-021-00385-9. URL https://www.nature.com/articles/s42255-021-00385-9. Number: 5 Publisher: Nature Publishing Group.

[250] S. Stern, R. Santos, M. C. Marchetto, A. P. D. Mendes, G. A. Rouleau, S. Biesmans, Q.-W. Wang, J. Yao, P. Charnay, A. G. Bang, M. Alda, and F. H. Gage. Neurons derived from patients with bipolar disorder divide into intrinsically different sub-populations of neurons, predicting the patients' responsiveness to lithium. *Molecular Psychiatry*, 23(6):1453–1465, June 2018. ISSN 1476-5578. doi: 10.1038/mp.2016.260.

[251] Nadine Stirmlinger, Jan Philip Delling, Stefanie Pfänder, and Tobias M. Boeckers. Elevation of SHANK3 Levels by Antisense Oligonucleotides Directed Against the 3′ Untranslated Region of the Human SHANK3 mRNA. *nucleic acid therapeutics*, November 2022. doi: 10.1089/nat.2022.0048. URL https://www.liebertpub.com/doi/10.1089/nat.2022.0048. Publisher: Mary Ann Liebert, Inc., publishers 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA.

[252] Rianne D. Stowell, Elissa L. Wong, Hanna N. Batchelor, Monique S. Mendes, Cassandra E. Lamantia, Brendan S. Whitelaw, and Ania K. Majewska. Cerebellar microglia are dynamically unique and survey Purkinje neurons in vivo. *Developmental Neurobiology*, 78(6):627–644, June 2018. ISSN 1932-846X. doi: 10.1002/dneu.22572.

[253] Tim Stuart, Andrew Butler, Paul Hoffman, Christoph Hafemeister, Efthymia Papalexi, William M. Mauck, Yuhan Hao, Marlon Stoeckius, Peter Smibert, and Rahul Satija. Comprehensive Integration of Single-Cell Data. *Cell*, 177(7):1888–1902.e21, June 2019. ISSN 0092-8674, 1097-4172. doi: 10.1016/j.cell.2019.05.031. URL https://www.cell.com/cell/abstract/S0092-8674(19)30559-8. Publisher: Elsevier.

[254] Aravind Subramanian, Pablo Tamayo, Vamsi K. Mootha, Sayan Mukherjee, Benjamin L. Ebert, Michael A. Gillette, Amanda Paulovich, Scott L. Pomeroy, Todd R. Golub, Eric S. Lander, and Jill P. Mesirov. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, 102(43):15545–15550, October 2005. doi: 10.1073/pnas.0506580102. URL https://www.pnas.org/doi/10.1073/pnas.0506580102. Publisher: Proceedings of the National Academy of Sciences.

[255] Debora Sugiaman-Trapman, Morana Vitezic, Eeva-Mari Jouhilahti, Anthony Mathelier, Gilbert Lauter, Sougat Misra, Carsten O. Daub, Juha Kere, and Peter Swoboda. Characterization of the human RFX transcription factor family by regulatory and target gene analysis.

BMC Genomics, 19(1):181, March 2018. ISSN 1471-2164. doi: 10.1186/s12864-018-4564-6. URL https://doi.org/10.1186/s12864-018-4564-6.

[256] Wenjie Sun, Jeremie Poschmann, Ricardo Cruz-Herrera del Rosario, Neelroop N. Parikshak, Hajira Shreen Hajan, Vibhor Kumar, Ramalakshmi Ramasamy, T. Grant Belgard, Bavani Elanggovan, Chloe Chung Yi Wong, Jonathan Mill, Daniel H. Geschwind, and Shyam Prabhakar. Histone Acetylome-wide Association Study of Autism Spectrum Disorder. *Cell*, 167 (5):1385–1397.e11, November 2016. ISSN 0092-8674, 1097-4172. doi: 10.1016/j.cell.2016.10.031. URL https://www.cell.com/cell/abstract/S0092-8674(16)31451-9. Publisher: Elsevier.

[257] Zhiqi Sun and Veit Hornung. cGAS-STING signaling. *Current biology: CB*, 32(13):R730–R734, July 2022. ISSN 1879-0445. doi: 10.1016/j.cub.2022.05.027.

[258] P. Swoboda, H. T. Adler, and J. H. Thomas. The RFX-type transcription factor DAF-19 regulates sensory neuron cilium formation in C. elegans. *Molecular Cell*, 5(3):411–421, March 2000. ISSN 1097-2765. doi: 10.1016/s1097-2765(00)80436-0.

[259] Kazutoshi Takahashi, Koji Tanabe, Mari Ohnuki, Megumi Narita, Tomoko Ichisaka, Kiichiro Tomoda, and Shinya Yamanaka. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*, 131(5):861–872, November 2007. ISSN 0092-8674. doi: 10.1016/j.cell.2007.11.019.

[260] Masayuki Takahashi, Viorica Raluca Contu, Chihana Kabuta, Katsunori Hase, Yuuki Fujiwara, Keiji Wada, and Tomohiro Kabuta. SIDT2 mediates gymnosis, the uptake of naked single-stranded oligonucleotides into living cells. *RNA Biology*, 14(11):1534–1543, April 2017. ISSN 1547-6286. doi: 10.1080/15476286.2017.1302641. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5785214/.

[261] F. Tavani, R. A. Zimmerman, G. T. Berry, K. Sullivan, R. Gatti, and P. Bingham. Ataxia-telangiectasia: the pattern of cerebellar atrophy on MRI. *Neuroradiology*, 45(5):315–319, May 2003. ISSN 0028-3940. doi: 10.1007/s00234-003-0945-9.

[262] Tuan Leng Tay, Sagar, Jana Dautzenberg, Dominic Grün, and Marco Prinz. Unique microglia recovery population revealed by single-cell RNAseq following neurodegeneration. *Acta Neuropathologica Communications*, 6(1):87, September 2018. ISSN 2051-5960. doi: 10.1186/s40478-018-0584-3. URL https://doi.org/10.1186/s40478-018-0584-3.

[263] Samuel Thudium, Katherine Palozola, Éloïse L'Her, and Erica Korb. Identification of a transcriptional signature found in multiple models of ASD and related disorders. *Genome Research*, 32(9):1642–1654, September 2022. ISSN 1088-9051, 1549-5469. doi: 10.1101/gr.276591.122. URL https://genome.cshlp.org/content/32/9/1642. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press

Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab.

[264] Sally M. Till, Raven D. L. Hickson, and Peter C. Kind. Cross-species considerations in models of neurodevelopmental disorders. *Trends in Neurosciences*, 45(3):171–172, March 2022. ISSN 0166-2236, 1878-108X. doi: 10.1016/j.tins.2021.12.005. URL https://www.cell.com/trends/neurosciences/abstract/S0166-2236(21)00253-8. Publisher: Elsevier.

[265] Itay Tirosh, Benjamin Izar, Sanjay M. Prakadan, Marc H. Wadsworth, Daniel Treacy, John J. Trombetta, Asaf Rotem, Christopher Rodman, Christine Lian, George Murphy, Mohammad Fallahi-Sichani, Ken Dutton-Regester, Jia-Ren Lin, Ofir Cohen, Parin Shah, Diana Lu, Alex S. Genshaft, Travis K. Hughes, Carly G. K. Ziegler, Samuel W. Kazer, Aleth Gaillard, Kellie E. Kolb, Alexandra-Chloé Villani, Cory M. Johannessen, Aleksandr Y. Andreev, Eliezer M. Van Allen, Monica Bertagnolli, Peter K. Sorger, Ryan J. Sullivan, Keith T. Flaherty, Dennie T. Frederick, Judit Jané-Valbuena, Charles H. Yoon, Orit Rozenblatt-Rosen, Alex K. Shalek, Aviv Regev, and Levi A. Garraway. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science (New York, N.Y.)*, 352(6282):189–196, April 2016. ISSN 1095-9203. doi: 10.1126/science.aad0501.

[266] Hélène Tran, Michael P. Moazami, Huiya Yang, Diane McKenna-Yasek, Catherine L. Douthwright, Courtney Pinto, Jake Metterville, Minwook Shin, Nitasha Sanil, Craig Dooley, Ajit Puri, Alexandra Weiss, Nicholas Wightman, Heather Gray-Edwards, Miklos Marosfoi, Robert M. King, Thomas Kenderdine, Daniele Fabris, Robert Bowser, Jonathan K. Watts, and Robert H. Brown. Suppression of mutant C9orf72 expression by a potent mixed backbone antisense oligonucleotide. *Nature Medicine*, 28(1):117–124, January 2022. ISSN 1546-170X. doi: 10.1038/s41591-021-01557-6.

[267] Kai-Hei Tse and Karl Herrup. DNA damage in the oligodendrocyte lineage and its role in brain aging. *Mechanisms of Ageing and Development*, 161(Pt A):37–50, January 2017. ISSN 1872-6216. doi: 10.1016/j.mad.2016.05.006.

[268] Peter J. Uhlhaas, Frédéric Roux, Eugenio Rodriguez, Anna Rotarska-Jagiela, and Wolf Singer. Neural synchrony and the development of cortical networks. *Trends in Cognitive Sciences*, 14(2):72–80, February 2010. ISSN 1364-6613, 1879-307X. doi: 10.1016/j.tics.2009.12.002. URL https://www.cell.com/trends/cognitive-sciences/abstract/S1364-6613(09)00282-4. Publisher: Elsevier.

[269] Ana Uzquiano, Amanda J. Kedaigle, Martina Pigoni, Bruna Paulsen, Xian Adiconis, Kwanho Kim, Tyler Faits, Surya Nagaraja, Noelia Antón-Bolaños, Chiara Gerhardinger, Ashley Tucewicz, Evan Murray, Xin Jin, Jason Buenrostro, Fei Chen, Silvia Velasco, Aviv Regev, Joshua Z. Levin, and Paola Arlotta. Proper acquisition of cell class identity in organoids allows definition of fate specification programs of the human cerebral cortex. *Cell*, 185(20):3770–3788.e27, September 2022. ISSN 0092-8674, 1097-4172. doi: 10.1016/j.

cell.2022.09.010. URL https://www.cell.com/cell/abstract/S0092-8674(22)01168-0. Publisher: Elsevier.

[270] Teunis J. P. van Dam, Julie Kennedy, Robin van der Lee, Erik de Vrieze, Kirsten A. Wunderlich, Suzanne Rix, Gerard W. Dougherty, Nils J. Lambacher, Chunmei Li, Victor L. Jensen, Michel R. Leroux, Rim Hjeij, Nicola Horn, Yves Texier, Yasmin Wissinger, Jeroen van Reeuwijk, Gabrielle Wheway, Barbara Knapp, Jan F. Scheel, Brunella Franco, Dorus A. Mans, Erwin van Wijk, François Képès, Gisela G. Slaats, Grischa Toedt, Hannie Kremer, Heymut Omran, Katarzyna Szymanska, Konstantinos Koutroumpas, Marius Ueffing, Thanh-Minh T. Nguyen, Stef J. F. Letteboer, Machteld M. Oud, Sylvia E. C. van Beersum, Miriam Schmidts, Philip L. Beales, Qianhao Lu, Rachel H. Giles, Radek Szklarczyk, Robert B. Russell, Toby J. Gibson, Colin A. Johnson, Oliver E. Blacque, Uwe Wolfrum, Karsten Boldt, Ronald Roepman, Victor Hernandez-Hernandez, and Martijn A. Huynen. CiliaCarta: An integrated and validated compendium of ciliary genes. *PLoS ONE*, 14(5):e0216705, May 2019. ISSN 1932-6203. doi: 10.1371/journal.pone.0216705. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6522010/.

[271] Teunis JP van Dam, Gabrielle Wheway, Gisela G. Slaats, Martijn A. Huynen, Rachel H. Giles, and SYSCILIA Study Group. The SYSCILIA gold standard (SCGSv1) of known ciliary components and its applications within a systems biology consortium. *Cilia*, 2(1):7, May 2013. ISSN 2046-2530. doi: 10.1186/2046-2530-2-7. URL https://doi.org/10.1186/2046-2530-2-7.

[272] Joyce van de Leemput, Jayanth Chandran, Melanie A Knight, Lynne A Holtzclaw, Sonja Scholz, Mark R Cookson, Henry Houlden, Katrina Gwinn-Hardy, Hon-Chung Fung, Xian Lin, Dena Hernandez, Javier Simon-Sanchez, Nick W Wood, Paola Giunti, Ian Rafferty, John Hardy, Elsdon Storey, R. J. McKinlay Gardner, Susan M Forrest, Elizabeth M. C Fisher, James T Russell, Huaibin Cai, and Andrew B Singleton. Deletion at ITPR1 Underlies Ataxia in Mice and Spinocerebellar Ataxia 15 in Humans. *PLoS Genetics*, 3 (6):e108, June 2007. ISSN 1553-7390. doi: 10.1371/journal.pgen.0030108. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1892049/.

[273] Katy Vandereyken, Alejandro Sifrim, Bernard Thienpont, and Thierry Voet. Methods and applications for single-cell and spatial multi-omics. *Nature Reviews Genetics*, 24(8):494–515, August 2023. ISSN 1471-0064. doi: 10.1038/s41576-023-00580-2. URL https://www.nature.com/articles/s41576-023-00580-2. Number: 8 Publisher: Nature Publishing Group.

[274] Silvia Velasco, Amanda J. Kedaigle, Sean K. Simmons, Allison Nash, Marina Rocha, Giorgia Quadrato, Bruna Paulsen, Lan Nguyen, Xian Adiconis, Aviv Regev, Joshua Z. Levin, and Paola Arlotta. Individual brain organoids reproducibly form cell diversity of the human cerebral cortex. *Nature*, 570(7762):523–527, June 2019. ISSN 1476-4687. doi: 10.1038/s41586-019-1289-x.

[275] Dmitry Velmeshev, Lucas Schirmer, Diane Jung, Maximilian Haeussler, Yonatan Perez, Simone Mayer, Aparna Bhaduri, Nitasha Goyal, David H. Rowitch, and Arnold R. Kriegstein. Single-cell genomics identifies cell type-specific molecular changes in autism. *Science (New York, N.Y.)*, 364(6441):685–689, May 2019. ISSN 1095-9203. doi: 10.1126/science.aav8130.

[276] S. Vickovic, B. Lötstedt, J. Klughammer, S. Mages, Å Segerstolpe, O. Rozenblatt-Rosen, and A. Regev. SM-Omics is an automated platform for high-throughput spatial multi-omics. *Nature Communications*, 13(1):795, February 2022. ISSN 2041-1723. doi: 10.1038/s41467-022-28445-y. URL https://www.nature.com/articles/s41467-022-28445-y. Number: 1 Publisher: Nature Publishing Group.

[277] Philipp Wahle, Giovanna Brancati, Christoph Harmel, Zhisong He, Gabriele Gut, Jacobo Sarabia del Castillo, Aline Xavier da Silveira dos Santos, Qianhui Yu, Pascal Noser, Jonas Simon Fleck, Bruno Gjeta, Dinko Pavlinić, Simone Picelli, Max Hess, Gregor W. Schmidt, Tom T. A. Lummen, Yanyan Hou, Patricia Galliker, David Goldblum, Marton Balogh, Cameron S. Cowan, Hendrik P. N. Scholl, Botond Roska, Magdalena Renner, Lucas Pelkmans, Barbara Treutlein, and J. Gray Camp. Multimodal spatiotemporal phenotyping of human retinal organoid development. *Nature Biotechnology*, pages 1–11, May 2023. ISSN 1546-1696. doi: 10.1038/s41587-023-01747-2. URL https://www.nature.com/articles/s41587-023-01747-2. Publisher: Nature Publishing Group.

[278] Tom Walsh, Jon M. McClellan, Shane E. McCarthy, Anjené M. Addington, Sarah B. Pierce, Greg M. Cooper, Alex S. Nord, Mary Kusenda, Dheeraj Malhotra, Abhishek Bhandari, Sunday M. Stray, Caitlin F. Rippey, Patricia Roccanova, Vlad Makarov, B. Lakshmi, Robert L. Findling, Linmarie Sikich, Thomas Stromberg, Barry Merriman, Nitin Gogtay, Philip Butler, Kristen Eckstrand, Laila Noory, Peter Gochman, Robert Long, Zugen Chen, Sean Davis, Carl Baker, Evan E. Eichler, Paul S. Meltzer, Stanley F. Nelson, Andrew B. Singleton, Ming K. Lee, Judith L. Rapoport, Mary-Claire King, and Jonathan Sebat. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science (New York, N.Y.)*, 320(5875):539–543, April 2008. ISSN 1095-9203. doi: 10.1126/science.1155174.

[279] Zhong Wang, Mark Gerstein, and Michael Snyder. RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews. Genetics*, 10(1):57–63, January 2009. ISSN 1471-0056. doi: 10.1038/nrg2484. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2949280/.

[280] Dianne Watters, Kum Kum Khanna, Heather Beamish, Geoffrey Birrell, Kevin Spring, Padmini Kedar, Magtouf Gatei, Deborah Stenzel, Karen Hobson, Sergei Kozlov, Ning Zhang, Aine Farrell, Jonathan Ramsay, Richard Gatti, and Martin Lavin. Cellular localisation of the ataxia-telangiectasia (ATM) gene product and discrimination between mutated and normal forms. *Oncogene*, 14(16):1911–1921, April 1997. ISSN 1476-5594. doi: 10.1038/sj.onc.1201037. URL https://www.nature.com/articles/1201037. Number: 16 Publisher: Nature Publishing Group.

[281] Anne E. West and Michael E. Greenberg. Neuronal Activity–Regulated Gene Transcription in Synapse Development and Cognitive Function. *Cold Spring Harbor Perspectives in Biology*, 3(6):a005744, June 2011. ISSN , 1943-0264. doi: 10.1101/cshperspect.a005744. URL http://cshperspectives.cshlp.org/content/3/6/a005744. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab.

[282] J. G. White, E. Southgate, J. N. Thomson, and S. Brenner. The structure of the nervous system of the nematode Caenorhabditis elegans. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 314(1165):1–340, November 1986. ISSN 0962-8436. doi: 10.1098/rstb.1986.0056.

[283] Tom Whitington, Martin C. Frith, James Johnson, and Timothy L. Bailey. Inferring transcription factor complexes from ChIP-seq data. *Nucleic Acids Research*, 39(15):e98, August 2011. ISSN 0305-1048. doi: 10.1093/nar/gkr341. URL https://doi.org/10.1093/nar/gkr341.

[284] A. Jeremy Willsey, Stephan J. Sanders, Mingfeng Li, Shan Dong, Andrew T. Tebbenkamp, Rebecca A. Muhle, Steven K. Reilly, Leon Lin, Sofia Fertuzinhos, Jeremy A. Miller, Michael T. Murtha, Candace Bichsel, Wei Niu, Justin Cotney, A. Gulhan Ercan-Sencicek, Jake Gockley, Abha R. Gupta, Wenqi Han, Xin He, Ellen J. Hoffman, Lambertus Klei, Jing Lei, Wenzhong Liu, Li Liu, Cong Lu, Xuming Xu, Ying Zhu, Shrikant M. Mane, Ed S. Lein, Liping Wei, James P. Noonan, Kathryn Roeder, Bernie Devlin, Nenad Sestan, and Matthew W. State. Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. *Cell*, 155(5):997–1007, November 2013. ISSN 1097-4172. doi: 10.1016/j.cell.2013.10.020.

[285] Tiffany Wu, Borislav Dejanovic, Vineela D. Gandham, Alvin Gogineni, Rose Edmonds, Stephen Schauer, Karpagam Srinivasan, Melanie A. Huntley, Yuanyuan Wang, Tzu-Ming Wang, Maj Hedehus, Kai H. Barck, Maya Stark, Hai Ngu, Oded Foreman, William J. Meilandt, Justin Elstrott, Michael C. Chang, David V. Hansen, Richard A. D. Carano, Morgan Sheng, and Jesse E. Hanson. Complement C3 Is Activated in Human AD Brain and Is Required for Neurodegeneration in Mouse Models of Amyloidosis and Tauopathy. *Cell Reports*, 28(8):2111–2123.e6, August 2019. ISSN 2211-1247. doi: 10.1016/j.celrep.2019.07.060.

[286] Zhijian Wu, Hongyan Yang, and Peter Colosi. Effect of Genome Size on AAV Vector Packaging. *Molecular Therapy*, 18(1):80–86, January 2010. ISSN 1525-0016. doi: 10.1038/mt.2009.255. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2839202/.

[287] Yangfei Xiang, Yoshiaki Tanaka, Bilal Cakir, Benjamin Patterson, Kun-Yong Kim, Pingnan Sun, Young-Jin Kang, Mei Zhong, Xinran Liu, Prabir Patra, Sang-Hun Lee, Sherman M.

Weissman, and In-Hyun Park. hESC-Derived Thalamic Organoids Form Reciprocal Projections When Fused with Cortical Organoids. *Cell Stem Cell*, 24(3):487–497.e7, March 2019. ISSN 1934-5909, 1875-9777. doi: 10.1016/j.stem.2018.12.015. URL https://www.cell.com/cell-stem-cell/abstract/S1934-5909(18)30605-2. Publisher: Elsevier.

[288] Methodios Ximerakis, Scott L. Lipnick, Brendan T. Innes, Sean K. Simmons, Xian Adiconis, Danielle Dionne, Brittany A. Mayweather, Lan Nguyen, Zachary Niziolek, Ceren Ozek, Vincent L. Butty, Ruth Isserlin, Sean M. Buchanan, Stuart S. Levine, Aviv Regev, Gary D. Bader, Joshua Z. Levin, and Lee L. Rubin. Single-cell transcriptomic profiling of the aging mouse brain. *Nature Neuroscience*, 22(10):1696–1708, October 2019. ISSN 1546-1726. doi: 10.1038/s41593-019-0491-3. URL https://www.nature.com/articles/s41593-019-0491-3. Number: 10 Publisher: Nature Publishing Group.

[289] Xiao Xu, Elitsa I Stoyanova, Agata E Lemiesz, Jie Xing, Deborah C Mash, and Nathaniel Heintz. Species and cell-type properties of classically defined human and rodent neurons and glia. *eLife*, 7:e37551, October 2018. ISSN 2050-084X. doi: 10.7554/eLife.37551. URL https://doi.org/10.7554/eLife.37551. Publisher: eLife Sciences Publications, Ltd.

[290] Nan Yang, Soham Chanda, Samuele Marro, Yi-Han Ng, Justyna A. Janas, Daniel Haag, Cheen Euong Ang, Yunshuo Tang, Quetzal Flores, Moritz Mall, Orly Wapinski, Mavis Li, Henrik Ahlenius, John L. Rubenstein, Howard Y. Chang, Arturo Alvarez Buylla, Thomas C. Südhof, and Marius Wernig. Generation of pure GABAergic neurons by transcription factor programming. *Nature Methods*, 14(6):621–628, June 2017. ISSN 1548-7105. doi: 10.1038/nmeth.4291.

[291] Ee-Lynn Yap and Michael E. Greenberg. Activity-Regulated Transcription: Bridging the Gap between Neural Activity and Behavior. *Neuron*, 100(2):330–348, October 2018. ISSN 1097-4199. doi: 10.1016/j.neuron.2018.10.013.

[292] Ling Ye, Yunlong Huang, Lixia Zhao, Yuju Li, Lijun Sun, You Zhou, Guanxiang Qian, and Jialin C. Zheng. IL-1β and TNF-α induce neurotoxicity through glutamate production: a potential role for neuronal glutaminase. *Journal of neurochemistry*, 125(6):897–908, June 2013. ISSN 0022-3042. doi: 10.1111/jnc.12263. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3747774/.

[293] Ofer Yizhar, Lief E. Fenno, Matthias Prigge, Franziska Schneider, Thomas J. Davidson, Daniel J. O'Shea, Vikaas S. Sohal, Inbal Goshen, Joel Finkelstein, Jeanne T. Paz, Katja Stehfest, Roman Fudim, Charu Ramakrishnan, John R. Huguenard, Peter Hegemann, and Karl Deisseroth. Neocortical excitation/inhibition balance in information processing and social dysfunction. *Nature*, 477(7363):171–178, July 2011. ISSN 0028-0836. doi: 10.1038/nature10360. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4155501/.

[294] Guangchuang Yu, Li-Gen Wang, Yanyan Han, and Qing-Yu He. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics: A Journal of Integrative Biology*, 16(5):284–287, May 2012. ISSN 1557-8100. doi: 10.1089/omi.2011.0118.

[295] Guangchuang Yu, Li-Gen Wang, and Qing-Yu He. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics*, 31(14):2382–2383, July 2015. ISSN 1367-4803. doi: 10.1093/bioinformatics/btv145. URL https://doi.org/10.1093/bioinformatics/btv145.

[296] Jacob M. Zahn, Suresh Poosala, Art B. Owen, Donald K. Ingram, Ana Lustig, Arnell Carter, Ashani T. Weeraratna, Dennis D. Taub, Myriam Gorospe, Krystyna Mazan-Mamczarz, Edward G. Lakatta, Kenneth R. Boheler, Xiangru Xu, Mark P. Mattson, Geppino Falco, Minoru S. H. Ko, David Schlessinger, Jeffrey Firman, Sarah K. Kummerfeld, William H. Wood, Alan B. Zonderman, Stuart K. Kim, and Kevin G. Becker. AGEMAP: a gene expression database for aging in mice. *PLoS genetics*, 3(11):e201, November 2007. ISSN 1553-7404. doi: 10.1371/journal.pgen.0030201.

[297] T. Zayats, K. K. Jacobsen, R. Kleppe, C. P. Jacob, S. Kittel-Schneider, M. Ribasés, J. A. Ramos-Quiroga, V. Richarte, M. Casas, N. R. Mota, E. H. Grevet, M. Klein, J. Corominas, J. Bralten, T. Galesloot, A. A. Vasquez, S. Herms, A. J. Forstner, H. Larsson, G. Breen, P. Asherson, S. Gross-Lesch, K. P. Lesch, S. Cichon, M. B. Gabrielsen, O. L. Holmen, C. H. D. Bau, J. Buitelaar, L. Kiemeney, S. V. Faraone, B. Cormand, B. Franke, A. Reif, J. Haavik, and S. Johansson. Exome chip analyses in adult attention deficit hyperactivity disorder. *Translational Psychiatry*, 6(10):e923, October 2016. ISSN 2158-3188. doi: 10.1038/tp.2016.196.

[298] Yichong Zhang, Ji-Hoon Lee, Tanya T. Paull, Sarah Gehrke, Angelo D'Alessandro, Qianhui Dou, Vadim N. Gladyshev, Elizabeth A. Schroeder, Samantha K. Steyl, Brooke E. Christian, and Gerald S. Shadel. Mitochondrial redox sensing by the kinase ATM maintains cellular antioxidant capacity. *Science Signaling*, 11(538):eaaq0702, July 2018. ISSN 1937-9145. doi: 10.1126/scisignal.aaq0702.

[299] Yingsha Zhang, ChangHui Pak, Yan Han, Henrik Ahlenius, Zhenjie Zhang, Soham Chanda, Samuele Marro, Christopher Patzke, Claudio Acuna, Jason Covy, Wei Xu, Nan Yang, Tamas Danko, Lu Chen, Marius Wernig, and Thomas C. Südhof. Rapid Single-Step Induction of Functional Neurons from Human Pluripotent Stem Cells. *Neuron*, 78(5):785–798, June 2013. ISSN 0896-6273. doi: 10.1016/j.neuron.2013.05.029. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3751803/.

[300] Yong Zhang, Tao Liu, Clifford A. Meyer, Jérôme Eeckhoute, David S. Johnson, Bradley E. Bernstein, Chad Nusbaum, Richard M. Myers, Myles Brown, Wei Li, and X. Shirley Liu. Model-based Analysis of ChIP-Seq (MACS). *Genome Biology*, 9(9):R137, September 2008.

ISSN 1474-760X. doi: 10.1186/gb-2008-9-9-r137. URL https://doi.org/10.1186/gb-2008-9-9-r137.

[301] Wei Zhao, Kevin G. Johnston, Honglei Ren, Xiangmin Xu, and Qing Nie. Inferring neuron-neuron communications from single-cell transcriptomics through NeuronChat. *Nature Communications*, 14(1):1128, February 2023. ISSN 2041-1723. doi: 10.1038/s41467-023-36800-w. URL https://www.nature.com/articles/s41467-023-36800-w. Number: 1 Publisher: Nature Publishing Group.

[302] Yizhang Zhu, Likun Wang, Yuxin Yin, and Ence Yang. Systematic analysis of gene expression patterns associated with postmortem interval in human tissues. *Scientific Reports*, 7:5435, July 2017. ISSN 2045-2322. doi: 10.1038/s41598-017-05882-0. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5511187/.

This thesis was typeset using LaTeX, originally developed by Leslie Lamport and based on Donald Knuth's TeX. The body text is set in 11 point Egenolff-Berner Garamond, a revival of Claude Garamont's humanist typeface. The above illustration, "Science Experiment 02", was created by Ben Schlitter and released under CC BY-NC-ND 3.0. A template that can be used to format a PhD thesis with this look and feel has been released under the permissive MIT (X11) license, and can be found online at github.com/suchow/Dissertate or from its author, Jordan Suchow, at suchow@post.harvard.edu.