Multimodal 3D approaches to studying neurodevelopmental disorders

CELL TYPE EMERGENCE AND CIRCUIT DISRUPTIONS IN FETAL MODELS OF 15q13.3 MICRODELETION BRAIN DEVELOPMENT

By

Savannah Kilpatrick, BSc.

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

McMaster University DOCTOR OF PHILOSOPHY (2023) Hamilton, Ontario
(Biochemistry)

TITLE: CELL TYPE EMERGENCE AND CIRCUIT DISRUPTIONS IN FETAL MODELS OF 15q13.3 MICRODELETION BRAIN DEVELOPMENT

AUTHOR: Savannah Kilpatrick B.Sc. (Honours) (McMaster University)

SUPERVISOR: Dr. Karun K. Singh

NUMBER OF PAGES: CCLXIX, 269

Abstract

The 15q13.3 microdeletion is a common genetic disorder associated with multiple neurodevelopmental disorders including autism spectrum disorder, epilepsy, and schizophrenia. Patients have diverse clinical presentations, often prompting genetic assays that identify the CNV in the clinic. This late-stage screening leaves a considerable gap in our understanding of the prenatal and prediagnostic developmental impairments in these individuals, providing a barrier to understanding the disease pathobiology. We provide the first investigation into embryonic brain development of individuals with the 15q13.3 microdeletion by generating multiple 3D neural organoid models from the largest clinical cohort in reported literature. We incorporated unguided and guided forebrain organoid models into our multi-transcriptomic phenotyping pipeline to uncover changes in cell type emergence and disruptions to circuit development, all of which had underlying changes to cell adhesion pathways.

Specifically, we identified accelerated growth trajectories in 15q13.3del unguided neural organoids and used single cell RNA sequencing to identify changes in radial glia dynamics that affect neurogenesis. We measured changes in the pseudotemporal trajectory of matured unguided neural organoids, and later identified disruptions in synaptic signaling modules amongst the primary constituents to neural circuitry, excitatory and inhibitory neurons.

We leveraged dorsal and ventral forebrain organoid models to better assess circuit dynamics, as they faithfully produce the excitatory and inhibitory neurons in the pallium and subpallium, respectively. We then used the entire 15q13.3del cohort and performed bulk RNA sequencing on each tissue type at two timepoints and discovered convergence

on transcriptional dysregulation and disruptions to human-specific zinc finger proteins localized to chromosome 19. We also identified cell type-specific vulnerabilities to DNA damage and cell migration amongst the dorsal and ventral organoids, respectively, which was consistent with the excitatory and inhibitory neural subpopulations amongst the unguided neural organoids scRNA Seq, respectively.

We then examined neuron migration in a 3D assembloid model by sparsely labeling dorsal-ventral forebrain organoids from multiple genotype-lineage combinations. Light sheet microscopy identified deficits in inhibitory neuron migration and morphology, but not migration distance, suggesting a complex disruption to cortical circuitry. This novel combination of cell type characterization, pathway identification, and circuitry phenotyping provides a novel perspective of how the 15q13.3 deletions impair prenatal development and can be applied to other NDD models to leverage understanding of early disease pathogenesis.

Keywords

Neurodevelopmental disorders, hiPSC modeling, Brain organoids, Single cell RNA sequencing, Bulk RNA sequencing, Tissue clearing, Light-sheet microscopy.

Summary for Lay Audience

The development of the human brain is a highly complex and tightly regulated process that requires the participation of multiple cell types throughout development. Disturbances to the emergence, differentiation, or placement of these cell types can cause disruptions and local miswiring of neural circuits, which is often associated with neurodevelopmental disorders (NDDs). The 15q13.3 microdeletion syndrome is a highly complex condition associated with multiple NDDs and has seldom been studied in a human context. To address this, we used stem cells derived from a 15q13.3 microdeletion syndrome cohort and their typically developing familial controls to generate unguided ("whole brain") and region-specific organoids to investigate early fetal development across time.

We used the largest 15q13.3 microdeletion cohort in reported literature to identify shared disruptions in early developmental milestones such as neurogenesis, neural migration, and neural patterning. We identified expansion of specific cell populations, including progenitors that later give rise to mature neurons. Abnormalities persisted in more mature cell populations, including the inhibitory neurons responsible for establishing critical microcircuitry in the human cortex. By generating guided organoids that enrich for excitatory and inhibitory neural populations, we were able to merge the models to form assembloids, where we captured early migratory and morphological deficits in inhibitory neuron populations, which is supported by the multi-transcriptomics experiments performed in both organoid models. This study provides a framework for examining fetal development in a neurodevelopmental disorder context. By using the 15q13.3 microdeletion background, we found novel disruptions in cell type emergence

and circuit formation previously unreported in mouse or 2D neuron models, highlighting the utility of the phenotyping platform for disease modeling.

Contributions to the thesis

Experimental design, data acquisition, and analysis was carried out by myself with the following exceptions:

Chapter 3: Loss of 15q13.3 genes in unguided neural organoids disrupts cellular trajectories and predicted cellular communication.

The core scRNA Seq analysis script was amended from the Seurat package by Jarryll Uy.

Chapter 4: Abnormal developmental trajectories in guided dorsal ventral 15q13.3del dorsal and ventral forebrain organoids converge on human-specific transcriptional dysregulation and abnormal circuitry.

Processing of the raw bulk RNA Seq data (adapter trimming, read alignment) was performed using an analysis pipeline from the Hope lab written by Tony Chen that incorporated CutAdapt and STAR2Pass, respectively.

Supplementary data

Supplementary Figure 1: Sendai viral transfections and colony expansion for reprogramming for iPSCs was performed by Dr. Alexandria Afonso (patient family 3) and Annie Cheng (patient families 4-6). They also conducted flow cytometry for pluripotency validations.

Supplementary Figure 1: Mycoplasma PCR tests were performed by Dr. Yi Pan.

Supplementary Figure 4: iPSCs from three 15q13.3 families were infected by Dr. Leon Chalil for NGN2 axon morphology experiments.

Acknowledgments

This project was carried out over the traditional territories of the Mississaugas of the Credit, the Anishnabeg, the Chippewa, the Haudenosaunee and the Wendat Nations, which remain unceded territory.

I would like to firstly thank my supervisor, Dr. Karun Singh, for his constant support over the years and throughout this entire project; for always keeping your door (or email, or phone) open and available for scientific discussions no matter the hour. Thank you for investing in lab culture and ensuring the environment in and outside the lab is always a supportive one. Thank you to my committee members, Dr. Leslie MacNeil and Dr. Brad Doble, whose insight over the years helped shape the project and my expectations of what great science is.

I'd like to thank my lab members, both past and present, for fostering an environment of support and friendship – with special thanks to the following people: The 15q13.3 veterans, Dr. Brianna Unda, and Dr. Leon Chalil: for teaching me the value of a good culture and for your patience when I was just starting out. Thank you, Dr. Nadeem Murtaza, for being a pillar in the lab and for your constant engagement in my and everyone's projects. Philadelphia's gain is our loss. To Annie Cheng, for constantly supporting me, both scientifically and with your insanely good cooking, and for always checking in on my mental health. It's hard to find a friend as loyal as you. Thank you Jarryll Uy, for teaching me the ways of single cell RNA sequencing and the dank art of dog memes. To Laura Botler, for always providing support and a picture of a hometown capybara when needed.

I'd like to thank members of the Wallace lab, Drs. Akshay Gurdita and Neno Pokrajac, for teaching me the magic of tissue clearing and for your patience whenever I crashed the IMARIS software. Thank you to Tony Chen for your insight into best practices in transcriptomics, and for your genuine interest in a project entirely outside of your field. I've been supported by wonderful staff at McMaster and Krembil – thanks Dr. Kevin Conway, for humoring some wild imaging experiments on my end, and for investing in my science.

I'm grateful to the friends I've lived with: Tanya Karyakina, Waverley Birch, and Kieran Lehan – thank you for taking me into your home and into the woods when needed; you really made a house a home. My family has provided unwavering support through this entire process, and I'm endlessly grateful for their love and understanding when science got in the way. Lastly, I need to thank my partner, Craig Matthews, for waking up at ungodly hours just to share a coffee with me before I head to lab. I couldn't have done this without your support and frequent reminder that better things are just around the corner.

TABLE OF CONTENTS

List of abbreviations and symbols

2D	Two dimensional
3D	Three dimensional
5HT3aR	Ionotropic serotonin receptor 5HT3a
ADHD	Attention deficit hyperactivity disorder
ARHGAP11B	Rho GTPase Activating Protein 11B
ASD	Autism spectrum disorder
bp	Base pairs
BrdU	5-bromo-2'-deoxyuridine
CAM	Cell adhesion molecule
cDNA	Complimentary DNA
ChIP-Seq	Chromatin immunoprecipitation sequencing
CHRNA7	Cholinergic receptor, nicotinic, alpha 7
CNS	Central nervous system
CNV	Copy number variant
CR	Cajal-Retzius
CRISPR	Clustered Regularly Interspaced short Palindromic Repeats
CUBIC	Clear, Unobstructed Brain Imaging Cocktails and Computational Analysis
Df(h15q13)/+	15q13.3 heterozygous mouse
DNA	Deoxyribonucleic acid
DD	Developmental delay
DEG	Differentially expressed gene

Dox	Doxycycline
D-V	Dorsal-ventral
ECM	Extracellular matrix
EEG	Electroencephalography
EIF3L	Eukaryotic Translation Initiation Factor 3 Subunit L
E/I	Excitatory/inhibitory
ENCODE	Encyclopedia of DNA Elements
ES	Enrichment score
ESC	Embryonic stem cell
ExN	Excitatory neuron
ExM	Excitatory mature neuron
FAN1	FANCD2 and FANCI associated nuclease 1
FDR	False discovery rate
FMR1	Fragile X mental retardation 1
GABA	Gamma-aminobutyric acid
GEM	Gel bead in emulsion
GNO	Guided neural organoid
GO	Gene ontology
GWAS	Genome wide association studies
HET	Heterozygous
Hind	Hindbrain
ICC	Immunocytochemistry
ID	Intellectual disability
IHC	Immunohistochemistry

IP	Intermediate progenitor
iPSC	Induced pluripotent stem cell
KD	Knockdown
KLF13	KLF Transcription factor 13
KO	Knockout
Mid	Midbrain
MRI	Magnetic resonance imaging
mRNA	Messenger RNA
NAHR	Non-allelic homologous recombination
NDD	Neurodevelopmental disorder
NGS	Next generation sequencing
Non-tel	Non-telencephalic
NPC	Neural progenitor cell
OCT	Optimal cutting temperature
oRG	Outer radial glia
oSVZ	Outer subventricular zone
OTUD7A	OTU deubiquitinase 7A
pAdj	Adjusted p value
PAM	Positive allosteric modulator
PBS/-T	Phosphate buffered saline/ -Supplemented with Tween-20
PCA	Principal component analysis
PFA	Paraformaldehyde
PNS	Peripheral nervous system
PV	Parvalbumin

RGC	Radial glia cell
RI	Refractive index
RIN	RNA integrity number
RNA	Ribonucleic acid
scATAC Seq	Single cell sequencing assay for transposase-accessible chromatin RNA sequencing
scRNA Seq	Single cell RNA sequencing
SCZ	Schizophrenia
SFARI	Simons Foundation Autism Research Initiative
Shh	Sonic Hedgehog
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variant
SST	Somatostatin
STRING	Search Tool for the Retrieval of Interacting Genes
SVZ	Subventricular zone
TRPM1	Transient receptor potential cation channel subfamily M member 1
UNO	Unguided neural organoid
UMI	Unique molecular identifier
VIP	Vasoactive Intestinal Peptide
VZ	Ventricular zone
Wnt	Wingless-related integration site
WT	Wild-type

List of Tables

List of Figures

CHAPTER 4

1 CHAPTER 1: INTRODUCTION

1.1 Neurodevelopment

The development of the human brain is a protracted process that begins within the third week of gestation and extends postnatally into adulthood. The neural plate represents the first neural structure formed in an embryo, and during neurulation it will fold into the neural tube and later give rise to the central and peripheral nervous systems (CNS and PNS, respectively).[1] As the neural tube closes, discrete regionalization occurs along the anterior-posterior axis to form brain regions such as the forebrain, midbrain, and hindbrain, which can be further divided within their dorsal-ventral compartments. These regions are produced by gradients of signaling molecules and patterning factors that drive neural fate and cell-type specification.[2,3]

1.1.1 Neurogenesis

Prior to neurogenesis, the neural tube is lined with a single layer of neuroepithelial cells, which later give rise to neural stem cells (NSCs), neural progenitor cells (NPCs), and finally the neurons and glia that populate the adult brain. First, NSCs rapidly undergo cellular division and expansion to form the tightly packed ventricular zones. Many of the cells take on an elongated radial morphology and are known as radial glia cells (RGCs).[4,5] RGCs accumulate in the ventricular zone (VZ) of the of the developing embryo, which is located above the fluid-filled ventricular systems of the neural tube. RGCs within the VZ divide asymmetrically to produce one radial glia daughter cell and either an intermediate neural progenitor (INP) cell type, or a neuron directly. Just above the VZ lies the subventricular zone (SVZ), which contains mostly INPs that also produce newborn neurons. Lastly, the outer subventricular zone (oSVZ) is

a primate-specific region that lies above the SVZ and greatly contributes towards primate (and specifically human) cortical expansion.[6] These ventricular zones produce primarily excitatory neurons; however, recent studies have suggested that a small subpopulation of inhibitory neurons are derived from cortical progenitors in early human development.[7] It should be noted that multiple proliferative or germinal zones exist outside of the cerebral cortex, including the ganglionic eminences of the forebrain, which produce inhibitory neurons that later migrate into the cortex to participate in neural circuitry.[5,8]

1.1.2 Neuronal migration

Before neurons can fully mature, they must first reach their final destination in the brain. The cerebral cortex is arranged in six layers and is occupied by distinct cell types unique to each layer. These cell types can broadly be classified as either excitatory pyramidal cells, originating from the cortical ventricular zone, or inhibitory neurons, which emerge later from the ganglionic eminences of the ventral telencephalon and migrate great distances to reach the cortex.[9]

Cortical excitatory migration occurs radially and is facilitated by radial glia progenitors that provide a scaffold from the base of the ventricular zone to the upper pial surface, which newborn neurons can travel along to their intended cortical layer. The lamination of the cortex occurs in an inside-out fashion, with the lower layers produced first, followed by upper layers. In this way, the older neurons can influence the migratory paths of newborn neurons through synaptic interactions and secretion of signaling factors, an example being the Reelin-emitting Cajal-Retzius cells that help guide later-born projection neurons in the cortex.[10] The older neurons also participate in subcortical circuit

development by extending centimeters-long neuronal projections that interact with brain regions such as the thalamus and brain stem.

In contrast to the excitatory neurons, cortical inhibitory interneurons emerge from the ganglionic eminences in the subpalium (below the cortex and in the ventral forebrain) and undergo tangential migration until an intermediate target is identified below the preplate; after which they switch to radial migration that is perpendicular to the cortical plate and guided by radial glia.[9,11] Here, they integrate within specific cortical layers to regulate the excitatory/inhibitory (E/I) balance of nearby circuitry through the release of inhibitory neurotransmitters such as gamma-aminobutyric acid (GABA). Most cortical inhibitory neurons are GABAergic and can be further classified based on the expression of cellularly-distinct genes such as: somatostatin (SST), parvalbumin (PV), and ionotropic serotonin receptor 5HT3a (5HT3aR). Further subclassifications of interneuron cell types are based on cellular morphology, firing patterns, and the cortical layers that they are located in. Neuron migration is reliant on multiple signaling molecules, cell types, adhesion molecules, cytoskeletal components, and scaffolding elements to ensure specific neural cell types are directed to proper brain regions.

Chemoattractants such as netrins,[12] neurotrophins, semaphorins, and Slit proteins[13–15] have all been implicated in facilitating NPC and neuron migration in the developing brain, often working with morphogens such as Sonic hedgehog (Shh) to provide a gradient of signaling cues for axonal pathfinding and cell migration. Motogenic factors such as Reelin can also stimulate a neuron into a migratory state and are released in a cell-type specific manner.[10] Some chemoattractants like netrins also have adhesive properties and act as cell adhesion molecules (CAMs) that allow migrating cells to

connect to various scaffolding and extracellular matrix (ECM) elements, in addition to providing proper pathfinding for the migratory route. Deficits in any of these processes can lead to failure in cellular motility and navigation, ultimately resulting in disruptions in the intended circuitry between cell types.

1.2 Neurodevelopmental disorders (NDDs)

Neurodevelopmental disorders are a heterogeneous group of disorders that arise from disruptions in the early developing brain. Individuals with NDDs include those with autism spectrum disorder (ASD), epilepsy, intellectual disability (ID), attention deficit hyperactivity disorder (ADHD), and schizophrenia (SCZ), of which many individuals have overlapping symptomatology or comorbidities.[16–18] The global prevalence of these NDDs has increased over the years,[19–22] which can largely be attributed to improvements in clinical diagnostics and more precise inclusion criteria.[23–25] Despite these improvements, the underlying pathobiology and mechanisms for these disorders remain unknown. NDD risk factors can occur prenatally, perinatally, and postnatally to include genetic perturbations and environmental insults, all of which can occur individually, or more likely, synergistically.

1.2.1 Autism spectrum disorder

Autism spectrum disorder (ASD) was first identified in 1925, when it was first described as a form of childhood schizophrenia. The initial diagnostic criteria included impairments in social interactions and repetitive and restricted behaviours and has since evolved to include impairments in sensory processing and responses to environmental stimuli. Changes in diagnostic methods have improved the accuracy of diagnoses and

have likely contributed to the increase in prevalence over the years. Clinical presentations of those with ASD vary widely, ranging from non-verbal individuals to mild intellectual impairments. The complexity of the disorder is furthered by the number of comorbidities, including epilepsy and seizure disorders (up to 30%),[26] intellectual disability (over 30%), ADHD (48%),[27] gastrointestinal disorders (9-9%), [28] anxiety (over 25%) and depression (%20-30).[29]

1.2.2 Epilepsy

Epilepsy and related seizure disorders are chronic neurological disorders that impact as many as 50 million individuals globally.[30] Seizures describe changes to neurological and/or bodily functions resulting from excessive and hypersynchronous neural activity; epilepsy is then a condition with recurrent, unprovoked seizures, and can be classified based on the category of seizure. Seizures are broadly classified into three groups based on the coverage of brain regions affected at the onset of the seizure: generalized (affecting regions across hemispheres) focal (formerly called partial,[31] affecting only a region of the brain), and epilepsies with unknown onset. Generalized seizures can be further subclassified based on the external presentation of the seizure, and include absence, tonic-clonic, myoclonic, and atonic generalized onset seizures. Absence seizures (formerly “petit mal”) often affect children and are typically brief experiences of altered consciousness and staring episodes, which can go unnoticed by the individual experiencing the seizure.[31,32] Roughly 70% of epilepsies can be treated with pharmacological intervention, while the remaining population is considered refractory to antiepileptic medication and may require surgical intervention.[32] Clinical assessments

must therefore be rigorous and explore the etiology from a structural and genetic perspective for effective treatments.

1.3 Genetics of NDDs

The above descriptions of just two neurodevelopmental disorders can make one appreciate the phenotypic diversity and clinical heterogeneity of NDDs. The high occurrence of comorbidities between NDDs, however, suggests a mechanistic overlap that is likely present at the genomic level. Indeed, large scale sequencing studies and consortia have identified causative genes and overlapping molecular processes to be implicated in these disorders,[33,34] including synaptic pathways and gene regulatory networks. The over-abundance of implicated NDD risk genes calls for more sophisticated interpretations of the data, which benefit from pairing multiomic (genome, transcriptome, epigenome, and proteome) and computational approaches (predictive modeling) to better elucidate convergent pathways and underlying disease pathobiology. [35]

1.3.1 Rare inherited and *de novo* variants.

Twin studies were the first indicator of a genetic component in NDDs,[36–40] with heritability estimates of roughly 80% (ASD),[41] 64-81% (SCZ)[42–44] and roughly 30% (epilepsy).[45] Clinical karyotyping, chromosomal microarrays, and genome wide association studies (GWAS) have since identified a number of critical genes and genomic hotspots that are enriched in NDD populations; however, the heterogeneous nature and incomplete penetrance of many of these risk genes suggest that monogenic forms of NDDs are quite rare. Instead, NDDs are more likely polygenic, arising from a combination of common, rare-inherited, or *de novo* variants of the disease-associated

genes. Common genetic variants are present in over one percent of the human population and affect genes with individually small effect sizes. Variants in these genes typically take the form of single nucleotide variants/polymorphisms (SNVs/SNPs) and do not face selective pressure in part because they have a low mutational burden,[46,47] whereas rare variants display more Mendelian patterns of disease inheritance and include more vulnerable genes with a high mutational burden.[48–51] Both rare and common variants can be inherited, whereas *de novo* variants occur in the germline and are specific to the individual. *De novo* variants are often the most penetrant in NDD populations, as they have a high degree of selective pressure,[48] however, they represent roughly two percent of variance liability of NDDs like ASD,[52] suggesting that inherited mutations are more common. Outside of SNPs, genomic variations can include structural variants such as copy number variants (CNVs), [34,53–56] of which a greater burden of *de novo* events is observed in NDD populations such as individuals with ASD.[57]

1.3.2 Copy number variants (CNVs).

Copy number variations are insertions and deletions of contiguous genomic content that range from 50 bp – 3 Mb in size[58] and result in a change in the number of copies in a DNA sequence. Similar structural variations can also include inversions, which do not change copy number of a genomic region, large tandem repeats, which contain repetitive segments of roughly 50 bp, and retrotransposon insertions. It is estimated that 4.8-9.5% of the human genome contributes towards CNVs,[59] and that nucleotide content of CNVs per genome exceeds SNVs in global populations, suggesting the evolutionary role of CNVs.[60] Population studies such as The 1000 Genomes Project have identified structural variation differences between geographically distinct human

populations, suggesting either a genetic drift or evolutionary advantage of CNVs to novel environments.[61] CNVs are formed through misalignment of highly repetitive sequences called segmental duplications during homologous recombination. Expansion of segmental duplications can be traced to a period of early neocortical expansion in our hominid ancestors,[62,63] and due to the evolutionary role of these regions, disruptions in their copy number are often associated with NDDs in a pleiotropic manner i.e., one CNV may confer risk to multiple NDDs. Over 20 NDD-pathogenic CNVs have been identified within the last two decades and can be either inherited or generated *de novo*.[64,65] Generally, NDD-pathogenic CNVs pose a greater disease risk or odds ratio (OR) than common genetic variants by a factor of 10-100,[66,67] in part because of the multiple genes that can be disrupted within a single CNV. Despite the strong pathogenicity of CNVs, there is still incomplete penetrance, suggesting that other polygenic or environmental factors influence neurodevelopmental disorders.[50,68]

1.3.3 The 15q13.3 Microdeletion syndrome.

The long arm of chromosome 15 is one of the least stable regions of the human genome and is highly susceptible to CNVs, such as the 15q13.3 microdeletion (chr15:30,910,306–32,445,407).[69] While duplications occur with a higher frequency, deletions are more highly associated with NDDs and are classified as a major genetic risk factor for epilepsy (idiopathic generalized epilepsy, accounting for > 1% of epilepsy cases, odds ratio of 68[70,71]), SCZ (odds ratio of 11), ASD (less common and amongst 10-30% of individuals) and intellectual disability (roughly 50%).[66,72] The instability of the region is in part due to the presence of many low copy repeats clustering into segmental duplications of the *GOLGA* gene, which arose during a human-specific expansion

roughly 0.5-0.9 million years ago.[68,70,73,74] The *GOLGA* repeats are mapped to each of the seven "break points" within the region, of which different combinations can result in the deletion of 1.5-2 Mb of genomic content (**Fig. 1**). Deletions typically present as heterozygous, although homozygous mutations have been reported in roughly 3% of 15q13.3 microdeletion cases and include more severe neurodevelopmental impairments including neonatal encephalopathy.[75,76] This NDD-pathogenic CNV is referred to as the 15q13.3 microdeletion syndrome, a heterogeneous group of disorders ranging from epilepsy, ASD, ID, ADHD, and schizophrenia.[74,77–79] The complexity of these clinical presentations has made treatment difficult, often targeting the core symptoms rather than the underlying etiology of the disorder.

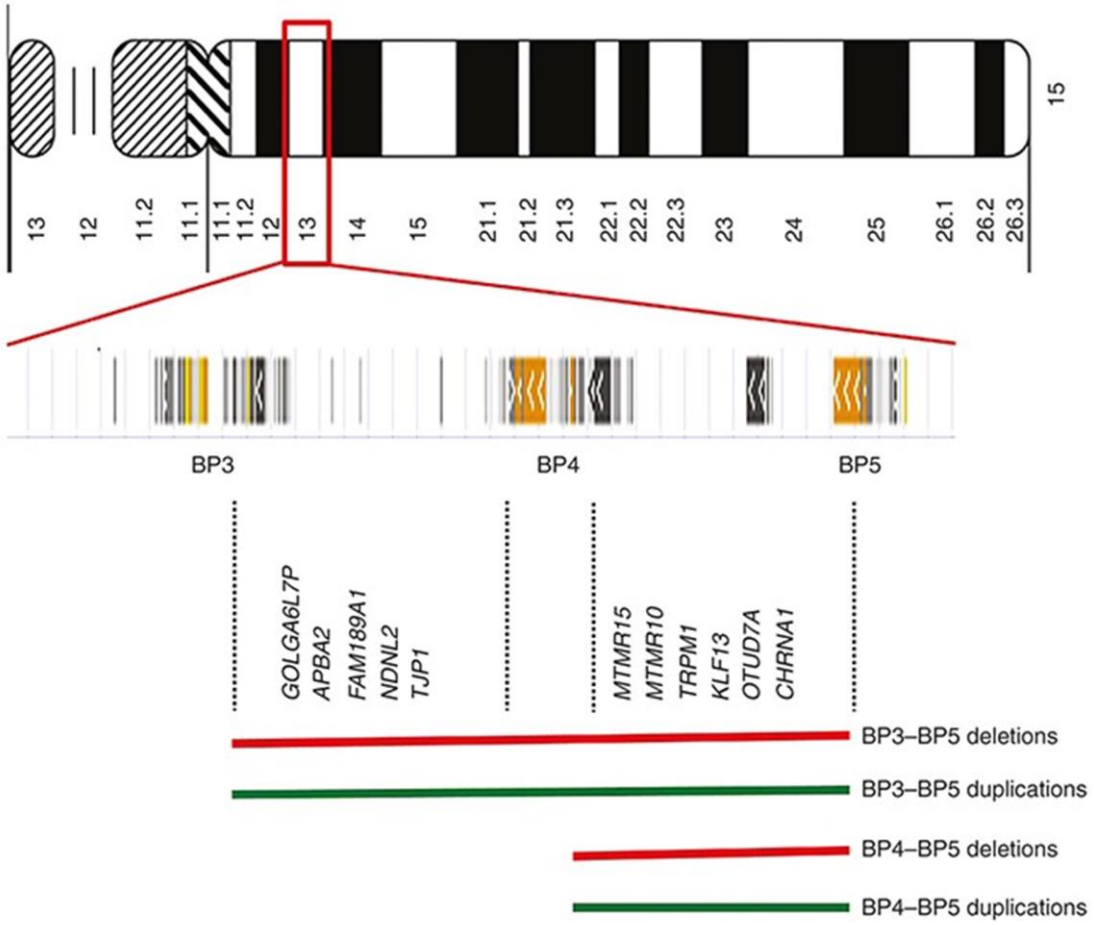


Figure 1: Genomic breakpoints of the 15q13.3 locus.

Schematic diagram showing the location of the human 15q13.3 locus and affected genes within breakpoints (BP) 3-5 highlighted in green or red to show duplications and deletions, respectively. Adapted from *Ziats et al., Genetics in Medicine 2016.*[80]

To delineate the pathophysiology of the disorder, researchers have sought to identify so called “driver genes” within deletions that contribute strongly to disease phenotypes, with multiple candidates emerging over the years. Within the deletion locus are several protein-coding genes, including: *ARHGAP11B, TRPM1, FAN1, KLF13*, *MTMR10, OTUD7A, CHRNA7,* as well one micro-RNA, *microRNA-211,* and two putative pseudogenes, *LOC100288637* and *LOC283710*. Of these genes, nested deletions have been reported to encompass only *CHRNA7* and *OTUD7A* and have influenced their candidacy as driver genes underlying the 15q13.3 microdeletion syndrome.

1.3.3.1 CHRNA7

CHRNA7 was first established as a putative driver gene due to its role in mediating synaptic transmission and potential contribution to the epileptic presentations in 1513.3del patients. The gene encodes the α7 nicotinic acetylcholine receptor (nAChR), a ligand-gated ion channel that is stimulated by acetylcholine (and choline prenatally) to result in the flux of sodium, calcium, and potassium cations. The resulting membrane depolarization activates voltage-dependent calcium channels to induce calcium release from organelles such as the endoplasmic reticulum.[81] α7 nAChRs are located pre-, post-, and extra-synaptically, and depending on the cell type expression they can facilitate the release of neurotransmitters such as glutamate, GABA, acetylcholine, and dopamine. *CHRNA7* variants have been identified in NDD populations such as ASD, epilepsy, ADHD, Rett Syndrome, and schizophrenia, as well as in neuropsychiatric disorders (namely bipolar disorder) and neurodegenerative disorders (Alzheimer’s disease). Importantly, treatment with α7 nAChR agonists and positive allosteric modulators (PAMs) have been shown to have clinically relevant improvements in NDD populations,

including increases in cognition and reductions in negative symptoms in schizophrenic populations.[82,83]

Recent evolutionary expansion of the *CHRNA7* gene has resulted in partial duplication and insertion of exons 5-10 into a region 1.6 Mb upstream, causing the interruption of two other genes. The fusion gene product, *CHRFAM7A*, participates in α7 subunit assembly to negatively regulate channel function. This fusion gene product and regulation of *CHRNA7* is a human-specific event and cannot be recapitulated in other mammalian models, which may underly the mild synaptic and absent behavioural abnormalities seen in *Chrna7* KO mice. In addition, studies have shown that *RIC3*, a chaperone required for biogenesis and trafficking of a7 nAChRs, functions differently in human *vs.* mouse models; only the latter appears to be affected CHRNA7 concentration-dependent, further suggesting human-specific properties of *CHRNA7* function.[84]

1.3.3.2 OTUD7A

OTUD7A (OTU deubiquitinase 7A) is a deubiquitinating enzyme from the ovarian tumour protein family whose expression is limited to the brain and testis. Of all the genes in the 15q13.3 locus, *OTUD7A* exclusively possesses brain-critical exons, and is least tolerant to loss of function mutations.[56] *OTUD7A* has more recently been proposed as a driver of the 15q13.3 deletion, with previous work in our own lab demonstrating its ability to regulate and rescue the dendritic spine and neurite deficits seen in human and mouse models of the deletion.[56,85] Multiple publications have shown that ASD- and SCZ-associated variants in *OTUD7A* likewise produce dendritic spine deficits, indicating its broad role in NDDs. Despite this, the neural function of *OTUD7A* is not fully characterized. Proximity proteomics revealed localization of OTUD7A in the

dendritic spines and axons of glutamatergic mouse neurons, as well as interactions with a key regulator of the axon initial segment, Ankyrin G (ANK3/AnkG).[85] Molecular studies in *Otud7a* KO mice have been shown to recapitulate the dendritic abnormalities as well as behavioural deficits of 15q13.3 mouse (Df(h15q13)/+) and human models, suggesting its candidacy as a primary driver gene for the deletion.

1.3.3.3 KLF13

KLF13 encodes a Krüppel-domain zinc finger (ZnF) transcription factor that targets GC-rich sites and CACCC boxes in the genome.[86,87] *KLF13* function is best characterized in the heart, where it serves as a transcriptional activator of many genes essential for cardiac differentiation and morphogenesis. *KLF13* variants are associated with congenial heart defects with considerable penetrance,[88–91] and KD of the protein in *Xenopus* models have major heart malformations.

Outside of the heart, *KLF13* is one of the few 15q13.3 genes expressed in the fetal mouse ganglionic eminence, where it has been shown to play a role in mediating the proliferation of CGE and MGE-derived interneuron precursors.[92] Using mouse hippocampal primary neurons and CRISPR-edited human cell lines, KLF13 has also been identified as a mediator of NPC proliferation, axonogenesis, and transcriptional regulation of the JAK/STAT pathway.[93–95] Amongst oligodendrocytes, KLF13 interacts with another Krüppel-domain transcription factor, KLF9, to activate known oligodendrocyte-specific regulatory domains and drive the differentiation and myelination of primary oligodendroglial cultures. This is supported by a reduction in myelin gene expression in early KLF13-deficient mouse pups. Despite these findings, the broad function of KLF13 in the brain has yet to be assessed and is known to change on a

cell-type and developmental timepoint specific basis. The dynamic nature of this transcription factor calls for a more detailed assessment of its role in the brain.[96]

1.3.3.4 FAN1

Fanconi-associated nuclease 1, or *FAN1*, encodes a DNA repair enzyme that helps preserve DNA stability in the CNS and in multiple peripheral tissues. The enzymatic activity of FAN1 is quite broad, as it contains both exo- and endonuclease activity to target DNA structures that occur as intermediates during DNA repair, including displacement-loops (D-loops) during homologous recombination (HR) and 5' flap branch structures during interstrand crosslinking (ICL).[97,98] These structures can form in response to DNA damage, and their prolonged presence can stall replication forks and hinder DNA repair. *FAN1* therefore functions to cleave these structures to promote replication fork recovery and enable repair to continue. This role is particularly evident in *FAN1* knockdown (KD) studies, which show an increased sensitivity to drugs that induce ICL and DNA damage, as well as mild chromosomal instability.[97,99]

The contributions of FAN1 to neurodevelopment have only recently been made apparent. Exome sequencing of NDD populations have identified *FAN1* as a risk gene for ASD and SCZ,[100–102] and more recently for repeat expansion diseases like Huntington's Disease and Fragile X.[103,104] FAN1's ability to slow repeat progression suggests it may have other unexplored roles in disease pathology.

1.3.3.5 TRPM1

TRPM1 is a visual transduction ion channel that is part of the melanoma-related transient receptor potential (TRPM) subfamily. Protein expression is highly enriched in the retina, and it may play a role in melanin synthesis in melanocytes. Mutations in *TRPM1* are associated with an autosomal recessive form of complete congenital stationary night blindness, supporting the role of *TRPM1* as a driver gene behind the visual impairments seen in homozygous 15q13.3del patients.

1.3.3.6 ARHGAP11B

ARHGAP11B encodes the human-specific protein Rho GTPase Activating Protein 11B, which emerged through partial duplication of the *ARHGAP11A* gene during early hominid evolution. The emergence of *ARHGAP11B* is thought to coincide with early neocortical expansion in human evolution, and indeed expression of the protein increases cortical size and folding in mouse and marmoset models,[105] whereas loss of ARHGAP11B in a human neural model results in reductions in basal radial glia populations.[106] ARHGAP11B is highly localized to the mitochondria, and its regulation of basal progenitor cells appears to be conditional upon the breakdown of glutamine to glutamate (glutaminolysis).[107,108]

1.3.3.7 Arguments for using human models to study the 15q13.3 microdeletion syndrome.

The expansion of the human neocortex marks an evolutionary divergence that occurred between 7-9 million years ago and has resulted in a threefold increase in brain size and a two-fold increase in total neuron populations, relative to that of the chimpanzee.[109,110] This increase in size is attributed to an increase in NPC proliferation

including the outer radial glial (oRG) present in the outer subventricular zone. oRGs are enriched in higher order primates and are thought to influence cerebral cortex size and folding (**Fig. 2**). [111] As mentioned previously, the proliferation of these cells is heavily influenced by the human-specific 15q13.3 gene, *ARHGAP11B*, and both this cell type and gene are completely absent in the murine model (Df(h15q13)/+) traditionally used to study the 15q13.3 microdeletion syndrome. The formation of the CNV itself relies on human-specific expansion of *GOLG8* repeats, creating incredibly complex genomic architecture that is not captured in the isogenic deletion model targeting mouse chromosome 5.[68,70,73,74] Indeed, Df(h15q13)/+ mice lack human/primate specific cell types (oRGs, primate-specific neuron types in deep layers IV[110]) and also do not mimic the unique and variable CNV size and composition that contribute towards the NDD heterogeneity. Human models therefore offer a unique opportunity to study the microdeletion while still retaining the unique (albeit heterogeneous) changes in genomic architecture caused by the CNV.

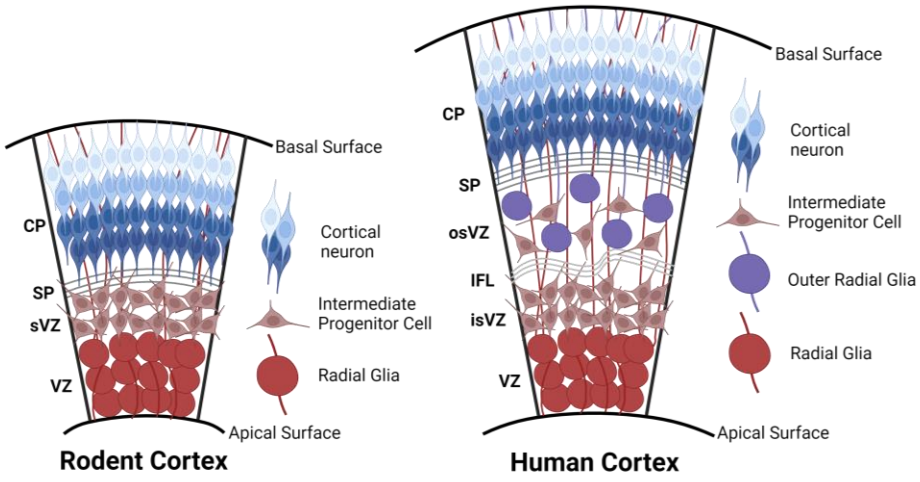


Figure 2: Structural comparison of rodent and human cortices.

1.4 Tools and limitations to studying NDDs.

1.4.1 Mouse models of NDDs.

Murine models are one of the first established mammalian models to study neurodevelopment and its disorders. Orthologous or syntenic regions can be manipulated in the mouse genome to provide a reliable, isogenic background to study various genetic abnormalities, which can then be assessed throughout mouse development. An early example is the *Fmr1* KO mouse, which was generated over 25 years ago to possess a deletion in the *Fmr1* gene to model Fragile X syndrome, a monogenic form of intellectual disability commonly associated with ASD. Using this model, researchers were able to shed light on sensory processing abnormalities that are found both in the mouse and human, which are now a diagnostic tool for diagnosing ASD.[112,113] CNVs have also been modeled in the mouse background to study the polygenic contribution of genes *in vivo*, and include 16p11.2 deletions, 15q13.3 microdeletions (Df(h15q13)/+), and the 22q11.2 microdeletion.

1.4.2 The Df(h15q13)/+ mouse model.

The ability to study CNVs in a living system allows researchers to investigate molecular, behavioural, and neurophysiological abnormalities throughout development. The Df(h15q13)/+ mouse is a genetic model of the 15q13.3 microdeletion with the orthologous deletion at chromosome 7qC and has been studied extensively at postnatal periods of development.[56,114,115] These mice recapitulate epileptic presentations found in patient cohorts, showing an increased propensity for myoclonic and absence seizures, and

a decrease in clonic and tonic seizures during acute seizure assays. Abnormal neural activity has also been reported within specific brain regions including the hippocampus and the cerebral cortex[116,117]. Population-wide functional assays in 2D cortical cultures have also demonstrated reductions in spontaneous neural activity and bursting patterns, further indicating disruptions to neural circuitry.[118]

Changes in the proportions and morphology of multiple cell types have been identified in the Df(h15q13)/+ mouse, including reductions in PV^+ interneurons in layers 2/3 of the somatosensory cortex,[119] and reductions of astrocytes in the medial prefrontal cortex.[117] While no proportion changes have been characterized in cortical excitatory neurons, there have been persistent reductions in their dendritic spine density and maturation morphology across multiple independent studies.[56,117,118]

1.4.3 iPSC and organoid modeling of NDDs.

For years, our understanding of NDDs was limited by inaccessibility of neural tissue at relevant periods of neurodevelopment, which has led to a poor understanding of the pathobiology of many of these disorders and in turn, inadequate treatment. Advancements in human disease modeling have led to the creation of induced pluripotent stem cells (iPSCs), which can be generated from somatic patient cells, typically blood or fibroblasts, and subsequently differentiated into complex cell types that include neural tissue.[120] The retention of genetic background in patient lines enables researchers to assess for patient- and NDD- specific mechanisms in a disorder, and when powered by a clinical cohort of individuals with the same genetic abnormalities, can provide insight in how these mechanisms can be explored for therapeutic treatment.

1.4.3.1 Human models of the 15q13.3 microdeletion syndrome

Presently, there are three studies that have characterized the 15q13.3 microdeletion in a patient background. Two of these studies utilized the NGN2-overexpression system, [118,121] which produces a homogeneous population of glutamatergic-like excitatory neurons through overexpression of the transcription factor Neurogenin-2 (NGN2). In contrast, the third study examined NPCs derived from both deletion and duplication patients using a dual-SMAD inhibition protocol[122]. NPCs from both deletion and duplication patients were found to have a decrease in CHRNA7-mediated calcium flux, a measure of synaptic activity and intracellular signaling. The authors went on to show a significant decrease in CHRNA7-specific chaperoning proteins amongst the 15q13.3del NPCs, whereas the duplication lines showed the opposite trend.[122]

Activity has since been assessed in mature NGN2 neurons, where our team demonstrated abnormal population-level spontaneous activity using multielectrode array recording across time. We also characterized morphological irregularities in these cells, where we found reductions in dendritic arborization, disruptions to the nanodomain organization of the axon initial segment (AIS), and I identified deficits in axon projection capacities amongst three 15q13.3del families.[118] The final and most recent assessment of 15q13.3del NGN2 neurons includes an impressive bulk RNA, methylation, and ATAC sequencing experiment on 15q13.3 neurons (all omics) and of isogenic 15q13.3 gene KO lines for *FAN1*, *KLF13*, *OTUD7A*, and *MTMR10* (bulk transcriptomics only). In this study they identified differential methylation of a protocadherin encoding region on chromosome 5, broad changes in chromatin accessibility, and increased DNA damage,

which they validated by showing an increased sensitivity to cisplatin-induced DNA damage.[121]

1.4.3.2 Organoids as a model system for neurodevelopment

iPSCs can also be used to create brain organoids, which are self-organizing 3D neural structures comprising stem cells, NPCs, and terminally differentiated neurons capable of modeling early periods human neurodevelopment. Gene co-expression analyses have shown a high degree of similarity between neural organoids and fetal brain tissue, producing transcriptional programs similar to post-conception weeks 9-35.[123–129] For this reason, brain organoids offer unprecedented access to study human neurodevelopmental processes such as neurogenesis, cellular differentiation, and circuit integration in a 3D context.

Brain organoids are classified based on the directedness of differentiation, either as unguided neural organoids (UNOs, formerly known as “cerebral” organoids), or guided neural organoids (GNOs),[130] which are directed towards brain-region specific lineages.[131] UNOs spontaneously differentiate into multiple regions of the developing telencephalon, expressing markers for forebrain, midbrain, hindbrain, hippocampal, choroid plexus, meningeal, and retinal cell identities. This differentiation is achieved through intrinsic signaling and is preceded by the formation of NPC-rich rosette structures reminiscent of the human cortical ventricular zone. Newborn neurons will emerge from the rosettes and migrate outwards to form laminar structures reminiscent of the developing cortex. After 2.5 months of growth. the resulting postmitotic neurons are mostly excitatory (roughly 50%), but that percentage decreases in half by 4 months, where subpopulations of include inhibitory neurons, astrocytes, oligodendrocytes, and

even retinal cell types increase.[127,132–134] The cell type diversity within mature UNOs is a key advantage in modeling cell type emergence and differentiation in disease backgrounds.

Unlike UNOs, GNOs require the use of exogenous signaling and patterning factors to direct the organoid towards specified cell fates and brain regions, including dorsal (pallium) and ventral (subpallium) forebrain, thalamus, hypothalamus, striatum, cerebellum, brain stem, and spinal cord.[131,135,136] Through directed differentiation, the system becomes more unified, and batch-to-batch heterogeneity is greatly reduced. Importantly, multiple GNOs can be combined to create assembloid systems to assess the neuronal circuitry between different brain regions. In this way, GNOs offer a unique opportunity to study circuitry dynamics including neuronal migration, cellular integration, and axonal innervation.

Organoid disease modeling has steadily increased over the years to include neurodegenerative disorders like multiple sclerosis and Alzheimer's disease, as well as viral and chemical infections like Zika virus and chronic nicotine exposure.[129,137,138] Due to the transcriptional similarity between brain organoids and fetal development, however, they are more commonly used to model NDDs such as ASD, epilepsy, Down syndrome, Tuberous Sclerosis, and Angelman syndrome [139–142].

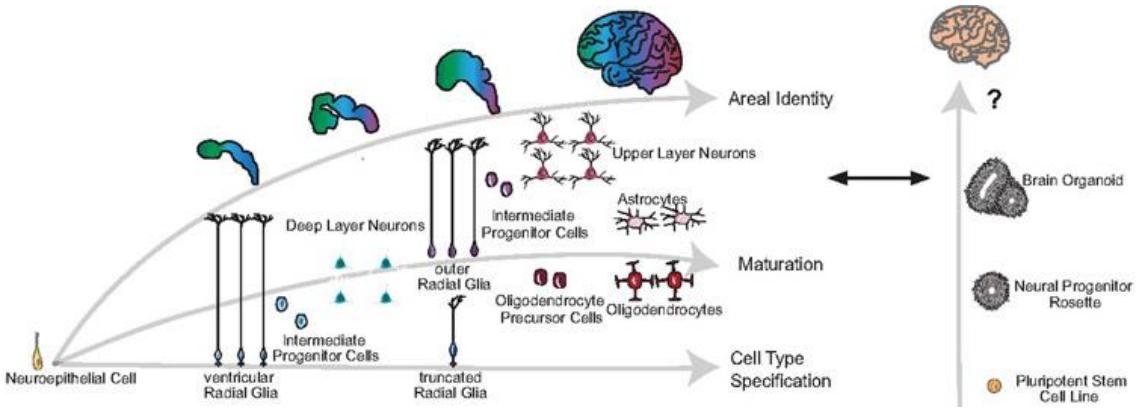


Figure 3: Mapping neurodevelopmental trajectories of human neural organoids to fetal development.

Ordered developmental milestones of cell type emergence, migration, and maturation in the human brain can be modeled using human iPSC systems. Adapted from *Bhaduri et al., Nature 2020.*

1.4.4 Bulk transcriptomic approaches to assess developmental trajectories.

Two of the major barriers to NDD therapeutics are a lack of established biomarkers and a poor understanding of disease trajectory. Bulk RNA sequencing (bulk RNA Seq) provides a scalable tool to examine the transcriptome of aggregate tissue and has evolved over the years to enable users to assess copy number variations, transposable elements, and cell type deconvolution (albeit with limitations). Depending on the tissue processing and chosen read depth, gene transcription can be assessed broadly for differential gene expression (DGE) or with isoform-specific resolution by capturing differential splicing of mRNA. A subset of the differentially expressed genes (DEGs) can then be assessed for over-representation of genes for a pathway of interest.

Gene set enrichment analysis (GSEA) is an alternative analysis pipeline for examining gene expression. Instead of providing a list of DEGs, the entire gene set is ranked and provided a score according to its gene significance and fold change (incorporating the direction of up or downregulation). An algorithm is then applied to determine first if there is enrichment amongst the genes for a biological pathway (incorporating the gene's score), and then assessed on how the genes are distributed in the ranked list. GSEA determines if there is a significant enrichment for a pathway of interest by "walking" down the ranked list, increasing a running-sum statistic whenever it encounters a gene within the pathway of interest, and decreasing when one is not present. Step-wise or continuous distribution of the enriched genes near the top or bottom of the ranked gene set (most up or downregulated), is generally enriched.[143] An enrichment score (ES) is then calculated to reflect the degree of overrepresentation of the genes,

either at the top or bottom of the list. In this way bulk RNA Seq can identify the affected biological pathways and biomarkers in a gene list that is provided in an unbiased manner, identifying differentially expressed transcripts that may otherwise be missed with conventional phenotyping methods such as immunofluorescence or single gene expression.

Bulk RNA Seq has recently been used to profile the developmental trajectory of human tissue by use of NDD patient-derived brain organoids and 2D cultures.[144] In this way, disruptions to critical periods of development have been identified, including stages of neurogenesis, synaptogenesis, and circuit formation. In addition to identifying critical developmental windows, bulk RNA Seq may provide insight into what early biomarkers may be present in a pre-symptomatic period, providing a means for supporting at-risk children. This would be a critical form of intervention, as there is an increased risk factor for neuropsychiatric illnesses for individuals with ASD diagnosis later in adulthood.[144]

1.4.5 Single cell transcriptomics as a tool to examine cellular composition and communication.

One of the major caveats to bulk transcriptomics is the aggregation of what is often heterogeneous tissue and cell types; the introduction of so much variability can dilute phenotypic effects and render the interpretation of data nearly impossible. Single cell RNA sequencing (scRNA Seq) offers the unique ability to profile individual cell transcriptomes and preserve the diversity of cell types within a sample. While multiple processing pipelines exist, ranging from: fixed frozen, flash frozen, fresh dissociation, single cell vs single nucleus, the core element to sample processing includes the immersion of single cells or nuclei in an oil solution containing reverse transcription

enzymes and oligonucleotide-barcoded beads. Using microfluidic partitioning the cells and beads are added in a 1:1 ratio to ensure each cDNA library is labelled with a barcode unique to each cell. The barcode is then retained during next generation sequencing (NGS) and can therefore be mapped back to the cell of origin. Within the bead itself are tagged nucleotide sequences called unique molecular identifiers (UMIs) which help differentiate between unique mRNA transcripts and therefore allow the user to screen out PCR duplicates and mutational artifacts *vs.* rare variants. For this reason, multiple reads of the same gene, barcode and UMI are collapsed into a single UMI count for downstream processing.

Single cell RNA sequencing analyses are constantly evolving, with predictive platforms like CellChat and pseudotime allowing scientists to infer cell-cell communication and developmental trajectories, respectively, of various cell types. CellChat was introduced by the Satija lab in 2021 to provide predictive modeling of ligand-receptor interactions. An advantage to CellChat is that it applies communication probabilities using the law of mass action, which considers the geometric means of ligand and receptor expressions (with their subunits), weighted by their agonists/antagonists and levelled with DEG analyses. It uses an *in-house* database of known ligand-receptor complexes (including multimeric classes), soluble agonists and antagonists, as well as stimulatory and inhibitory membrane-bound co-receptors to create inferences about the proportion and strength of cell-cell communication.

Another powerful but inference-based analysis in scRNA Seq transcriptomics is pseudotime, which arranges individual cells across a temporal trajectory in order to understand cellular dynamics that govern differentiation and disease.[145,146] Monocle is

one of the earliest pseudotime pipelines, and functions by performing unsupervised clustering of cells based on their differential gene expression along a differentiation trajectory. It reconstructs biological branch sites that are often representative of a new differentiation state, thereby allowing the user to infer lineage trajectories between two conditions. In this way, pseudotime can be used to assess cellular trajectories, the molecular pathways that underlie cell fate decisions, and the temporal dynamics of disease. Importantly, pseudotime data must be rigorously examined, as it can occasionally make inferences that do not reflect true biological trajectories.[147]

1.5 Thesis objectives

The primary objective of my thesis was to characterize human-specific neurodevelopmental deficits in the 15q13.3 microdeletion background using various fetal model systems (3D neural organoids). The project can be further divided into three core aims, which are discussed below. The common tool used in each chapter includes the 15q13.3del patient-derived iPSCs, which we have reprogrammed *in house* and have previously phenotyped a subset of, by using Neurogenin-2 (NGN2)-induced neurons.[118] Aim 1 is addressed in chapter 3 and uses the unguided neural organoid model while applying single cell transcriptomics to a subset of the cohort (3-4 families). Aims 2-3 are discussed in chapter 4 and include the entire 15q13.3 cohort for bulk transcriptomic and circuit phenotyping using guided neural organoid modeling, followed by validations in 1-3 families.

1.5.1 Aim 1: Characterizing vulnerable cell types and cellular trajectories in the 15q13.3 microdeletion cohort.

In our previous study (which I was the fourth author of) we identified synaptic and morphological impairments in 2D NG2N neurons derived from three unique 15q13.3del families. These neurons are generated through overexpression of the transcription factor, Neurogenin-2, to yield highly homogenous glutamatergic-like excitatory neurons. The morphological and functional abnormalities that we observed in the three 15q13.3del families are congruent with the neurodevelopmental deficits observed in the patients, who present with epilepsy, developmental delay, and autism spectrum disorder. However, the NGN2 model system uses an accelerated differentiation protocol that fails to capture early neurodevelopmental processes such as neural progenitor cell emergence, cell fate decisions, and neurogenesis,[148] and therefore presents a gap in our knowledge of 15q13.3del developmental dynamics.

This study focuses on capturing early neurodevelopmental processes that were not characterized in our 15q13.3 microdeletion patient cohort, all of whom received diagnoses postnatally following behavioural tests or epileptic presentations. Specifically, we wanted to examine fetal neurodevelopmental milestones such as the formation of the ventricular zone (VZ) cell type emergence (neurogenesis) and cell maturation. Using unguided neural organoids, we exploited their ability to stochastically generate numerous cell types from multiple brain regions and performed the first examination of human 15q13.3del early fetal brain development across time. Using an established *in house* scRNA Seq analysis pipeline, we captured changes in early progenitor cell populations and their intracellular structures, as well as late-stage impairments to the temporal

trajectory and inferred synaptic signaling of mature excitatory and inhibitory neurons. We found changes in cell adhesion to persist throughout many of these processes, suggesting an underlying mechanism for early developmental impairments.

1.5.2 Aim 2: Characterization of cell-type specific contributions to 15q13.3 microdeletion syndrome progression and pathology.

Given the profound impairments in mature excitatory and inhibitory neurons from our previous study, we wanted to identify the shared and unique abnormalities within each population across development. To achieve this, we generated guided forebrain organoids of dorsal and ventral lineages and used the entire 15q13.3del cohort to perform bulk RNA sequencing across two timepoints in development. By collaborating with Tony Chen, a PhD candidate from Dr. Kristin Hope's lab, we generated a comprehensive dataset of 120 samples across two tissue types and timepoints. Amongst the tissue types, we discovered convergence on cell adhesion and transcriptional dysregulation, which we later found to include several human-specific zinc finger protein loci at 19p13.2 and 19q13.43.

While sharing disruptions in transcription, the dorsal forebrain organoids exclusively showed enrichment for DNA damage and repair pathways, which we confirmed as an excitatory neuron-specific phenomena amongst the UNO excitatory neurons as well. Conversely, the ventral organoids and the UNO inhibitory neurons population converged on disruptions to neuron migration, a phenomenon that has never been modeled or assessed in a 15q13.3del background. This study represents the first examination of both excitatory and inhibitory neurons in a human-specific background

and uses the largest clinical 15q13.3del cohort in reported literature. Using this system, we have uncovered convergent and cell-type specific abnormalities in early neurodevelopment, necessitating the need for circuit exploration in assembloid models.

1.5.3 Aim 3: Applying assembloid models to assess 15q13.3del circuit abnormalities.

The objective of the last aim was to examine 15q13.3del circuit dynamics in a dorsal-ventral forebrain assembloid model to better understand the consequences to cell type-specific abnormalities. We performed tissue clearing and light sheet microscopy on fluorescently tagged forebrain assembloids using the same three 15q13.3del families from **aim 1**. We exploited different lineage (D *vs.* V) and genotype (WT *vs.* HET) combinations to explore cell intrinsic *vs.* extrinsic contributions to cell migration and morphology. We found persistent alterations in interneuron neurite length in all but the WT-WT combination, however, we only saw migration impairments in combinations that contained a 15q13.3del ventral organoid, suggesting that cell autonomous effects have some interplay with the dorsal microenvironment. The findings from this study and those above are currently being prepared for submission pending the analysis of dorsal-ventral spatial sequencing (samples submitted) and patch clamp electrophysiology phenotyping in collaboration with Dr. Zahra Dargaei.

2 CHAPTER 2: MATERIALS AND METHODS

The techniques used in this report are a combination of stem cell culturing, transcriptomics, and microscopy (traditional immunocytochemistry and light sheet), the details of which can be found below.

2.1 hiPSC Reprogramming and maintenance.

2.1.1 Patient sample collection and Cell Oversight Committees.

All pluripotent stem cell work was approved by the Canadian Institutes of Health Research. Blood samples were collected from individuals with the approval from SickKids Research Ethics Board after informed consent was obtained (REB approval file #1000050639). Additionally, this study was approved by the Hamilton Integrated Research Ethics Board, (REB approval file #2707). Fibroblast and blood samples were acquired from locations globally and are outlined per family/sample (**Table 1)**.

2.1.2 Patient cohort

Patients heterozygous for the 15q13.3 microdeletion were approached at sites across Canada, the United States, and Italy. Microarray data determined that the deletion sizes ranged from 1.5-2 Mb in size, which is within the typical range of the deletion. Clinical assessments were provided upon request and revealed a heterogeneous population of individuals with ASD, intellectual disability, and epilepsy (most often presenting with absence seizures). Information on clinical background and deletion size to our cohort can be found on **Table 1**.

Table 1: Clinical presentations and background of patients heterozygous for the 15q13.3 microdeletion.

Table 1

Patient ID	Sex	Proband relation	Clinical presentations	Tissue collection and mode of reprogramming	Deletion size (kbp)
Family 1 control	Female	Mother	Neurotypical	PMBC isolation followed by Sendai viral reprogramming	No 15q breakpoints
15q13.3del Patient 1	Female		ASD		(1530.69)
Family 2 control	Male	Brother	Neurotypical		No 15q breakpoints
15q13.3del Patient 2	Female		Epilepsy (absence seizures), DD, ID		2000.669
Family 3 control	Male	Father	Neurotypical		No 15q breakpoints
15q13.3del Patient 3	Female		Severe epilepsy (absence seizures), ADHD, ASD		1348.142
Family 4 control	Female	Mother	Neurotypical		No 15q breakpoints
15q13.3del Patient 4	Female		Epilepsy (myoclonic absence seizures), abnormal EEG and MRI, thick corpus callosum		Microarray data not available
Family 5 control	Female	Mother	Neurotypical		No 15q breakpoints
15q13.3del Patient 5	Female		DD, Thick corpus callosum, T2 hyperintensity in tegmental tracks, language, and motor delay		Microarray data not available
15q13.3del Patient 6	Female	No familial control	Global DD and ID		Microarray data not available

ADHD - attention deficit hyperactivity disorder, ASD - Autism spectrum disorder, DD - developmental delay, EEG – Electroencephalography, ID - Intellectual disability, kbp - kilobase pairs.

Importantly, discovery of the CNV in each patient was well after the first year of life, leaving a developmental window uncharacterized for this disorder. We sought to profile the developmental continuum of the 15q13.3 microdeletion using unguided neural organoids, which are capable of generating multiple telencephalic brain regions and cell types from the developing fetal brain in three-dimensional space, providing an unbiased model that can also be probed for relevant cytoarchitectural changes across development.[130,134,149–151]

2.1.3 Reprogramming and hiPSC culturing.

$CD34^+$ blood cells were assessed via flow cytometry and collected for iPSC reprogramming. iPSCs from Family 1 and 2 were reprogrammed by the Centre for Commercialization of Regenerative Medicine (CCRM, MaRS Centre, Toronto, ON) and iPSCs from families 3-6 were generated *in-house*. iPSCs were generated by Sendai viral reprogramming and clonal expansion using the CytoTune – iPSC 2.0 kit (Thermo Fisher) according to manufacturer instructions. In brief, colonies were expanded for 2-3 weeks, and once large enough were transferred to a single well of a 12-well plate coated with irradiated mouse embryonic fibroblasts (MEFs). iPSC media consisted of DMEM/F12 supplemented with 10% KO serum, 1× non-essential amino acids, 1× GlutaMAX, 1 mM β-mercaptoethanol, and 16 ng/mL basic fibroblast growth factor. Once expanded, iPSCs were grown on Matrigel-coated 6-well plates in mTeSR1 (STEMCELL Technologies), where all subsequence passaging was performed with ReLeSR (STEMCELL Technologies). Line validation was performed through flow cytometry and immunocytochemistry of pluripotency markers, and cytogenomic karyotyping was assessed via G-Banding analysis at The Centre for Applied Genomics (TCAG, SickKids, Toronto, Canada). Data for hiPSC validations can be found in **Supplementary Fig. 1**.

2.2 Organoid generation and tissue handling

2.2.1 Organoid and assembloid generation

Unguided and forebrain-specific organoids were generated with the STEMdiff ™ Cerebral Organoid and the STEMdiff ™ Dorsal/Ventral Forebrain Organoid kits, respectively, from STEMCELL Technologies ™ using low-passage (P20-P30) hiPSCs under the manufacturer instructions.

Assembloids were generated from dorsal and ventral forebrain organoids on day 25 of differentiation; using a wide-bore P10 pipette, individual organoids from dorsal and ventral lineage were placed into one ultra-low attachment U-bottom 96 well plate (Corning) in 200 µL of media. Plates were then centrifuged for 15 seconds on the "Quick spin" setting. A half media change was performed the following day, and on day 4 assembloids were moved into a 6 well ultra-low attachment plate (Corning) using a wide-bore P1000 tip. 25 days post-assembly (day 5 of differentiation), maturation media was supplemented with 1 µg/mL of doxycycline to induce expression of fluorophore constructs. Induction proceeded for five days, after which the assembloids were processed for immunocytochemistry or light sheet microscopy.

2.2.2 Organoid BrdU pulse, fixation, and immunocytochemistry

On day 40 of differentiation, unguided neural organoids media was supplemented with 0.1 mM of BrdU (Sigma) and left to incubate for 24 hours. Following the 24 hours, organoids were fixed and processed for immunocytochemistry.

Organoids were washed 3x in Phosphate Buffered Saline (PBS) for 10 minutes prior to a 48-hour fixation in 4% paraformaldehyde (PFA). Samples were dried in 30% and then 50% sucrose in PBS overnight, and then embedded in either Optimal Cutting Temperature (OCT) (1-month unguided neural organoids) or gelatin (all other samples) and stored at -80°C. Sections were made on the Leica CM3050S at 20 µm on SuperFrost™ Plus microscopy slides and stored at -30°C.

The slides were washed 3x in 1x Phosphate Buffered Saline supplemented with Tween-20 (PBS-T) and incubated in a blocking/permeabilization solution comprised of

10% Donkey serum (Millipore) and 0.03% Triton X-100 (Fisher Scientific) in PBS for 1 hour at room temperature. Samples were incubated in this solution with primary antibodies overnight at 4℃. Samples were washed 3x with 1x PBS-T, followed by a 2-hour incubation at room temperature with secondary antibodies. Samples were washed 2x with 1x PBS-T, incubated with DAPI nuclear dye for 15 minutes, followed by a final wash with 1x PBS-T. Samples were mounted on Vista Vision glass microscope slides (VWR) using Prolong Gold antifade reagent (Life Technologies). Antibody sources and dilutions can be found in **Table 2**.

Table 2: List of antibodies used in this project.

Target	Host	Manufacturer	Working dilution	Comments (what it was used on)
BrdU	Rat	Abcam (Cat. ab6326)	1:500	Proliferation assay
CTIP2	Mouse	Abcam (Cat. ab233713)	1:500	D90 UNO IHC
GFP	Chicken	Thermo Fisher (Cat. A-11122)	1: 5000 3D, 1: 1000 2D	Assembloid IHC (3D) and NGN2 axon staining (2D)
Ki67	Rabbit	Abcam (Cat. ab15580)	1:500	Proliferation assay
MAP2	Chicken	Abcam (Cat. ab92434)	1:500 3D 1:1000 2D	Assembloid IHC (3D) and NGN2 axon staining (2D)
mCherry	Goat	Thermo Fisher (Cat. PA5-143590)	1:2000	Assembloid IHC
NKX2-1	Rabbit	Abcam (Cat. ab76013)	1:500	Ventral organoid IHC
SMI-312	Mouse	BioLegend (Cat. 837904)	1:1000	NGN2 axon staining
SOX2	Goat	R&D Systems (Cat. AF2018)	1:100	UNO and dorsal IHC
SST	Mouse	Abcam (Cat. ab140665)	1:500	Ventral organoid IHC
TBR1	Rabbit	Abcam (Cat. ab183032)	1:1000	UNO and dorsal IHC
RedDot	N/A	Biotium (Cat. 40060-1)	1:100	Light sheet imaging

2.2.3 Confocal microscopy and ICC cell population analyses

Confocal Z-stack images were taken on a Zeiss LSM 800 with a 2048x2048 resolution. Images were processed with ImageJ 1.44 Software (Fiji) and analyzed using CellProfiler 4 (BMC Bioinformatics). For population analyses, manual regions of interest (ROIs) were drawn, and maximum intensity projections of each channel were processed and thresholded to report cell populations normalized to $DAPI^+$ nuclei (unless otherwise stated).

2.2.4 Tissue clearing (CUBIC Protocol)

Fixed samples were processed using the CUBIC protocol[152] with minor modifications based on organoid and assembloid size. Optical clearing reduces the refractive index (RI) differences between different tissue components to allow the passage of light without scattering and with minimal refraction. The removal of highly refractive components such as lipid is achieved by using detergents and is followed by RI homogenization. Tissue with a singular RI can then be submerged in an oil solution with an identical RI and imaged with confocal or light sheet microscopes.

The CUBIC protocol (Clear, Unobstructed Brain Imaging Cocktails and Computational Analysis) was chosen as it has been heavily optimized for neural tissue,[152] and for its ability to preserve fluorescent molecules with minimal quenching (as compared to harsher protocols such as iDISCO). The iDISCO method was previously attempted (**data not shown**), however the endogenous fluorophore signal was completely quenched by the detergents, rendering it incompatible with our model system.

Fixed assembloid samples were incubated overnight at room temperature with the nuclear stain RedDot 647 (Biotrend, 1:100, shaking RPM), followed by 3x PBS washes. Samples were incubated in 50% CUBIC-1 in H_2O for 2 hours at 37℃, followed by a 48-hour incubation with 100% CUBIC-1 at 37℃. Samples were washed 3x with PBS for 30 minutes, and then incubated in 50% CUBIC-2 for 3 hours, followed by an overnight incubation in 100% CUBIC-2 at 37℃. Samples were washed in type FF immersion oil (16212, Cargille Laboratories Inc. Cedar Grove, NJ, USA) for 10 minutes at room temperature prior to imaging.

Confocal imaging was performed using the Ultramicroscope II light sheet microscope (LaVision BioTec GmbH, Bielefeld, Germany) at 2x magnification to capture the entire assembloid. Images were processed using IMARIS software (V10.0.0, Bitplane Inc. Zurich, Switzerland) and the proportion of hSYN-eGFP$^+$ ventral neurons was quantified using the Spots tool and Chi sum of squares statistical test for cell nuclei identification. We chose to report the proportion of cells migrating rather than raw values to mitigate confounding variables caused by differences in labeling efficiencies between cell lines. The NeuriteTracer package was then used to measure the neurite length of migrating hSYN-GFP$^+$ neurons.

2.3 Transcriptomics

2.3.1 Bulk RNA Sequencing processing

Organoid tissue was processed on days 50 and 100 of growth in vitro, unless otherwise specified. 3-5 organoids were combined per sample, and RNA was extracted

from 3 samples/patient line using the RNeasy mini kit and RNase-Free DNase set (Qiagen). RNA integrity was assessed with the 2100 Bioanalyzer (Agilent Technologies, Inc.), which performs chip-based capillary gel electrophoresis to calculate the RNA integrity number (RIN) based on proportions of the total 18S and 28S ribosomal RNA. Low RIN values (< 4) can indicate RNA degradation, and while there is no universal cutoff, we ensured all samples had RIN values above 8 before proceeding with library prep (**Supplementary Figures 10-11**). All samples of a tissue type were processed simultaneously at TCAG (SickKids, Toronto, Canada). with NEBNext Ultra II Direction RNA Library Prep kit for Illumina (New England Biolabs), which specifically enriches for protein-coding, polyadenylated messenger RNA (mRNA) fragments. mRNA is captured with purification beads and fragmented to 150 base pairs (bp) for cDNA synthesis. cDNA fragments are then ligated to adapters that enable library amplification through polymerase chain reaction (PCR) for downstream next generation sequencing (NGS). Stranded RNA was chosen for amplification given the increase in accuracy for transcript expression estimation[153]. The samples were run through a single lane on the Illumina NovaSeq 6000 S4 flow cell, which recognizes adapter sequences and performs NGS at specified read depths. A target capture of 30 million reads/sample was performed, and raw data was further processed using the CutAdapt package to trim residual adapter content, and the STAR2Pass package to align each read using the human genome GRCh38 as a reference. Quality control was performed using the FastQC platform, which enables high-throughput assessment of read quality using metrics such as read length, depth, Phred scores,[154] and adapter contamination.

2.3.2 Bulk RNA Sequencing Analysis

Read integrity was assessed on raw RNA-seq data by using FastQC (V0.11.5), followed by adapter trimming and filtering of low-quality reads via CutAdapt. Reads were mapped to hg38 using STAR (2-pass mode, V2.7.0A - outSAMtype BAM SortedByCoordinate) using default settings and aligned read content was log2 normalized and assessed for differential gene expressed (DGE) through the RStudio package, DESeq2.

Additional filters were applied to remove genes with low read counts (<10 reads in ≥50% of the samples) to reduce technical noise and improve DEG detection sensitivity. Normalization and differential gene expression (DGE) analysis was performed using the DESeq2 platform created by Dr. Michael Love (V1.38.3). We applied the covariate regression function to regressing out batch effects, sex, and ancestry to minimize unnecessary variability in the dataset. This variance is especially evident in the principal component analysis (PCA) of the raw, unprocessed data, which grouped loosely by family followed by timepoint (**Fig. 21, 25**).

DESeq2 uses negative binomial distribution estimates to determine differentially expressed genes (DEGs) within a dataset. First it applies normalization or “size factors” to the raw gene counts to account for differences in library depth between samples. It then estimates dispersions, which are inversely related to the mean expression and directly related to the variance; it operates under the assumption that genes with similar expression values have comparable dispersion and allows for shrinkage of the estimates to enable more accurate gene count modeling. DESeq2 then fits the negative binomial model and uses the Wald test for hypothesis testing. Importantly, the analysis pipeline

allows users to regress out covariates that may skew the data such as batch effects, sex, and ancestorial to minimize unnecessary variability in the dataset. After regressing out timepoint, sex, age of individual, and tissue type covariates, genes with an FDR <0.05 and pAdj <0.05 were deemed differentially expressed. Data was plotted using BioConductor and ggPlot2 packages in RStudio. DGE analysis was performed using ShinyGO (V0.8), a bioinformatic mapping tool that performs gene ontology (GO) analysis, chromosomal distribution mapping, and KEGG pathway enrichment for a gene input list with a custom background to overcome selection bias.[155] GO analysis was performed using default settings, selecting pathways within the range of 2-2 000 genes and FDR <0.05 (based on the nominal p-value from the hypergeometric test). Our total matrix gene set post-filtering was used as a background to yield more accurate results.

2.3.3 Gene set enrichment analysis (GSEA)

Gene set enrichment analysis was performed using the GSEA desktop application (http://software.broadinstitute.org/gsea/downloads.jsp)[143] on all genes ranked by fold change. The default parameters used were 1 000 permutations, a minimal set size of 30, and a maximum set size of 500 to account for the large DEG datasets. Gene sets with FDR <0.05 were considered significant.

2.3.4 Cell dissociation for single cell RNA sequencing (scRNA Seq)

Neural organoids beyond 3 months of differentiation were cut first and washed 1x with PBS to remove potential necrotic tissue. All organoids were dissociated in live culture using a Papain dissociation kit (Worthington), using a modified protocol from Dr. Paola Arlotta's laboratory.[151,156] Cells with viability > 80% were resuspended in 0.01%

BSA (PBS) at a concentration of 1 200 cells/µL and shipped on ice for gel bead in emulsion (GEM) generation and downstream processing (SickKids). 5 000 – 10 000 cells were FACS sorted and run through the 10x Genomics 3' V3.1 Gene Expression pipeline. 50 000 reads/cell were sequenced and provided in FASTQ format.

2.3.5 Single cell RNA sequencing analysis

Raw FASTQ files were first aligned to the human genome hChg38 using Cell Ranger's default alignment parameters. The Seurat R Package V.3.0 was then used for downstream processing.[157] First, low quality cells were identified through identification of high mitochondrial content (>15%) and abnormal read counts (<200 for poor read depth and >1 200 for dublets or multiplets).[157] The matrix was then log-normalized, and regression was performed on the poor-quality cells and ancestral covariant as these are known contributors to variation[158]. Individual samples were aggregated into a single Seurat object, and the aggregate gene expression was scaled and normalized for each gene using a linear model that accounts for and normalizes by the sequence depth of each cell.[157] We used the FindVariableFeatures function to identify genes ("features") that are highly variable in the dataset, and these variable features were used to perform principal component analysis (PCA). The top 30 PCs were selected for clustering using Seurat's FindNeighbors function, followed by FindClusters with a resolution of 0.15.

Cell variation was visualized by a Uniform Manifold Approximation and Projection (UMAP) plot, and later annotated with cluster identities. Using a longitudinal UNO dataset as a training dataset,[159] we applied the cell classification function to identify the top three most likely cell types amongst the predefined clusters. This list was supplemented with manual annotation based on an assessment of canonical cell type-

specific marker expression (Dot plots from S**upplementary Figures 2-3**). Once the populations were identified, cell proportions were assessed by two separate statistical tests: first using the permutation test with the scProportions function, and then supplemented by a Fisher's exact test (two sided) comparing cell frequencies between genotypes.

DGE analysis was performed using a non-parametric Wilcoxon rank sum test with an adjusted p value based on the Bonferroni correction using all features in the dataset. Gene lists were filtered for pAdj < 0.05 and average log2fold change > log(1.2) for pathway analysis using g:Profiler (V0.2.2) including enrichment terms for biological pathway, cellular compartment, reactome, KEGG, and HP, followed by REVIGO semantic similarity reduction (V1.8.1)

The CellChat (V1.6.0) R package was used to gain insight on potential intercellular communication networks from the scRNA Seq dataset, which estimates cellular crosstalk based on proportions of known receptor-ligand complexes through a combination of network analysis, pattern recognition, and manifold learning.[160] We first assessed aggregate interaction probabilities to understand how cellular communication may be behaving more broadly in our system. We measured the total number of predicted interactions, which represents interactions whose probability passes a permutation test that incorporates the law of mass action, a principle in chemistry that considers the ratios signaling ligands and receptors weighted by their agonists and antagonists. Pathway overrepresentation was determined using the rank similarity test, which assesses the similarity between rankings of multiple variables (cell types, interaction strengths, and signaling pathways) and applies a relative (non-numerical) ranking for visualization.

We used the Monocle3 package for pseudotime analysis, which is an algorithm that incorporates differential gene expression dynamics with dimensionality reduction to link cells along a pseudotemporally ordered path through Louvarian clustering.[147,161] To perform pseudotemporal ordering of the cells, a biologically-relevant "root node" or starting point must be chosen, which we assigned as the radial glia population. Monocle3 then measures the distance between the starting points of each cell to the root and will calculate a pseudotemporal value that represents the distance between a cell population of interest and the starting point.

3 **CHAPTER 3:** Loss of 15q13.3 genes in unguided neural organoids (UNOs) disrupts developmental trajectories and predicted cellular communication.

3D neural organoids have provided new avenues for studying neurodevelopment, enabling users to generate and examine tissues previously inaccessible to biological research. In this way, human-specific features of brain development and disease have been identified, although disease modeling is still in its infancy. Neural organoids provide enormous promise in studying complex disorders such as the 15q13.3 microdeletion syndrome, given their ability to model fetal neurodevelopmental milestones that are not captured in the clinic. Specifically, our objective was to use neural organoids to identify vulnerable windows of development and to pinpoint cell types that may confer vulnerabilities to these abnormalities.

3.1 **scRNA Seq of immature 15q13.3del UNOs reveals altered proportion and structure of radial glia populations.**

In order to create 3D neural organoids, iPSCs were first generated using Sendai viral delivery of the Yamanaka factors to reprogram extracted peripheral tissues ($CD34^+$ white blood cells) from our patient cohort into iPSCs. Pluripotency was validated using flow cytometry and immunocytochemistry (ICC) of known pluripotent markers OCT4, NANOG, FITC, SE, and Tra-160 (**Supplementary Fig. 1**). Following karyotyping quality control passing, the cells were then used to generate UNOs from four unique 15q13.3del probands and their typically developing familial controls.

The UNO model was chosen for its ability to stochastically generate multiple cell types from various brain regions,[130,134,162] providing a broad and unbiased overview of the

developing fetal brain. Importantly, UNOs can also form cortical-like structures reminiscent of a developing ventricular zone where neurogenesis occurs within the basal most rosette structures. As we began to grow the 5q13.3del UNOs, we noticed within the first month of growth an increase in 15q13.3del organoid size relative to the control lines, (**Fig. 5C**) and that this size change persisted for upwards of three months.

To assess the contributions of cell proliferation to the UNO size increase, we performed a 24-hour 5-bromo-2'-deoxyuridine (BrdU) pulse to label dividing cells in the UNOs. BrdU is incorporated into the DNA of dividing cells during the S phase of the cell cycle and can be quantified with Ki67 (present at all active phases of the cell cycle) to identify actively dividing cells. We performed ICC on three 15q13.3del families and probed for these two markers within the DAPI-rich rosette structures and found no change in the proportion of doble positive ($Ki67^+$ $BrdU^+$) cells nor the proportion of cells re-entering the cell cycle ($Ki67^-$ $BrdU^+$) within the aggregate dataset (**Fig. 6A**), suggesting that at this timepoint there is no change in the proliferative capacity of cells within the rosettes.

To profile the causative cell types for this growth and probe for signaling changes on a cellular level, we performed single cell RNA sequencing (scRNA Seq) on three unique 15q13.3del families on day 40 (D40) of development (**Fig. 7**). This timepoint represents a period when neural cell types emerge from neural progenitor cells (NPC) and begin to mature,[130] and is thought to resemble early mid-fetal (13–16 gestational weeks) through late mid-fetal (19–20 gestational weeks) neural development.[134] To profile all the neural and progenitor cell types at this timepoint, we used the 10x Genomics 3’ Gene

Expression V3.1 platform to capture and sequence 50 000 reads/cell from ~5 000 cells/line.[149]

After filtering out low quality cells, we were left a total of 29 735 cells (<1% dropout) for downstream analysis. Unsupervised clustering identified 13 distinct cell populations, which were annotated using a training dataset of UNOs across multiple developmental timepoints[159] and validated through assessment of canonical cell type markers (**Supplementary Fig. 2**, cell counts per proportion can be found in **Supplementary Table 1**).

We identified a diverse population of cell types including radial glia populations, neural progenitors, as well as newborn and mature post-mitotic neurons (**Fig. 8A**). To get an overview of the populations, we performed cell composition comparison using both a Fisher's exact test and a permutation test to examine cell proportions.[157,163] From these tests, six populations were identified to have significant differences (FDR <0.05 with absolute fold change>0.58) in proportions between genotypes (two populations decreased in 15q13.3del UNOs, four populations increased). Of the increased populations, half were composed of radial glia populations, which comprise the ventricular-like rosette structures responsible for making newborn neurons (**Fig. 8C-D)**.

To examine gene expression changes more broadly, we performed DGE analysis using the FindMarkers function based on the non-parametric Wilcoxon rank sum test. After filtering for an adjusted p value (pAdj) <0.05, over 8 000 DEGs were identified per cell type cluster. We first probed for 15q13.3 gene expression to ensure heterozygosity of the model, and indeed *OTUD7A*, *CHRNA7*, *FAN1*, *KLF13*, and *MTMR10* were all

significantly downregulated in multiple cell populations (aggregate expression shown in **Supplementary Fig. 2**).

To assess the cell type-specific biological consequences to this heterozygosity, gene ontology (GO) enrichment analysis was then performed on g:Profiler (V2), which showed over 3 000 altered pathways amongst the up and downregulated DEGs. We focused on the radial glia populations, as multiple subtypes were differentially populated in the 15q13.3 UNOs. We used ShinyGO to examine the top 250 up- and downregulated gene pathways and found strong enrichment for neurogenesis and neuron differentiation amongst the downregulated gene set in the apical radial glia population (**Fig. 8E,** biological pathway, FDR<0.001, 2.21-fold enrichment). We visualized the enriched genes from this pathway on STRING (Search Tool for the Retrieval of Interacting Genes), a database of known protein-protein interactions and biological pathways to see if any of the genes interact with each other outside of their role in neurogenesis. After plotting the 37 genes, we confirmed that they were indeed associated with neurogenesis (FDR= 5.68e-16) and noticed the enrichment of multiple cell adhesion proteins (FDR= 0.0019), including laminin (*LAMB1*), *Contactin-5* (*CNTN5*), and *Dystrophin* (*DMD*) (**Fig. 8F)**. In this way, changes in neurogenesis may be influenced by cell adhesion proteins.

SOX2 is a known marker of neural progenitors that is particularly enriched in radial glia and may be altered in states of reduced neurogenesis and differentiation.[164] In support of this, the UMAP plot of relative SOX2 mRNA expression showed the highest expression in the four radial glia subpopulations (**Fig. 9A**). We measured the global mRNA expression of SOX2 and found a significant increase in mRNA transcript

expression (**Fig. 10B,** $p < 0.001$), prompting us to examine the gene at the protein level. We performed ICC on a total of four 15q13.3del UNO families to target the $SOX2^+$ radial glia that comprise the rosette structures within the organoids. We found consistent increases in rosette size (nearly double that of the controls) (**Fig. 9C**), and dysmorphic rosette shapes, whereby rosette shape was not circular but rather caved in or turned inward.

In addition to forming the ventricular-like rosettes, UNO radial glia will proliferate and eventually give rise to newborn neurons that migrate to the outer edge of the organoid to produce laminar structures reminiscent of the developing cortex. Given the neurogenesis enrichment amongst the downregulated aRG genes, we probed for the expression of *TBR1*, a marker for cortical newborn neurons,[165] and found that this population was significantly decreased amongst the four 15q13.3del UNO families (**Fig. 9C**), indicating potential disruptions in neuronal maturation or differentiation, as indicated by our previous GO enrichment analysis.

This is the first time radial glia populations have been assessed in the 15q13.3del background, where we found structural abnormalities including differences in cell density, rosette size, and the proportion of newborn neurons amongst four 15q13.3del families. Early changes in precursor cell development suggests that there may be further impairments in early neuron development that have not been sufficiently characterized in the postnatal mouse models, and an assessment of the neurogenesis-enriched gene set suggests the role of cell adhesion in that process.

3.2 Mature 15q13.3del UNOs have global changes in inferred intercellular communication and enrichment for cell adhesion pathways.

Since radial glia can give rise to multiple neural subtypes (excitatory, inhibitory, oligodendrocytes, etc.), we aged the UNOs to a timepoint that recapitulates late mid-fetal gestation weeks 19-24 to capture fate specification and cellular trajectory of both newborn neurons and mature excitatory and inhibitory neurons.[134,166] During this time, we observed persistent size changes amongst 15q13.3del UNOs, as well as a decrease in cell density and increase in nuclei diameter consistent with the D40 timepoint (**Fig. 11**).

We sought to characterize the cellular dynamics on the single cell level and used the same three patient families to perform scRNA Seq using identical sample preparations to the previous experiment. To account for the increase in cell diversity at this timepoint, we doubled the number of cells sequenced from 5 000 to 10 000/sample and were able to capture a total of 68 455 cells post-filtering. Using the same analysis pipeline and training dataset,[159] we identified 11 unique cell clusters, including: neural inhibitory and excitatory progenitors; mature excitatory neurons, mature inhibitory interneurons, astrocytes, and oligodendrocyte precursors (OPCs) (**Fig. 11B,** cell counts per proportion can be found in **Supplementary Table 2**). A small population (345 cells, ~ 0.5%) matched the transcriptome of $COL1A1^{+}/GFAP^{-}$ mesenchymal-like cells, which are likely neural crest cells and were excluded from downstream DEG analysis.

We first sought to confirm 15q13.3 gene heterozygosity again at this later timepoint, and from aggregate mRNA transcript level (**Supplementary Fig. 2**) it was apparent that

the transcript levels remained reduced. We used the non-parametric Wilcoxon rank sum test with an adjusted p value to perform DGE analysis (filtering for p adj< 0.05), which produced lists for each cell type that ranged from ~ 600 to over 6 000. Each 15q13.3 gene with high expression in multiple cell types (**Supplementary Fig. 3**) was significantly downregulated in over half of the cell types, including *KLF13*, *FAN1*, *MTMR10*, and *OTUD7A*. Amongst the DEGs included ASD risk genes, axonal regulators, and OTUD7A interaction partners *ANK2* and *ANK3*, which were significantly downregulated in 8/10 and 9/10 cell types, respectively. Since *ANK3* was expressed in multiple cell types, we confirmed its aggregate reduction on the protein level through western blot of 2-month 15q13.3del UNOs (**Fig. 12**) as a validation to the scRNA Seq pipeline.

Given the diversity of cell types in this current model, we questioned how cell interactions and intercellular communication may be altered in aggregate and amongst distinct populations. We used the R package CellChat to gain insight on potential intercellular communication networks from the scRNA Seq dataset, which quantitatively infers cellular crosstalk based on proportions of known receptor-ligand complexes.[160] We first assessed aggregate interaction probabilities to understand how cellular communication may be behaving more broadly in our system. We measured the total number of predicted interactions and observed a reduction in the number and strength of significant cell interactions amongst the 15q13.3del dataset (**Fig. 13B**, 4676 in the control *vs.* 3224 in the 15q13.3del for the number of total significant interactions, and 0.164 interaction strength for the control vs. 0.068 for the 15q13.3del), suggesting broad reductions in cell communication. We stratified the total interactions and strength by cell type (**Fig 14C**) and found multiple populations to contribute towards the inferred

reduction in communication, with the exception of the outgoing signal from select excitatory neuron populations (non-telencephalic midbrain excitatory neurons signaling towards non-telencephalic hindbrain excitatory neurons and mature excitatory neurons; mature excitatory neurons towards apical radial glia and non-telencephalic midbrain excitatory neurons; and non-telencephalic hindbrain excitatory neurons towards apical radial glia, indicated by an increase in relative interaction between these cell types. We questioned what signaling pathways may underlie these changes and determined the overrepresented pathways by using the rank similarity test, which assesses the similarity between rankings of multiple variables (cell types, interaction strengths, and signaling pathways) and applies a relative (non-numerical) ranking for visualization. The top overrepresented pathway amongst the aggregate tissue in the rank similarity plot was Nectin signaling, a family of cell-adhesion molecules that are regulated in a calcium-dependent manner (**Fig. 14A**). To determine how cell adhesion signaling may be disrupted amongst single cell populations, we visually assessed the outgoing signaling of Nectin, the top cell adhesion pathway, in addition to Ephrin-A (EphA) and Ephrin-B (EphB) signaling. The resulting circle plots showed outcoming signaling from multiple cell populations, where the circle size and edge width are proportional to the number of cells in each cell cluster and the communication score between interacting cell clusters, respectively (**Fig. 14B**). Amongst the control cell populations, cell adhesion signaling was present between each cell type, with the radial glia populations producing the strongest Nectin and EphB signaling. In contrast, the 15q13.3del population showed a general reduction in communication scores (circle edges), with some cell types showing an absence of detectible signal, such as the inhibitory neurons input and output of Nectin

signaling, and the absence of detectible EphA signaling form all but the Cajal-Retzius and non-telencephalic excitatory hindbrain neurons.

We were interested if biological pathways from the DEGs of each cell type would support this predictive analysis, and performed GO analysis followed by REVIGO semantic similarity reduction to produce tree plots of the core biological features of each of the cell types sets,[167] excluding inhibitory neurons in the downregulated plots and Cajal-Retzius cells in the upregulated plots due to a lack of unique GO terms to aggregate. The terms varied between cell types, but a repeated feature amongst the downregulated terms included cellular adhesion pathways (5/9 plots) and nervous/systems/developmental processes (identified in 8/9 REVIGO plots, **Fig. 14C**), confirming that cell adhesion was disrupted at the transcript level.

Cellular communication is critical to the formation and development of neurons; without the appropriate input from nearby cell types, a neuron's excitability is compromised and thus, also its ability to integrate into a neural network.[168] Cell adhesion molecules play a critical role in the establishment of neuronal connections at the pre and post synapse, but also contribute more broadly to neuron maturation by participating in processes such as axon guidance, dendritic spine morphology, as well as synaptic plasticity, maintenance, and homeostasis.

3.3 Pseudotime analysis predicts alterations in cell type emergence and identifies putative imbalances to excitatory and inhibitory signaling

We exploited pseudotime analysis to better understand the emergence and trajectory of each cell type within our dataset. We used the Monocle3 plugin, an algorithm the

incorporates differential gene expression dynamics with dimensionality reduction to link cells along a pseudotemporally ordered path through Louvarian clustering.[147,161] To perform pseudotemporal ordering of the cells, a biologically-relevant "root node" or starting point must be chosen, which we assigned as the radial glia population. Monocle3 then measures the distance between the starting points of each cell to the root and calculates a pseudotemporal value that represents the distance between a cell population of interest and the starting point. After choosing the radial glia root node, 24 104 DEGs were identified to fit along the pseudotemporal trajectory with time-dependent expression (FDR <0.05). The unsupervised clustering of these genes resulted in 44 total pseudotime modules, which were used to produce the pseudotime UMAP partition plot in **Fig. 15**. The pseudotime partition plot mirrored the maturation trajectory that is expected of the developing brain, beginning with radial glia populations, and branching off into various post mitotic neural subtypes, each with a higher pseudotime value across the differentiation trajectory.

Given the consistencies between the differentiation state and cell type, we proceeded to quantify the pseudotime distributions to get a broad idea of cell type emergence. We found that the aggregate distribution between genotypes was significantly different by the Kolmogorov–Smirnov test ($p < 2.2e10\text{-}16$), suggesting changes in the developmental trajectory. In case a specific cell type may be driving this change in trajectory, we stratified the pseudotime data by individual cell types (**Supplementary Fig. 7**). We found that all but the mesenchymal-like cell populations had significantly different cell trajectories when analyzed independently, with the most significantly altered cell type

being the non-telencephalic hindbrain excitatory neurons ($p < 2.2e10\text{-}16$, pseudotime distance (D) = 0.21562). **Table 3** provides a ranked list of affected cell types.

Table 3: Ranked list of cell type-specific pseudotemporal changes in 15q13.3 D120 UNOs.

Statistics derived from Asymptotic two-sample Kolmogorov-Smirnov test, two sided.

Cell Type	Pseudotime distance value	P value
Non-telencephalic hindbrain excitatory neurons	0.21562	< 2.2e-16
Radial glia	0.18382	< 2.2e-16
Apical radial glia	0.13583	>2.2e-16
Non-telencephalic midbrain excitatory neurons	0.1194	>2.2e-16
Cajal Retzius cells	0.12903	4.463e-14
Intermediate progenitors and excitatory neurons	0.099327	2.234e-11
Excitatory mature neurons	0.097355	4.28e-09
Oligodendrocyte precursors	0.25044	1.872e-07
Cycling radial glia	0.096134	1.378e-05
Inhibitory neurons	0.079755	4.03e-05

We decided to further examine excitatory and inhibitory neuron emergence, as both cell types are required for the formation of cortical circuits in the developing brain. We examined the mature excitatory neurons over the non-telencephalic cell types, as they

more closely resemble the expected cell types of the cortex. After confirming significant changes to the pseudotemporal distribution of this cell type (p =4.28e-09, pseudotime distance (D) = 0.097355), we assessed the temporal modules for gene clusters that were enriched in mature excitatory neurons. We identified three distinct clusters (cluster 14, 18, and 19, **Fig. 17B**) and proceeded to examine cluster 19 due to its specific expression in the excitatory mature neuron cluster (**Supplementary Fig. 8A**). We classified the modules based on GO term enrichment, where synaptic signalling was assigned as the module identifier. We then tested to see if the genes within the module were significantly different between genotypes by performing a Wilcoxon test and found a significant increase in the aggregate gene expression in the 15q13.3del excitatory mature neurons (**Fig. 16,** p = 9.161339e-63). We also probed the gene set and confirmed a significant increase in ASD-risk genes *SYNGAP1*, *NRXN1*, and *CACNA1G*. We noticed that *NRXN* was also one of the over-represented signaling pathways reported in our previous CellChat analysis and assessed for the putative signaling network using the circle plot function. We found that, consistent with the pseudotime inference, the CellChat analysis also predicted an increase in general NRXN signaling between multiple cell types, with apical radial glia, inhibitory neurons, and excitatory mature neurons showing the most apparent increases (**Fig. 18)**.

We then repeated the same analysis for the inhibitory neuron population, in which there was also significant differences in pseudotime distribution (p = 4.03e-05, pseudotime distance (D) = 0.079755). We observed a highly specific gene module (module 37, **Supplementary Fig. 8B**) from the heatmap and found it to also be associated with synaptic signalling, however, it was more specified to GABAergic

signalling and transmission. Using the inhibitory neurons, we assessed the differential gene expression of module 37 genes and found it had the opposite trend than the excitatory mature neurons, in that there was a decrease in GABAergic signaling in the 15q13.3del inhibitory neurons (**Fig. 17**). We then confirmed the reduction of multiple GABAergic modulators in the 15q13.3del inhibitory neurons, including *GAD1*, *SLC6A1*, and *ERBB4* (a receptor highly enriched in PV interneurons).

We returned to CellChat to visualize significant ($p < 0.05$, one-sided permutation test) ligand-receptor pair expression between the inhibitory neurons and remaining cell types to see if the pseudotime predictions (reduced GABAergic signaling) were reflected by inferences made by CellChat. We could not detect any significant changes involving GABA; however, we were surprised to find several significant reductions in NRXN-NLGN ligand-receptor signaling (**Fig. 18**). This included NRXN 1-3 expression in the inhibitory neurons (acting presynaptically), and NLGN 1-3 in the remaining cell types (as the postsynaptic ligand), with the biggest difference between homotypic inhibitory-inhibitory interactions. We then visualized the normalized expression of the NRXN receptor and ligand in each cell type, confirming a decrease in the inhibitory neurons, with NRXN3 as the most prominent reduction. Importantly, and in agreement with the pseudotime analysis, we also saw an increase of the NRXN receptors amongst the excitatory mature neuron population, further bolstering the idea of excessive excitatory neuron signaling.

The possibility of 15q13.3del E/I imbalances at the synaptic level supports the epileptic phenotype observed in patients and mouse models but has not yet been characterized in human E/I cocultures or organoid models. Instead, the majority of

15q13.3 literature has focused on excitatory cortical neurons, which have generally shown immature functional phenotypes in mice. In contrast, human 15q13.3.del NGN2 glutamatergic-like neurons show early hyperexcitable phenotypes followed by a reduction in activity, suggesting that the neural activity phenotypes may be more complex in a human background, calling for more precise human modeling.

3.4 Chapter summary and considerations

This study represents the first exploration of early neural development and cell type emergence in a human 15q13.3del background. We exploited an unbiased neural organoid model for its ability to spontaneously produce progenitors and mature neural cell types from multiple brain regions, where we found changes in early radial glia population and the ventricular-like rosette structures that they occupy. After maturing the model further and enabling these radial glial cells to mature into more diverse neural subtypes, we identified altered pseudotemporal trajectories of the principal neurons that populate the cortex: excitatory neurons and inhibitory interneurons, and later predicted disrupted and inverse changes to their synaptic transcriptomic signatures. We then sought to examine communication more broadly, and by using a predictive cellular communication platform we identified global change is intercellular communication, and in particular of nectin and ephrin signaling.

The model system and analyses featured in this chapter are not without their caveats, however. The use of inference-based platforms such as a CellChat and pseudotime analysis require rigorous validations, both functionally and on the protein level to hold biological meaning. Due to the loss of spatial information in scRNA Seq, predicted

changes between two cell types (such as the proposed changes to E/I dynamics) require functional validations of their physical circuits, including recordings of baseline activity and following stimulation of the presynaptic cell.

Given the stochastic nature of UNO cell differentiation, heterogeneity is inherent in the system and only increases as the organoids mature and differentiate further. The diversity in cell types makes circuit phenotyping difficult, as the proportion of altered cell types such as excitatory mature neurons and inhibitory neurons can be diluted or out-competed by non-telencephalic neural populations. In addition, UNO models do not provide the signaling cues and microenvironment necessary for directed interneuron migration, which is a developmental milestone that largely shapes circuit function in the cortex and is worth assessing in this genetic background.

An alternative to UNO modeling that produces more pure populations of the intended neural cell type are guided neural organoids, which can be generated by adding patterning factors that direct the organoid towards a particular lineage or brain region. Using this system, we will better characterize abnormalities that are specific to excitatory and inhibitory neurons and combine them into assembloids to assess the circuit development and intercommunication between the two cell types.

3.5 Chapter 3 Figures

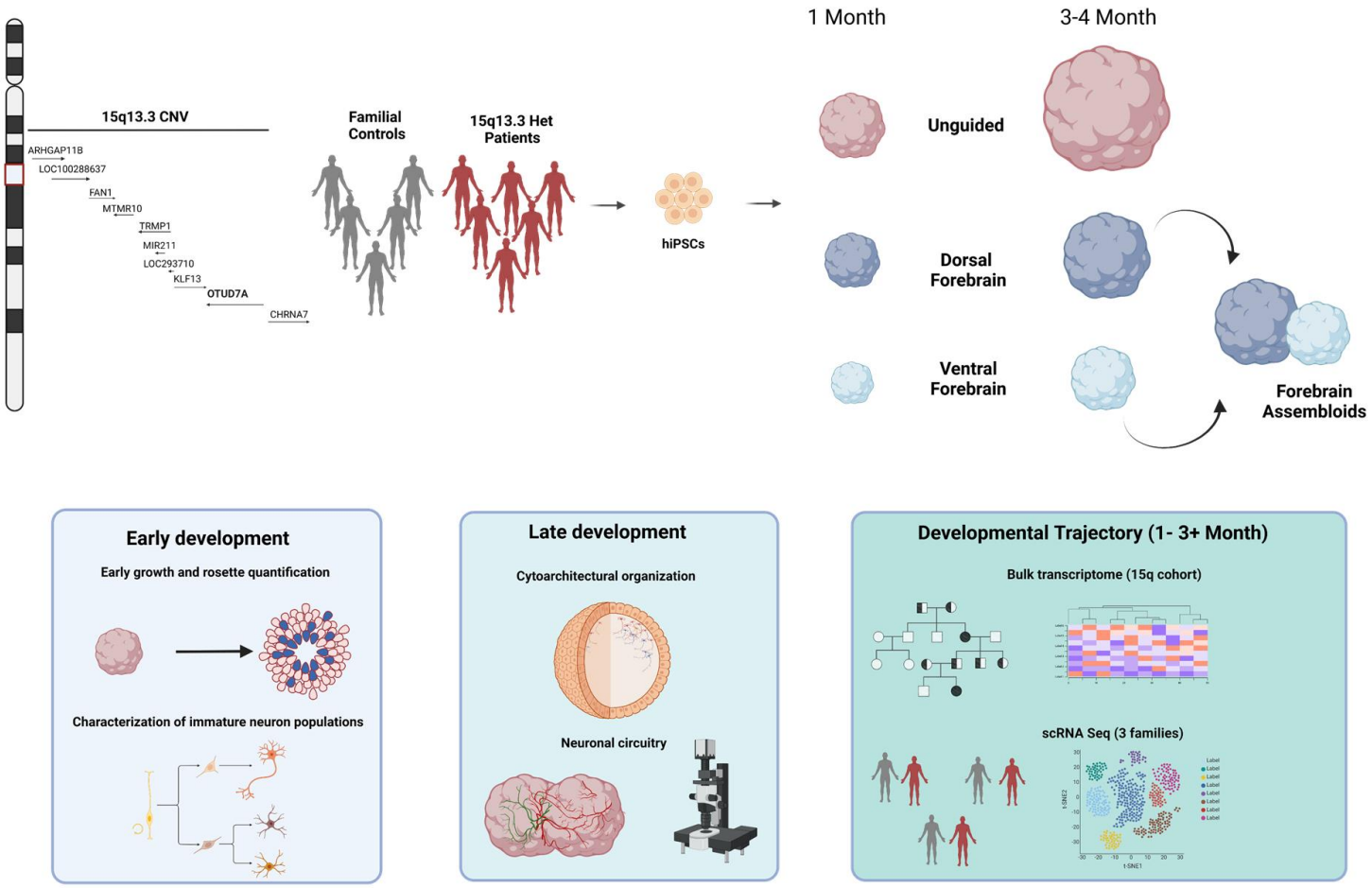


Figure 4: Organoid models and technical approaches.

Patient cohort comprised of 15q13.3 microdeletion patients and familial controls is used to generate human induced pluripotent stem cells (hiPSCs) and later 3D neural organoids for trajectory phenotyping.

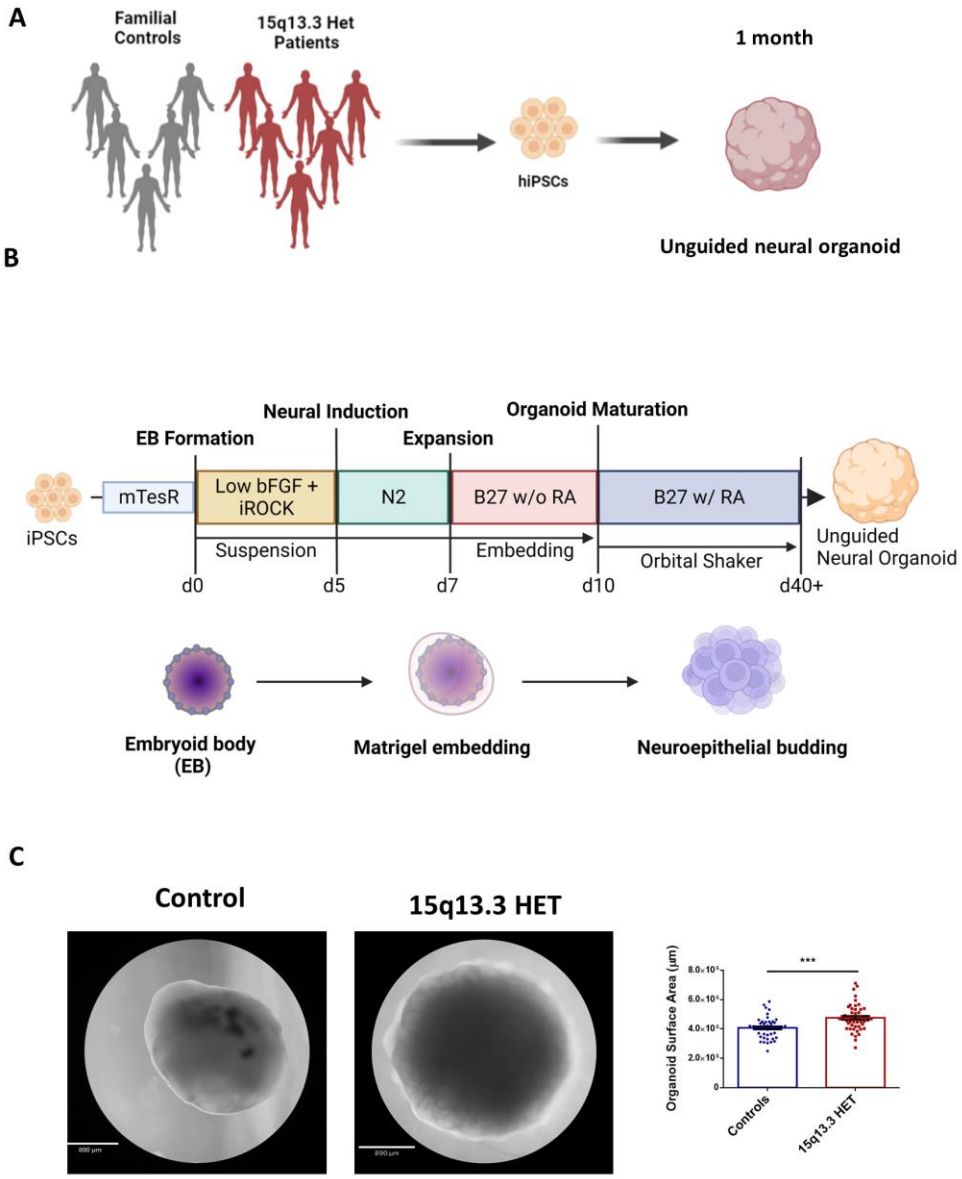


Figure 5: 15q13.3del unguided neural organoids (UNOs) show an increase in organoid size.

(A) Schematic of 15q13.3del cohort used for unguided neural organoid (UNO) phenotyping.
(B) Unguided neural organoid protocol requires three core steps: embryoid body (EB) formation, embedding in Matrigel, and neuroepithelial budding. Bovine fibroblast growth factor 2 (bFGF), (N2), retinoic acid (RA), Y-27632 (Rock inhibitor, iROCK).
(C) Representative 4x brightfield images of day 40 15q13.3del UNOs from families 1-4 and measured surface area (combined data, (WT 1 n = 14-18, HET 1 n = 15-19/patient). *Data analyzed by two-tailed student's t-test, *** $p < 0.001$, values represent mean ± S.E.M.*

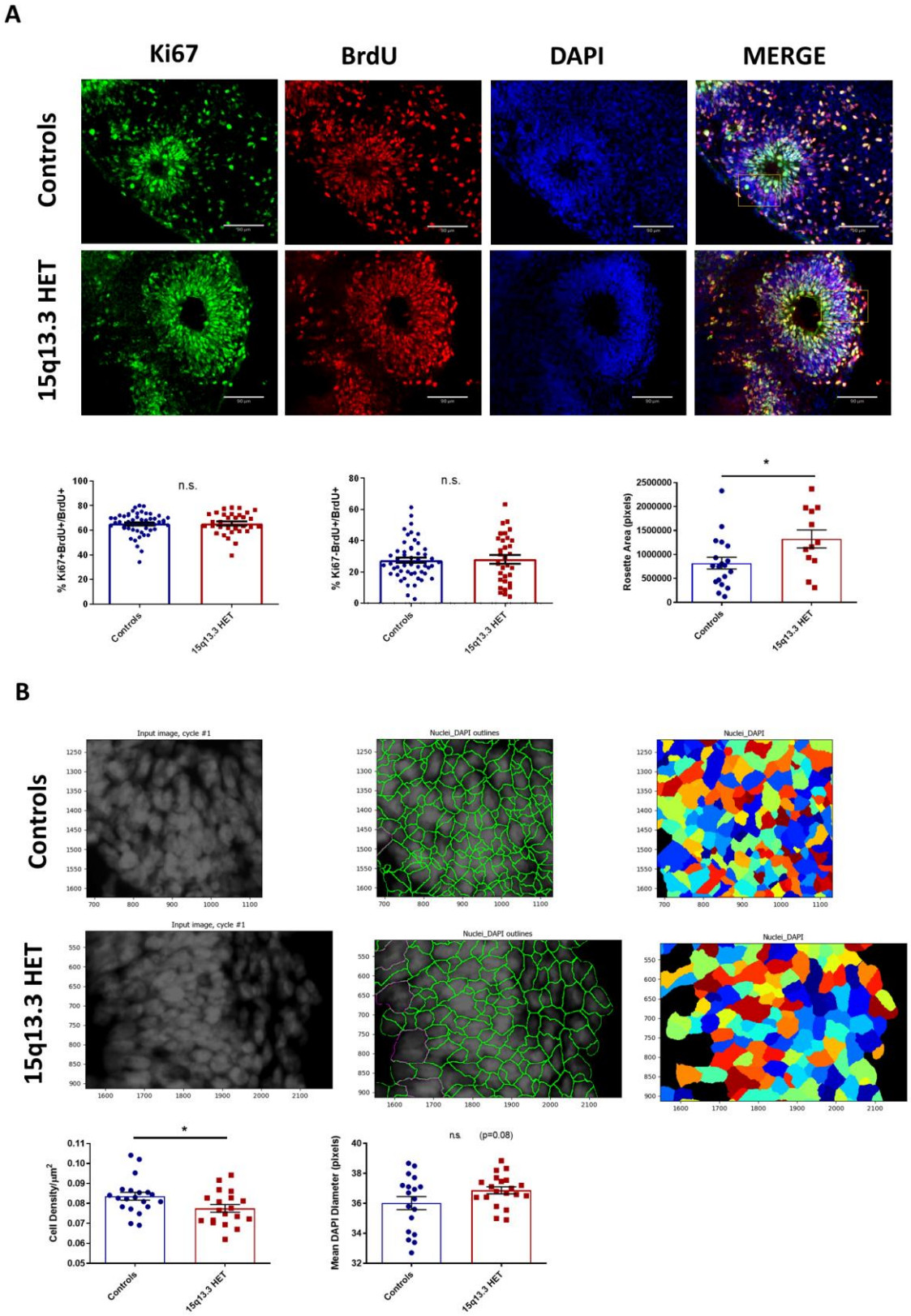

A
Ki67
BrdU
DAPI
MERGE
Controls
15q13.3 HET
% Ki67+BrdU+/BrdU+
n.s.
% Ki67-BrdU+/BrdU+
n.s.
Rosette Area (pixels)
*
Controls
15q13.3 HET
B
Input image, cycle #1
Nuclei_DAPI outlines
Nuclei_DAPI
Cell Density/μm²
*
Mean DAPI Diameter (pixels)
n.s. (p=0.08)

Figure 6: 15q13.3 unguided neural organoids display size differences and reductions in cell density that are not due to proliferation changes at one month of growth.

(A) (Top) Representative 20x images from Day 40 BrdU-pulsed15q13.3 cerebral organoids (Bottom) quantification of double positive $Ki67^{+}BrdU^{+}$ cell populations, cell cycle exit ($Ki67^{-}BrdU^{+}$) and $DAPI^{+}$ nuclei density; Control n = 9 ventricles, 15q13.3 proband n= 12 ventricles across 3 organoids from families 1-3.
(B) (Top) Representative $DAPI^{+}$ nuclei traces processed in CellProfiler and (right) quantification of nuclei density. *Data analyzed by two-tailed student's t-test, n.s. =* nonsignificant, * $p < 0.05$, *** *$p < 0.001$, values represent mean ± S.E.M.*

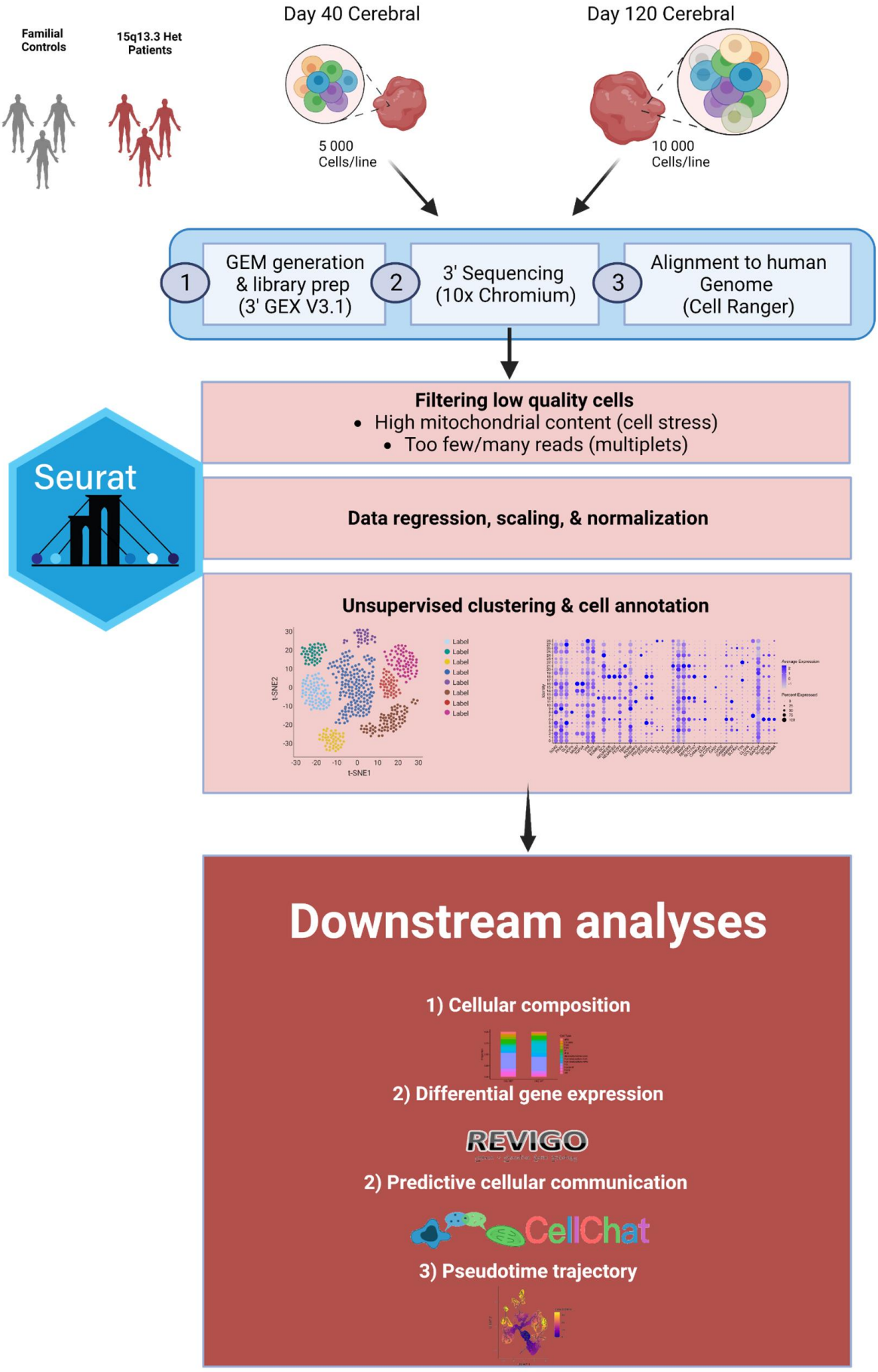

Familial Controls
15q13.3 Het Patients
Day 40 Cerebral
5 000 Cells/line
Day 120 Cerebral
10 000 Cells/line
1
GEM generation & library prep (3' GEX V3.1)
2
3' Sequencing (10x Chromium)
3
Alignment to human Genome (Cell Ranger)
Seurat
Filtering low quality cells
• High mitochondrial content (cell stress)
• Too few/many reads (multiplets)
Data regression, scaling, & normalization
Unsupervised clustering & cell annotation
t-SNE1
t-SNE2
Downstream analyses
1) Cellular composition
2) Differential gene expression
REVIGO
2) Predictive cellular communication
CellChat
3) Pseudotime trajectory

Figure 7: Pipeline schematic for processing and analyzing single cell RNA sequencing for unguided neural organoids.

Single cell sequencing (scRNA Seq) of three 15q13.3del families includes tissue dissociation at two timepoints, followed by library generation, mRNA sequencing, and read alignment to the human genome. Seurat package was used for downstream processing of integrated dataset, which includes data normalization, cell clustering, and downstream analyses such as CellChat and pseudotime analysis. Analysis pipeline designed by PhD student Jarryll Uy.

Supplementary Table 1: Population cell counts in Day 40 UNO scRNA Seq.

Population	**Cell count**
Radial glia	10009
Radial glia in S phase	2160
Non-telencephalic excitatory neurons	4587
Intermediate progenitors	2775
Excitatory mature neurons	1276
Mesenchymal-like cells	1271
Apical radial glia	1232
Radial glia in G2M phase	1875
Intermediate progenitors - S phase	758
Non-telencephalic neural progenitor cells	1735
Excitatory neurons	1735
Cajal-Retzius cells	417
Unknown	639
Total	**30469**

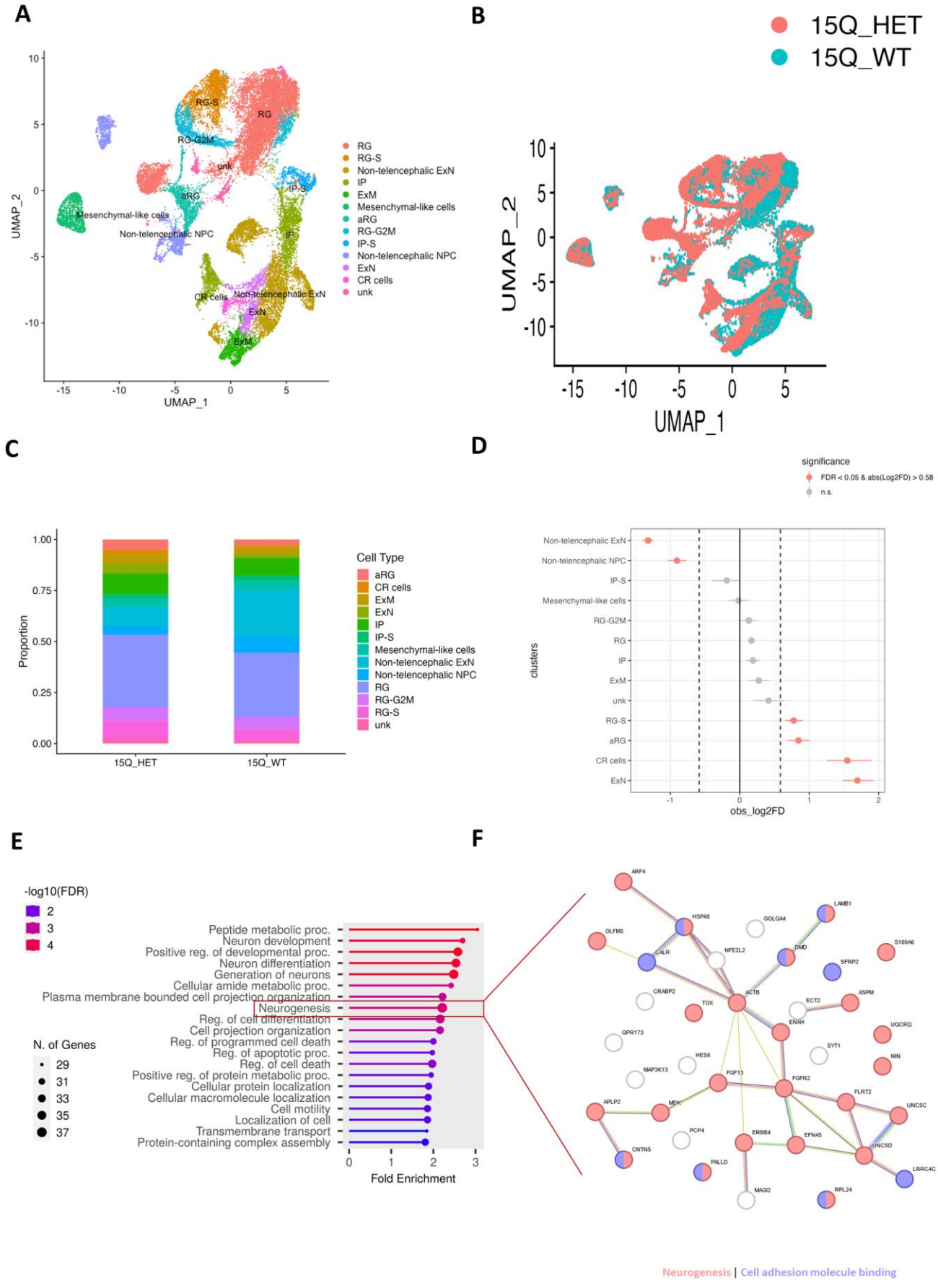
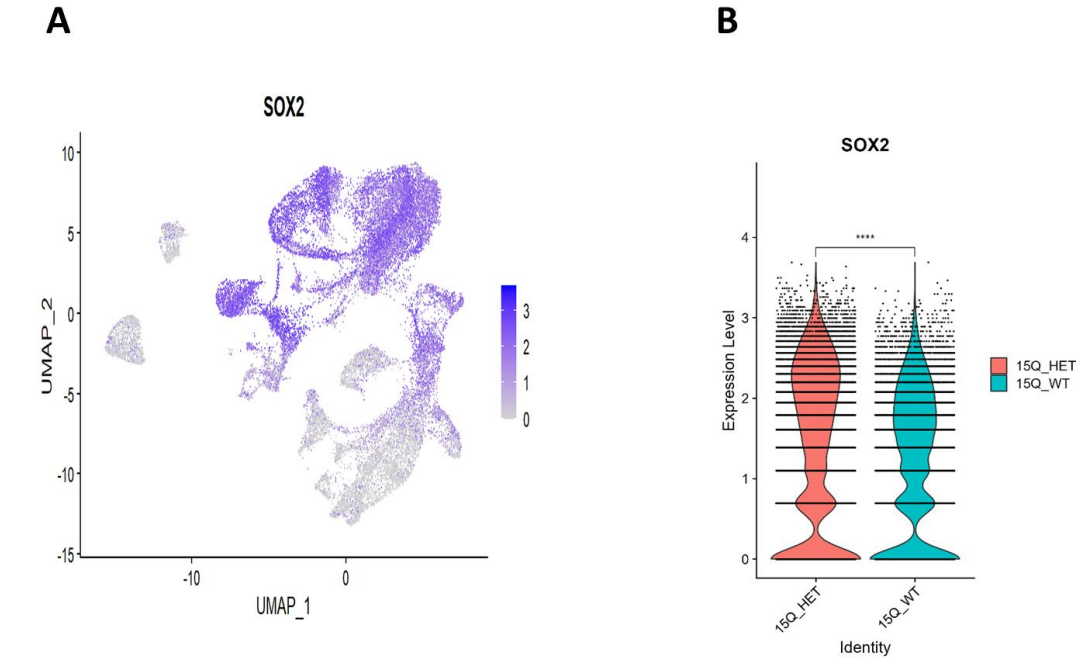
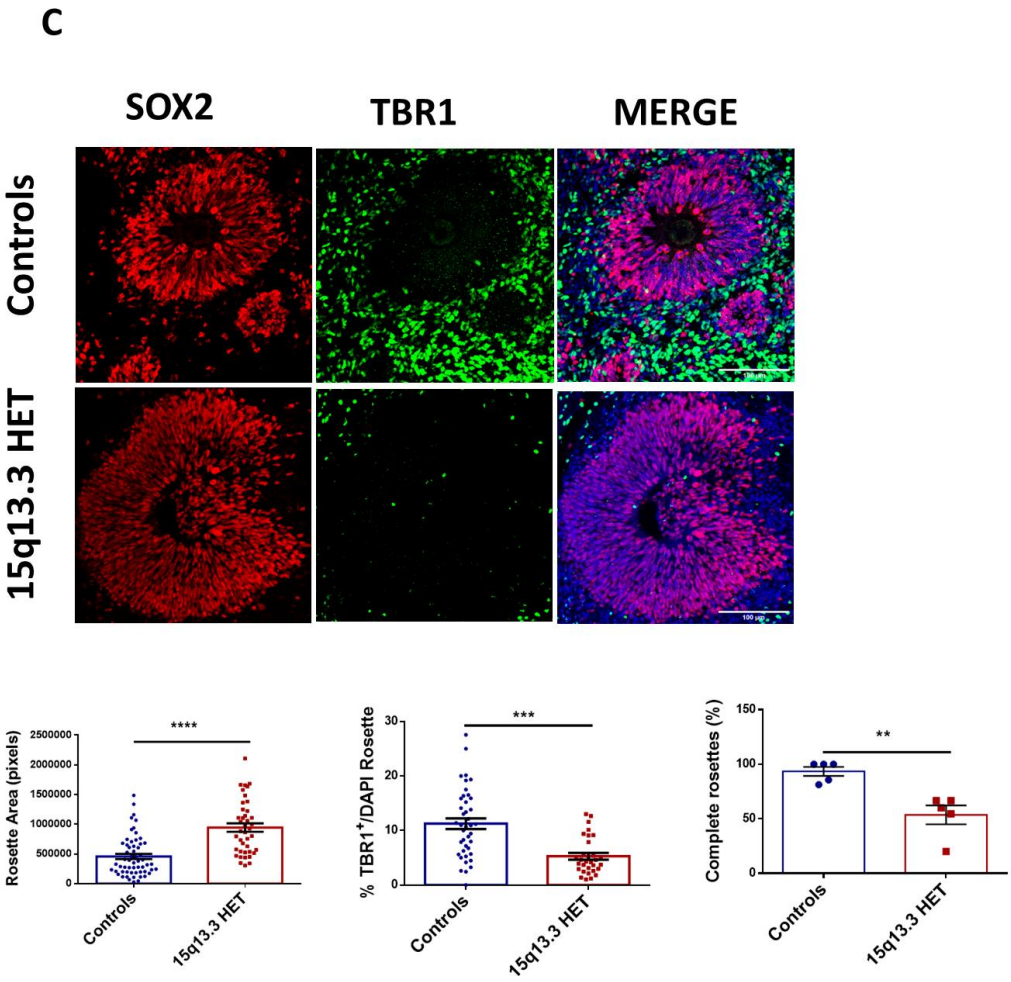

A
B
C
D
E
F
15Q_HET
15Q_WT
UMAP_1
UMAP_2
RG
RG-S
Non-telencephalic ExN
IP
ExM
Mesenchymal-like cells
aRG
RG-G2M
IP-S
Non-telencephalic NPC
ExN
CR cells
unk
Cell Type
Proportion
significance
FDR < 0.05 & abs(Log2FD) > 0.58
n.s.
clusters
obs_log2FD
-log10(FDR)
N. of Genes
Peptide metabolic proc.
Neuron development
Positive reg. of developmental proc.
Neuron differentiation
Generation of neurons
Cellular amide metabolic proc.
Plasma membrane bounded cell projection organization
Neurogenesis
Reg. of cell differentiation
Cell projection organization
Reg. of programmed cell death
Reg. of apoptotic proc.
Reg. of cell death
Positive reg. of protein metabolic proc.
Cellular protein localization
Cellular macromolecule localization
Cell motility
Localization of cell
Transmembrane transport
Protein-containing complex assembly
Fold Enrichment
Neurogenesis | Cell adhesion molecule binding

Figure 8: Immature unguided neural organoids display an increase in radial glia populations with reductions in neurogenesis-associated pathways.

(A) Day 40 UMAP plot (0.15 resolution, Control N = 13 475 cells, 15q13.3del N = 16 260 cells) across 3 families. a/RG/S/G2M (apical/radial glia/interphase/G2-Mitosis), OPC (oligodendrocyte precursor cell), ExM (excitatory mature neuron), IP/S (intermediate progenitor/in interphase), NPC (neural progenitor cell), CR (Cajal-Retzius) cells, unk (unknown).
(B) UMAP visualization of cell population densities filtered by genotype.
(C) (Left) Cell proportion test and (right) point-range plot from permutation test results with bootstrapping for cellular proportions. Dashed line represents FDR < 0.05 and absolute Log2 fold change > 0.58.
(D) Aggregate mRNA expression of SOX2 transcript stratified by genotype.
(E) GO enrichment analysis of apical radial glia downregulated DEGs shows enrichment for neurogenesis and cell projection (FDR <0.05).
(F) STRING visualization of neurogenesis-associated genes (FDR = 5.68e-16) also share enrichment for cell adhesion (FDR = 0.0019)

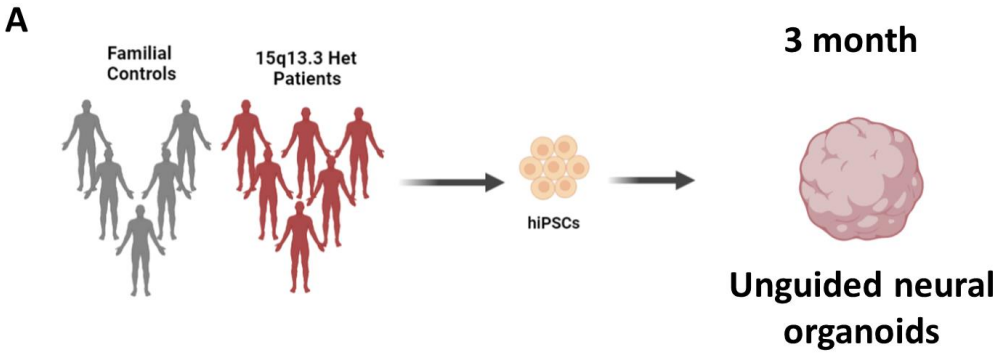

A
SOX2
UMAP_2
UMAP_1
B
SOX2
Expression Level
****
15Q_HET
15Q_WT
Identity

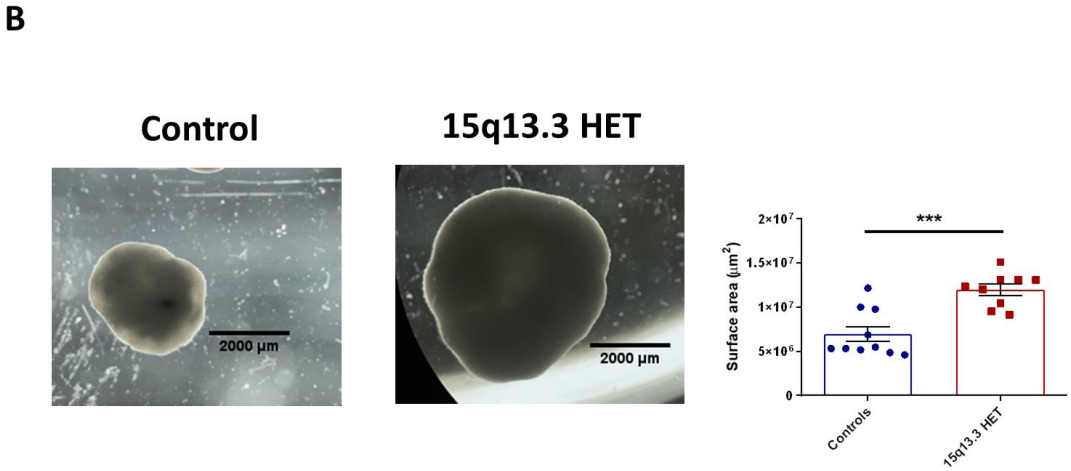

C
SOX2
TBR1
MERGE
Controls
15q13.3 HET

Rosette Area (pixels)
****
Controls
15q13.3 HET
% TBR1+/DAPI Rosette
***
Controls
15q13.3 HET
Complete rosettes (%)
**
Controls
15q13.3 HET

Figure 9: Radial glia dynamics are altered at the structural level in 15q13.3del UNOs.

(A) UMAP visualization plot of normalized SOX2 mRNA expression.
(B) Aggregate mRNA expression of SOX2 transcript stratified by genotype ($p < 0.001$, non-parametric Wilcoxon rank sum test).
(C) (Top) Representative 20x images from Day 40 15q13.3del UNOs Families 1-4. (Bottom) quantification of rosette area, proportion of $TBR1^+$ newborn neurons within organoid rosettes, and proportion of rosettes with circular formation. Control = 11 rosettes, N 15q13.3del = 13 rosettes. Data represent mean ±SEM in organoid slices; **p<0.01, ***p<0.001; student's unpaired *t-test*.


A
Familial Controls
15q13.3 Het Patients
hiPSCs
3 month
Unguided neural organoids


B
Control
15q13.3 HET
2000 µm
2000 µm
***
Surface area (µm²)
Controls
15q13.3 HET

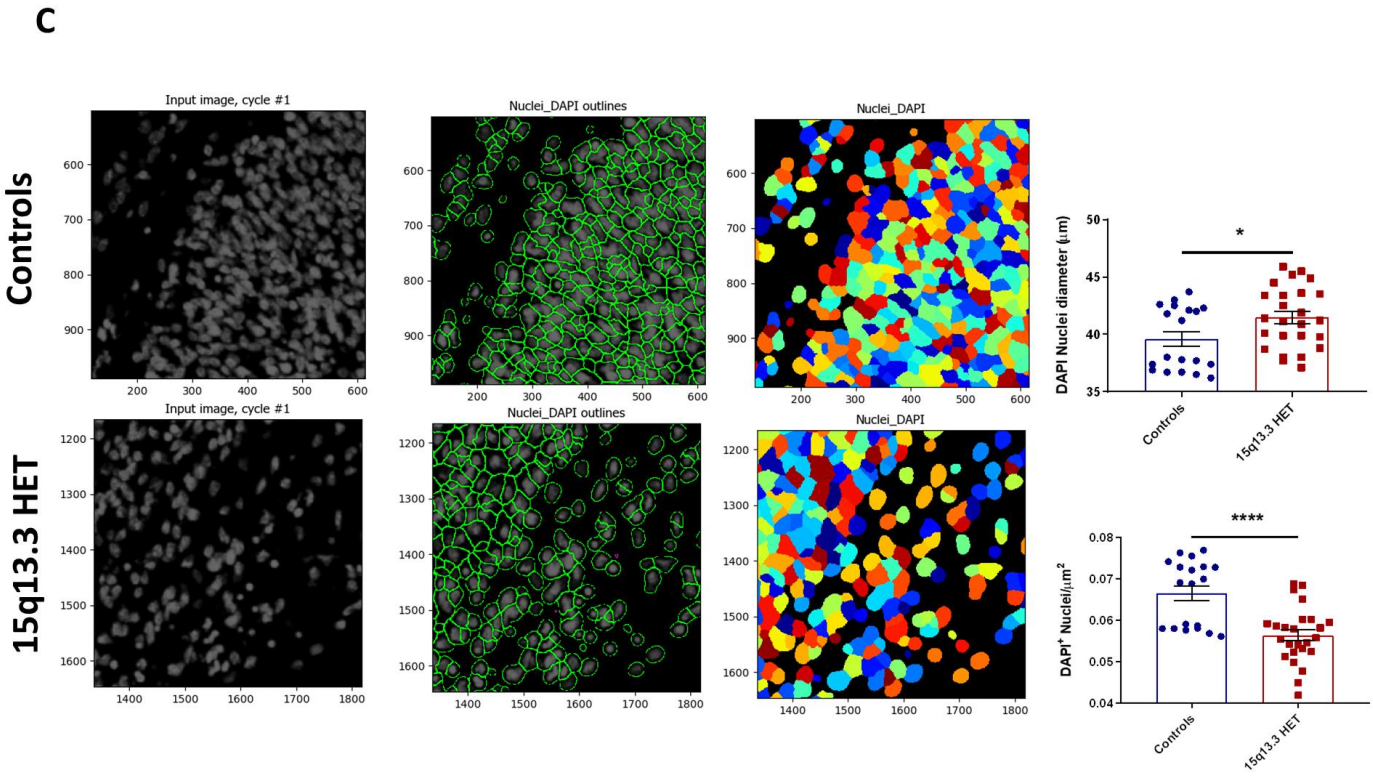
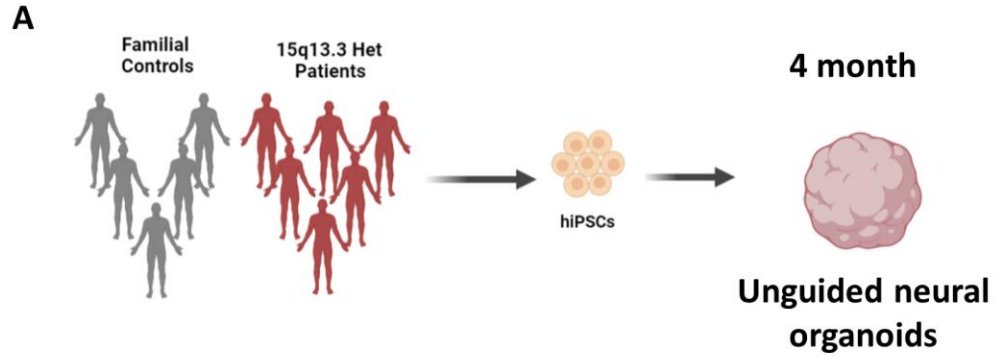
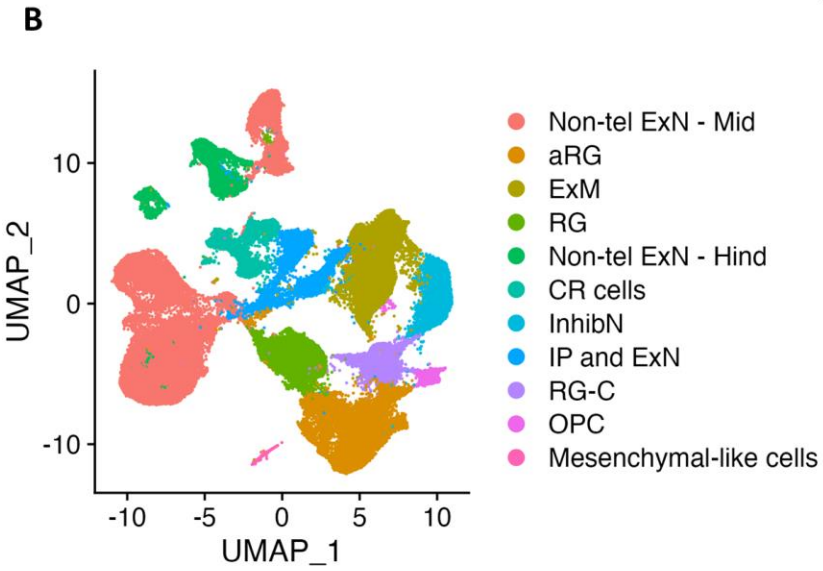

C
Controls
15q13.3 HET
Input image, cycle #1
Nuclei_DAPI outlines
Nuclei_DAPI
*
DAPI Nuclei diameter (µm)
****
DAPI⁺ Nuclei/µm²
Controls
15q13.3 HET

Figure 10: Mature 15q13.3 unguided neural organoids retain increases in size and reductions in DAPI+ nuclei density.

(A) Schematic of single cell experiment at day 120 timepoint.
(B) Representative 4x brightfield images of day 90 15q13.3del UNOs from family 3 and 5, and measured surface area (combined data, (WT 3/5 n = 10, HET 3/5 1 n = 9).and measured surface area.
(C) (Left) Representative $DAPI^{+}$ nuclei traces processed in CellProfiler and (right) quantification of nuclei density from three 15q13.3del families. Data represent mean ±SEM in organoid slices; *p<0.05, ***p<0.01, ***p<0.001; student's unpaired *t-test*.


A
Familial Controls
15q13.3 Het Patients
hiPSCs
4 month
Unguided neural organoids


B
UMAP_2
UMAP_1
Non-tel ExN - Mid
aRG
ExM
RG
Non-tel ExN - Hind
CR cells
InhibN
IP and ExN
RG-C
OPC
Mesenchymal-like cells

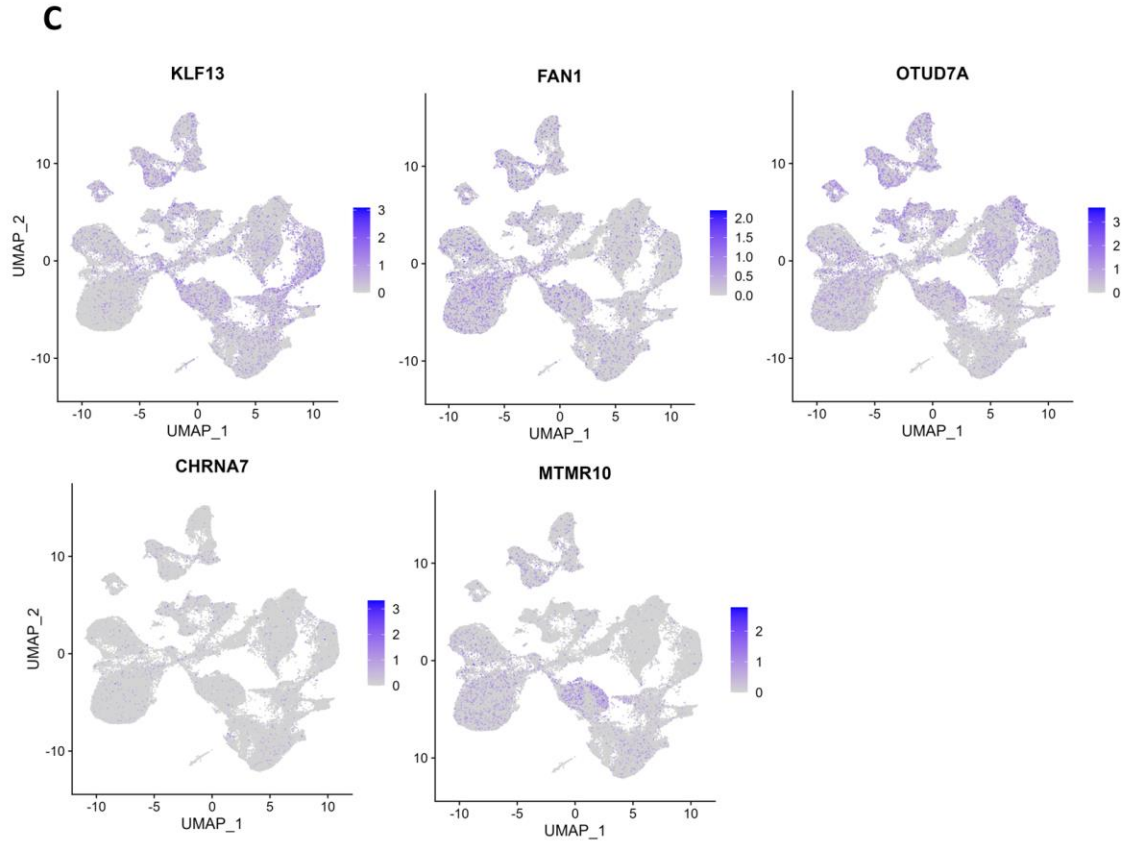

C
KLF13
FAN1
OTUD7A
CHRNA7
MTMR10
UMAP_2
UMAP_1

Figure 11: Mature unguided neural organoids produce diverse neural populations that express the 15q13.3 genes.

(A) Schematic of single cell experiment at day 120 timepoint.
(B) Day 120 UMAP plot (0.15 resolution, Control N =31 784 cells, 15q13.3del N = 3 6671 cells) across 3 families. a/o/RG/c (apical/outer/radial glia/cycling), OPC (oligodendrocyte precursor cell), ExM (excitatory mature neuron), IP and ExN (intermediate progenitor and excitatory neurons), Non-tel ExN – Hind (non-telencephalic hindbrain excitatory neuron), Non-tel ExN – Mid (non-telencephalic midbrain excitatory neuron), , InhibN (inhibitory neuron), IP (intermediate progenitor), IP-InhibN (inhibitory intermediate progenitor), CR (Cajal-Retzius) cells.
(C) UMAP representation highlighting normalized gene expression of 15q13.3 genes *KLF13*, *FAN1*, *OTUD7A*, CHRNA7, and *MTMR10*.

Supplementary Table 2: Population cell counts in Day 120 UNO scRNA Seq.

Population	**Cell Count**
Apical radial glia	9915
Excitatory mature neurons	8803
Radial glia	6070
Non-telencephalic midbrain excitatory neurons	21177
Non-telencephalic hindbrain excitatory neurons	5405
Cajal-Retzius cells	3955
Inhibitory neurons	3916
Intermediate progenitors and excitatory neurons	5157
Cycling radial glia	2832
Oligodendrocyte precursors	881
Mesenchymal-like cells	345
Total	68456

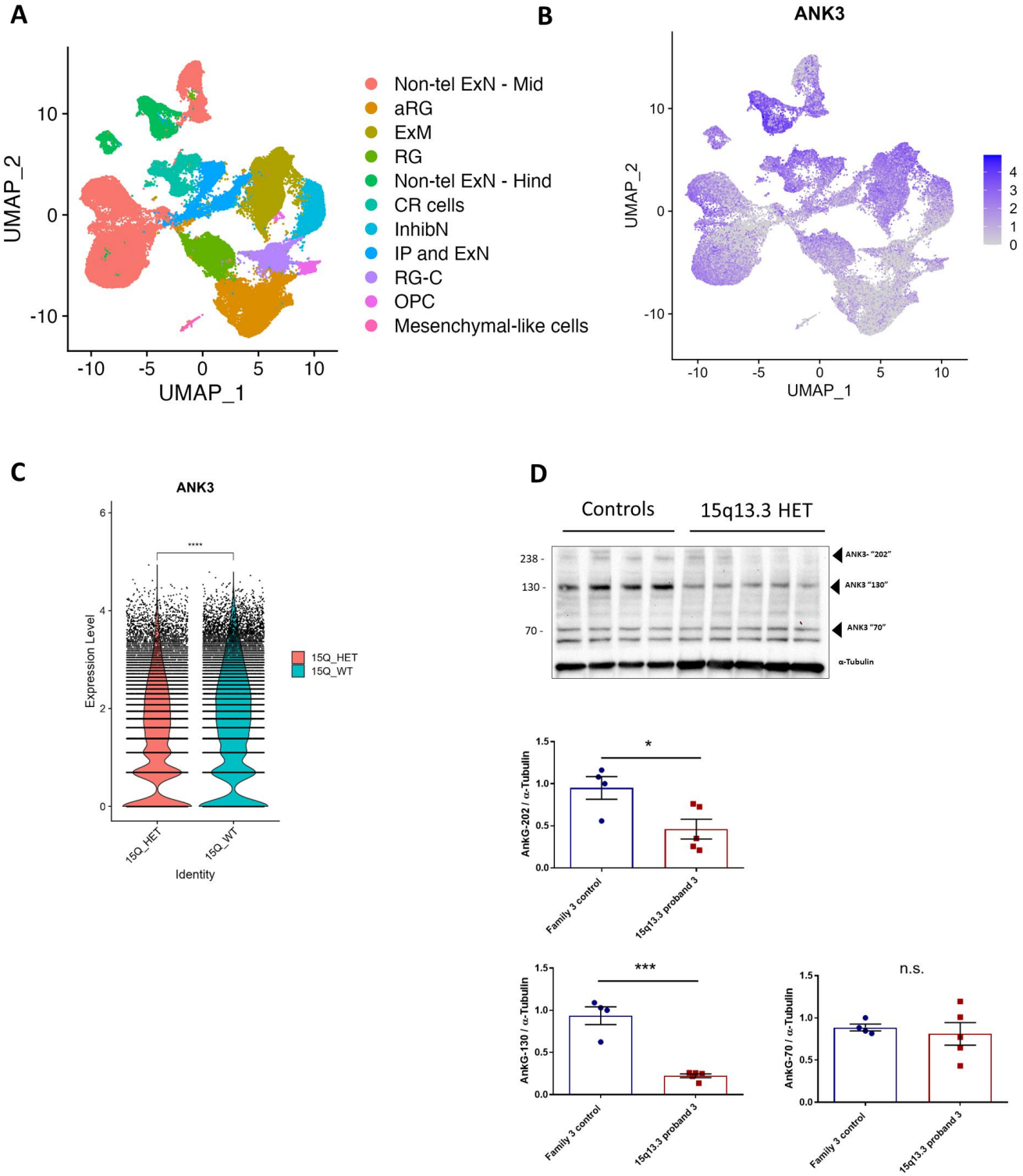

A
B
ANK3
UMAP_2
UMAP_1
Non-tel ExN - Mid
aRG
ExM
RG
Non-tel ExN - Hind
CR cells
InhibN
IP and ExN
RG-C
OPC
Mesenchymal-like cells
C
ANK3
Expression Level
15Q_HET
15Q_WT
Identity
D
Controls
15q13.3 HET
238 -
130 -
70 -
ANK3- "202"
ANK3 "130"
ANK3 "70"
α-Tubulin
AnkG-202 / α-Tubulin
AnkG-130 / α-Tubulin
AnkG-70 / α-Tubulin
Family 3 control
15q13.3 proband 3
n.s.

Figure 12: Validation of scRNA Seq DEG and OTUD7A interactor, ANK3.

(A) Day 120 UMAP plot (0.15 resolution, Control N =31 784 cells, 15q13.3del N = 3 6671 cells) across 3 families. a/o/RG/c (apical/outer/radial glia/cycling), OPC (oligodendrocyte precursor cell), ExM (excitatory mature neuron), IP and ExN (intermediate progenitor and excitatory neurons), Non-tel ExN – Hind (non-telencephalic hindbrain excitatory neuron), Non-tel ExN – Mid (non-telencephalic midbrain excitatory neuron), , InhibN (inhibitory neuron), IP (intermediate progenitor), IP-InhibN (inhibitory intermediate progenitor), CR (Cajal-Retzius) cells.
(B) UMAP representation of normalized gene expression of ANK3.
(C) Aggregate mRNA expression of ANK3 (Ankyrin-G) transcript stratified by genotype ($p < 0.001$, non-parametric Wilcoxon rank sum test).
(D) ANK3 protein is reduced in 2-month 15q13.3 family 3 UNOs. N = WT n = 4, 15q13.3 HET proband = 5. *Data analyzed by two-tailed student's t-test,* n.s = nonsignificant, * p <0.05, *** *p < 0.001, values represent mean ± S.E.M.*

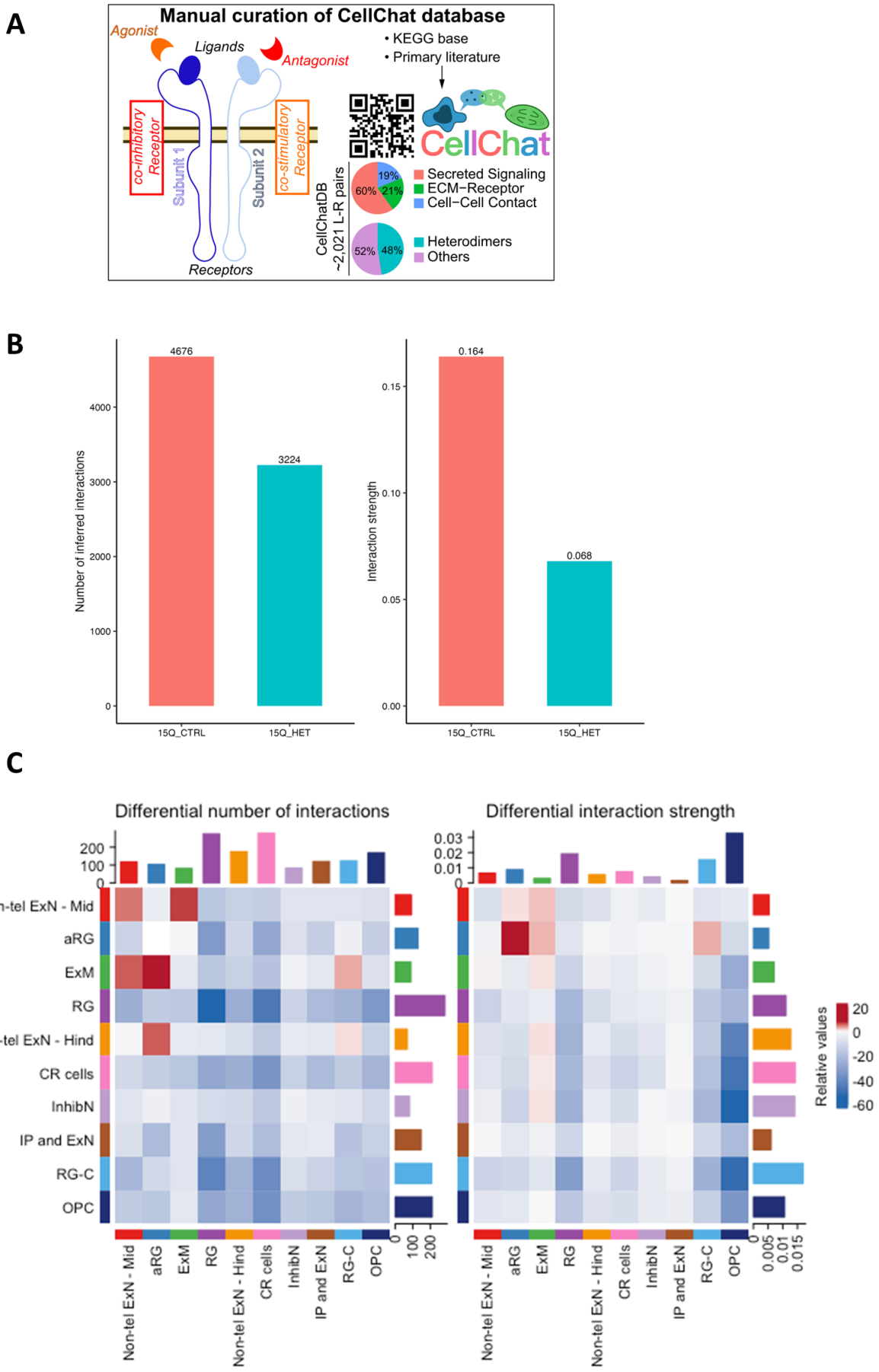

A
Manual curation of CellChat database
Agonist
Ligands
Antagonist
co-inhibitory Receptor
Subunit 1
Subunit 2
co-stimulatory Receptor
Receptors
• KEGG base
• Primary literature
CellChat
CellChatDB
~2,021 L-R pairs
19%
60%
21%
Secreted Signaling
ECM−Receptor
Cell−Cell Contact
52%
48%
Heterodimers
Others
B
Number of inferred interactions
4676
3224
15Q_CTRL
15Q_HET
Interaction strength
0.164
0.068
C
Differential number of interactions
Differential interaction strength
Sources (Sender)
Non-tel ExN - Mid
aRG
ExM
RG
Non-tel ExN - Hind
CR cells
InhibN
IP and ExN
RG-C
OPC
Relative values

Figure 13: CellChat analysis predicts global deficits in cell-cell communication in mature 15q13.3del unguided neural organoids.

(A) Schematic of CellChat and pseudotime processing adapted from *Jin et al., 2021 Nature Communications.*
(B) CellChat predicts reductions in 15q13.3del 4-month UNOs interaction number and strength. (Top) bar plots of weighted interaction number and strength for the predicted ligand-receptor interactions in aggregate. **(C)** Heatmaps show the signal intensity (relative strength) of each pathway in each cell type for outgoing or incoming signaling (y axis = sender/ligand, x axis = recipient/receptor).

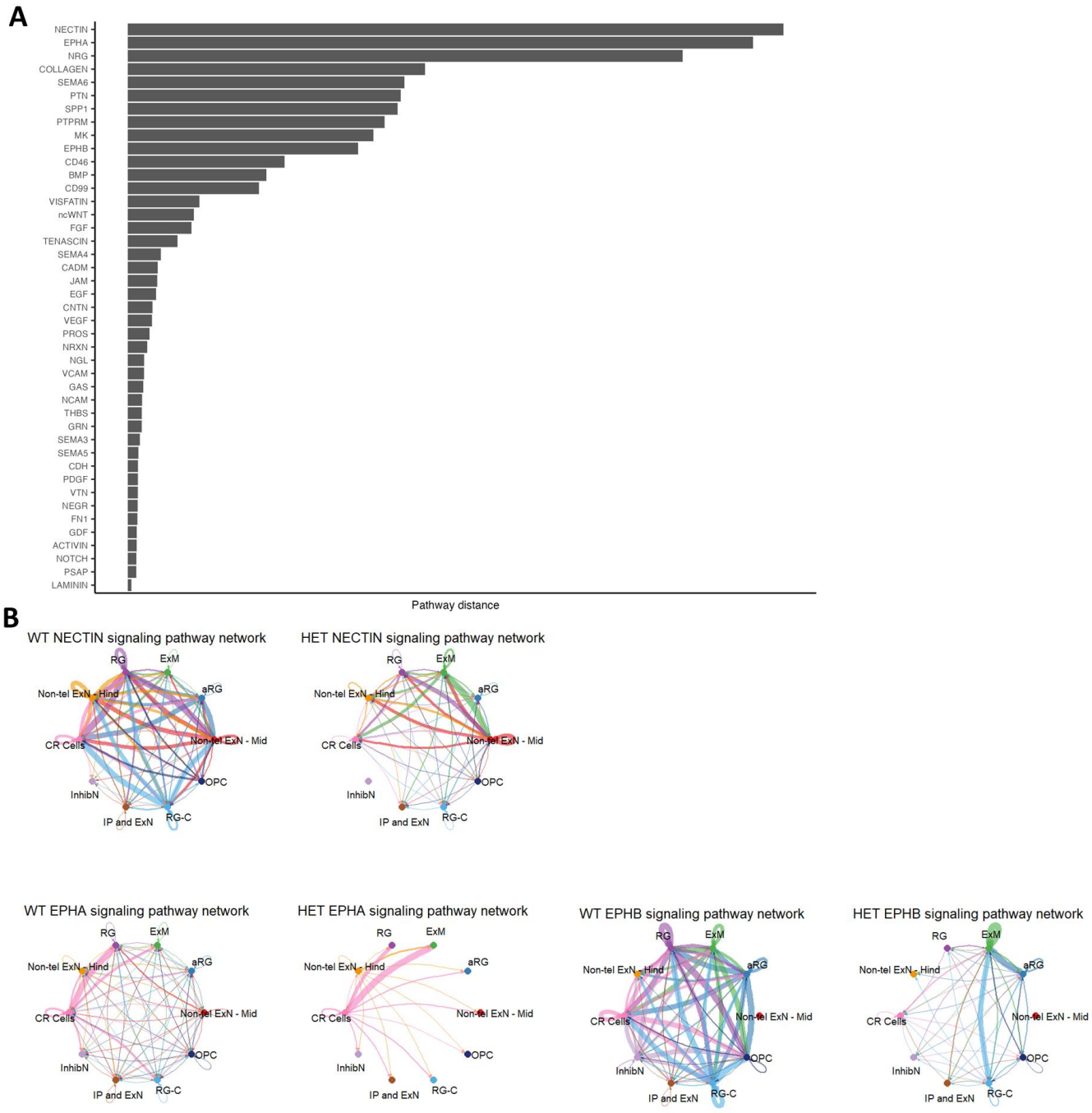
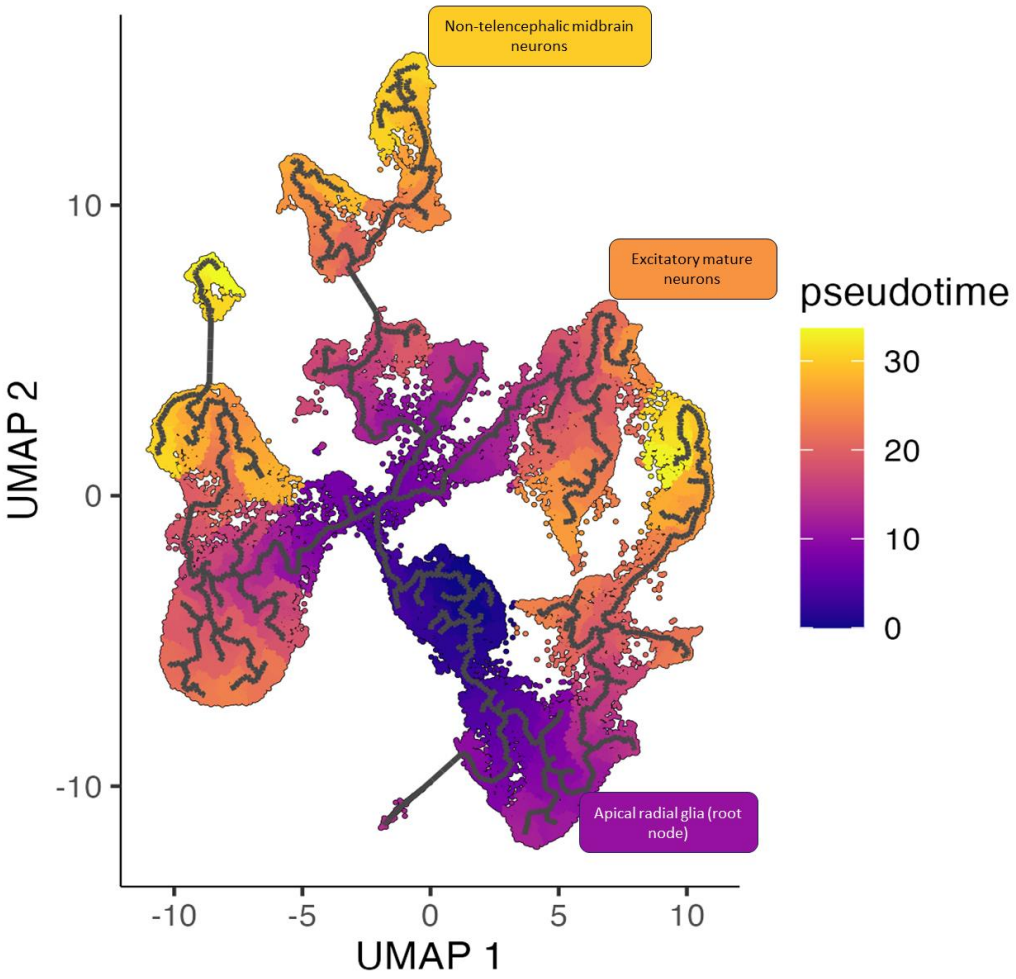

A
NECTIN
EPHA
NRG
COLLAGEN
SEMA6
PTN
SPP1
PTPRM
MK
EPHB
CD46
BMP
CD99
VISFATIN
ncWNT
FGF
TENASCIN
SEMA4
CADM
JAM
EGF
CNTN
VEGF
PROS
NRXN
NGL
VCAM
GAS
NCAM
THBS
GRN
SEMA3
SEMA5
CDH
PDGF
VTN
NEGR
FN1
GDF
ACTIVIN
NOTCH
PSAP
LAMININ
Pathway distance
B
WT NECTIN signaling pathway network
HET NECTIN signaling pathway network
WT EPHA signaling pathway network
HET EPHA signaling pathway network
WT EPHB signaling pathway network
HET EPHB signaling pathway network
RG
ExM
aRG
Non-tel ExN - Hind
CR Cells
Non-tel ExN - Mid
OPC
InhibN
IP and ExN
RG-C
C
Downregulated GO Terms
Radial glia
Excitatory mature neurons
Intermediate progenitors and excitatory neurons
developmental process
regulation of signaling
regulation of cell communication
regulation of cell population proliferation
regulation of cellular process
regulation of response to stimulus
negative regulation of cellular process
response to growth factor
cell migration
regulation of cell motility
locomotion
regulation of plasma membrane bounded cell projection organization
positive regulation of cellular process
response to endogenous stimulus
endothelial cell migration
neuron death
cell adhesion
nervous system development
generation of neurons
eye development
visual system development
regulation of neuron differentiation
central nervous system development
neurogenesis
sensory system development
multicellular organism development
multicellular organismal process
regulation of transcription by RNA polymerase II
head development
brain development
cell differentiation
cellular developmental process
neuron cell-cell adhesion
sensory organ development
neurogenesis
neuron differentiation
cerebellar granule cell differentiation
multicellular organismal process
neuron development
synapse assembly
cell junction assembly
system development
anatomical structure development
cell morphogenesis
modulation of chemical synaptic transmission
developmental process
vocalization behavior
neuron cell-cell adhesion

Figure 14: Deficits in cell adhesion pathways and communication are predicted in aggregate and in a cell type specific manner.

(A) Rank similarity plot shows overrepresentation of common cell adhesion pathways aggregate cellular communication based on joint manifold learning. EPHA (Ephrin-A), EPHB (Ephrin-B)
(B) Circle plots of top adhesion-based cell-cell signaling pathways show aggregate and cell-type specific changes in cell adhesion signaling. Circle size and edge width are proportional to the number of cells in each cluster and the communication score between interacting cell clusters, respectively.
(C) REVIGO treemap plots of over-represented gene ontology terms in downregulated gene sets show enrichment for cell-adhesion biological pathways.

A


Non-telencephalic midbrain neurons
Excitatory mature neurons
Apical radial glia (root node)
pseudotime
30
20
10
0
UMAP 2
UMAP 1
10
0
-10
-10
-5
0
5
10

B

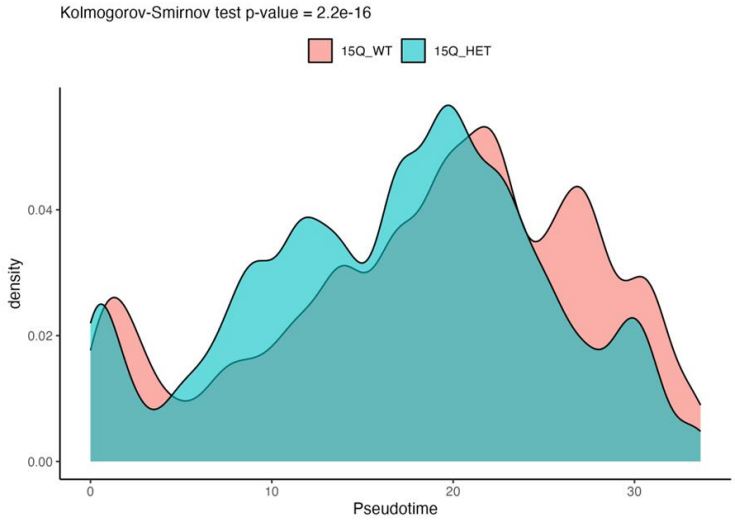

Kolmogorov-Smirnov test p-value = 2.2e-16
15Q_WT
15Q_HET
density
0.04
0.02
0.00
0
10
20
30
Pseudotime

Figure 15: Mature unguided neural organoids have global changes in pseudotemporal trajectories.

(A) Pseudotime UMAP partition plot of cell trajectories, chosen root node = radial glia cluster.
(B) Pseudotime ridge plot reveals abnormal cellular trajectory density amongst 15q13.3del cell types in aggregate (Kolmogorov–Smirnov test p = 2.2e10-16)

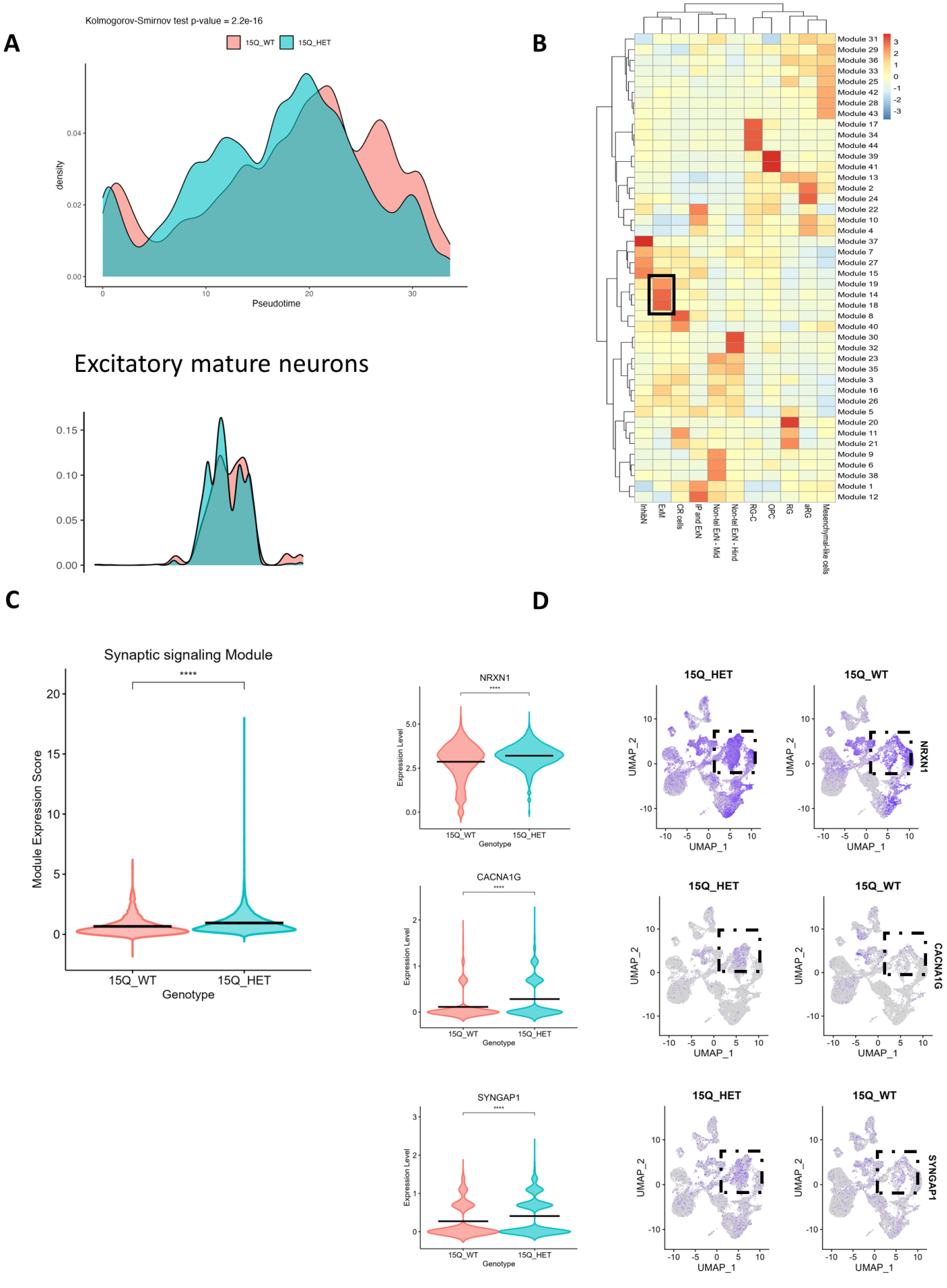

A
Kolmogorov-Smirnov test p-value = 2.2e-16
15Q_WT
15Q_HET
density
Pseudotime
Excitatory mature neurons
B
Module 31
Module 29
Module 36
Module 33
Module 25
Module 42
Module 28
Module 43
Module 17
Module 34
Module 44
Module 39
Module 41
Module 13
Module 2
Module 24
Module 22
Module 10
Module 4
Module 37
Module 7
Module 27
Module 15
Module 19
Module 14
Module 18
Module 8
Module 40
Module 30
Module 32
Module 23
Module 35
Module 3
Module 16
Module 26
Module 5
Module 20
Module 11
Module 21
Module 9
Module 6
Module 38
Module 1
Module 12
InhibN
ExM
CR cells
IP and ExN
Non-tel ExN - Mid
Non-tel ExN - Hind
RG-C
OPC
RG
aRG
Mesenchymal-like cells
C
Synaptic signaling Module
Module Expression Score
Genotype
D
NRXN1
CACNA1G
SYNGAP1
Expression Level
15Q_HET
15Q_WT
UMAP_1
UMAP_2

Figure 16: Excitatory mature neurons have enrichment for synaptic signaling gene modules.

(A) Pseudotime ridge plot reveals abnormal cellular trajectory density amongst 15q13.3del cell types in aggregate (Kolmogorov–Smirnov test p = 2.2e10-16) and amongst excitatory mature neurons (p = 4.28e-09)
(B) Heatmap of pseudotime modules based on aggregate DEG expression with highlighted excitatory neuron-specific (black box) modules 14, 18, and 19.
(C) Gene modules associated with synaptic signaling are enriched in 15q13.3del excitatory mature neurons (Module expression score 15Q_WT = 0.6625506, 15Q_HET = 0.9431373, Wilcox test, p = 9.161339e-63).
(D) Genes within synaptic transmission module are increased in 15q13.3del excitatory mature neurons (Wilcox test, CACNA1G (p= 8.421278e-41), NRXN1 (p = 1.982136e-18), SYNGAP1 (p=7.984738e-53).

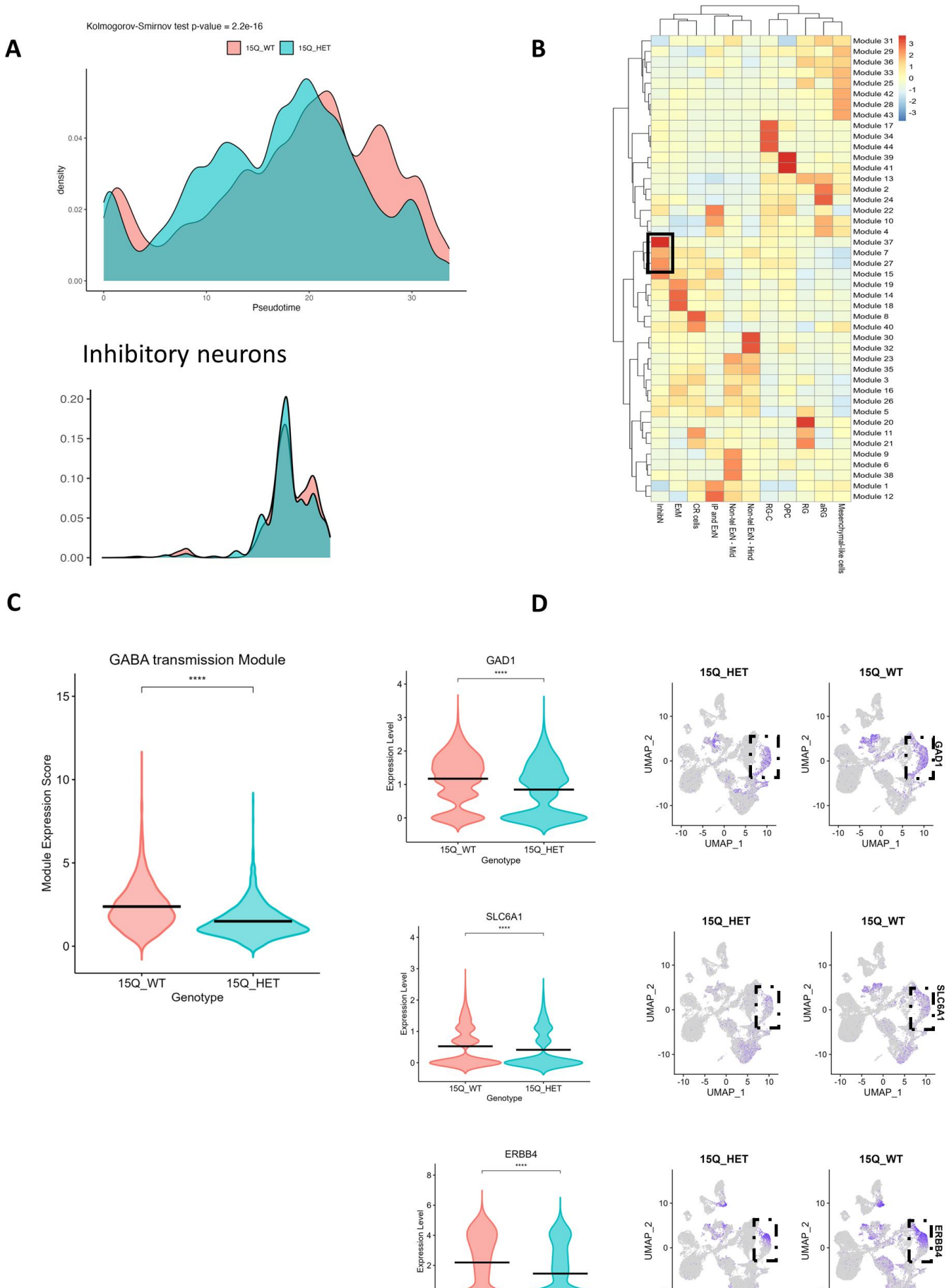

A
Kolmogorov-Smirnov test p-value = 2.2e-16
15Q_WT
15Q_HET
density
Pseudotime
Inhibitory neurons
B
Module 31
Module 29
Module 36
Module 33
Module 25
Module 42
Module 28
Module 43
Module 17
Module 34
Module 44
Module 39
Module 41
Module 13
Module 2
Module 24
Module 22
Module 10
Module 4
Module 37
Module 7
Module 27
Module 15
Module 19
Module 14
Module 18
Module 8
Module 40
Module 30
Module 32
Module 23
Module 35
Module 3
Module 16
Module 26
Module 5
Module 20
Module 11
Module 21
Module 9
Module 6
Module 38
Module 1
Module 12
InhibN
ExM
CR cells
IP and ExN
Non-tel ExN - Mid
Non-tel ExN - Hind
RG-C
OPC
RG
aRG
Mesenchymal-like cells
C
GABA transmission Module
Module Expression Score
Genotype
D
GAD1
SLC6A1
ERBB4
Expression Level
15Q_HET
15Q_WT
UMAP_1
UMAP_2

Figure 17: GABAergic transmission modules are disrupted in 15q13.3dle inhibitory neuron populations.

(A) Pseudotime ridge plot reveals abnormal cellular trajectory density amongst 15q13.3del cell types in aggregate (Kolmogorov–Smirnov test p = 2.2e10-16) and amongst inhibitory neurons (p = 4.03e-05)
(B) Heatmap of pseudotime modules based on aggregate DEG expression with highlighted inhibitory neuron-specific (black box) modules 15, 27, and 37.
(C) Gene modules associated with GABAergic transmission are reduced in 15q13.3del inhibitory neurons (Module expression score 15Q_WT = 2.379354, 15Q_HET = 1.497246, Wilcox test, p = 1.495263e-101).
(D) Genes within GABAergic transmission module are decreased in 15q13.3del inhibitory neurons (Wilcox test, GAD1 (p= 1.257692e-34), SLC6A1 (p=8.09127e-09), ERBB4 (p = 3.245731e-33).

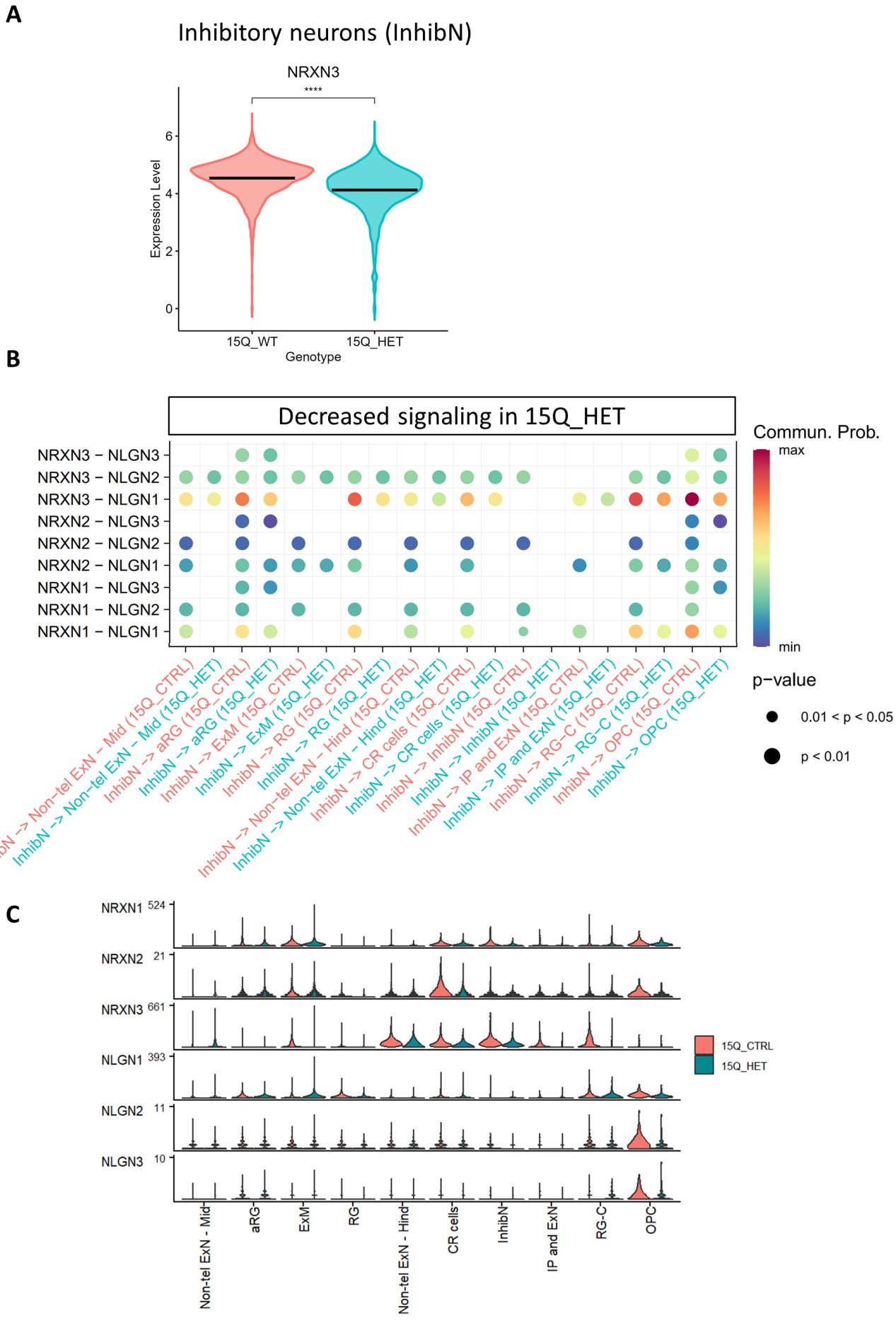
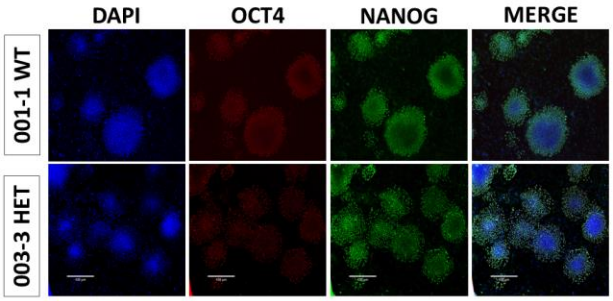

A
Inhibitory neurons (InhibN)
NRXN3
****
Expression Level
0
2
4
6
15Q_WT
15Q_HET
Genotype
B
Decreased signaling in 15Q_HET
NRXN3 – NLGN3
NRXN3 – NLGN2
NRXN3 – NLGN1
NRXN2 – NLGN3
NRXN2 – NLGN2
NRXN2 – NLGN1
NRXN1 – NLGN3
NRXN1 – NLGN2
NRXN1 – NLGN1
InhibN –> Non–tel ExN – Mid (15Q_CTRL)
InhibN –> Non–tel ExN – Mid (15Q_HET)
InhibN –> aRG (15Q_CTRL)
InhibN –> aRG (15Q_HET)
InhibN –> ExM (15Q_CTRL)
InhibN –> ExM (15Q_HET)
InhibN –> RG (15Q_CTRL)
InhibN –> RG (15Q_HET)
InhibN –> Non–tel ExN – Hind (15Q_CTRL)
InhibN –> Non–tel ExN – Hind (15Q_HET)
InhibN –> CR cells (15Q_CTRL)
InhibN –> CR cells (15Q_HET)
InhibN –> InhibN (15Q_CTRL)
InhibN –> InhibN (15Q_HET)
InhibN –> IP and ExN (15Q_CTRL)
InhibN –> IP and ExN (15Q_HET)
InhibN –> RG–C (15Q_CTRL)
InhibN –> RG–C (15Q_HET)
InhibN –> OPC (15Q_CTRL)
InhibN –> OPC (15Q_HET)
Commun. Prob.
max
min
p–value
0.01 < p < 0.05
p < 0.01
C
NRXN1 524
NRXN2 21
NRXN3 661
NLGN1 393
NLGN2 11
NLGN3 10
15Q_CTRL
15Q_HET
Non–tel ExN - Mid
aRG
ExM
RG
Non–tel ExN - Hind
CR cells
InhibN
IP and ExN
RG–C
OPC

Figure 18: NRXN-NLGN signaling is significantly reduced in 15q13.3del inhibitory neurons.

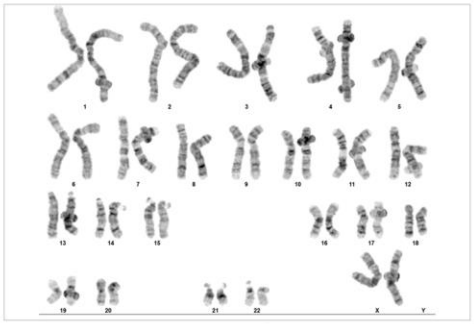
(A) NRXN3 mRNA is significantly reduced in 15q13.3del inhibitory neurons (Wilcox test, p = 1.679445e-68)
(B) Dot plot of significantly decreased NRXN-NLGN signaling in 15q13.3del inhibitory neurons. The dot color and size represent the communication probability and p-values, respectively. p-values were computed from one-sided permutation test.
(C) Normalized mRNA expression of NRXN-NLGN signaling components.

3.6 Chapter 3 Supplementary Data

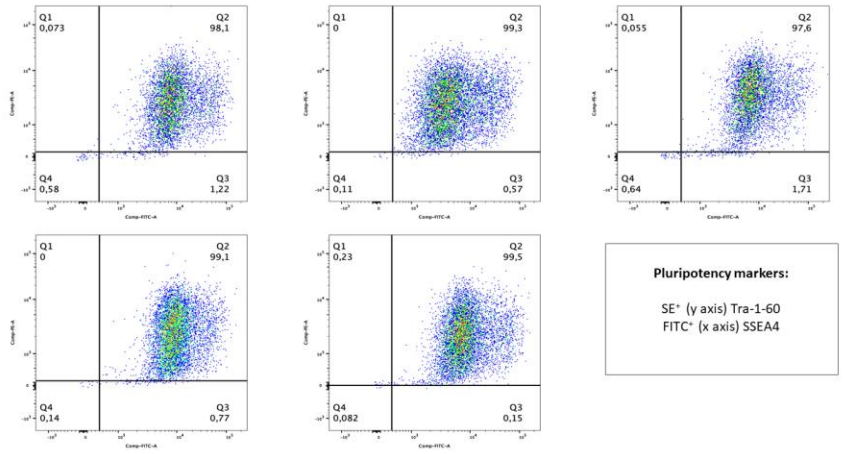
A



B



C

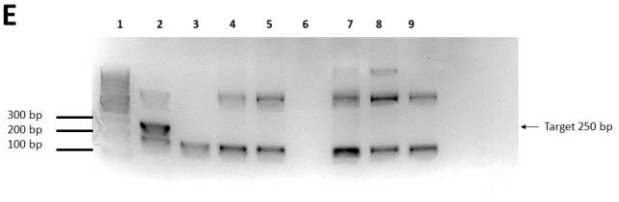


D

Segments Details

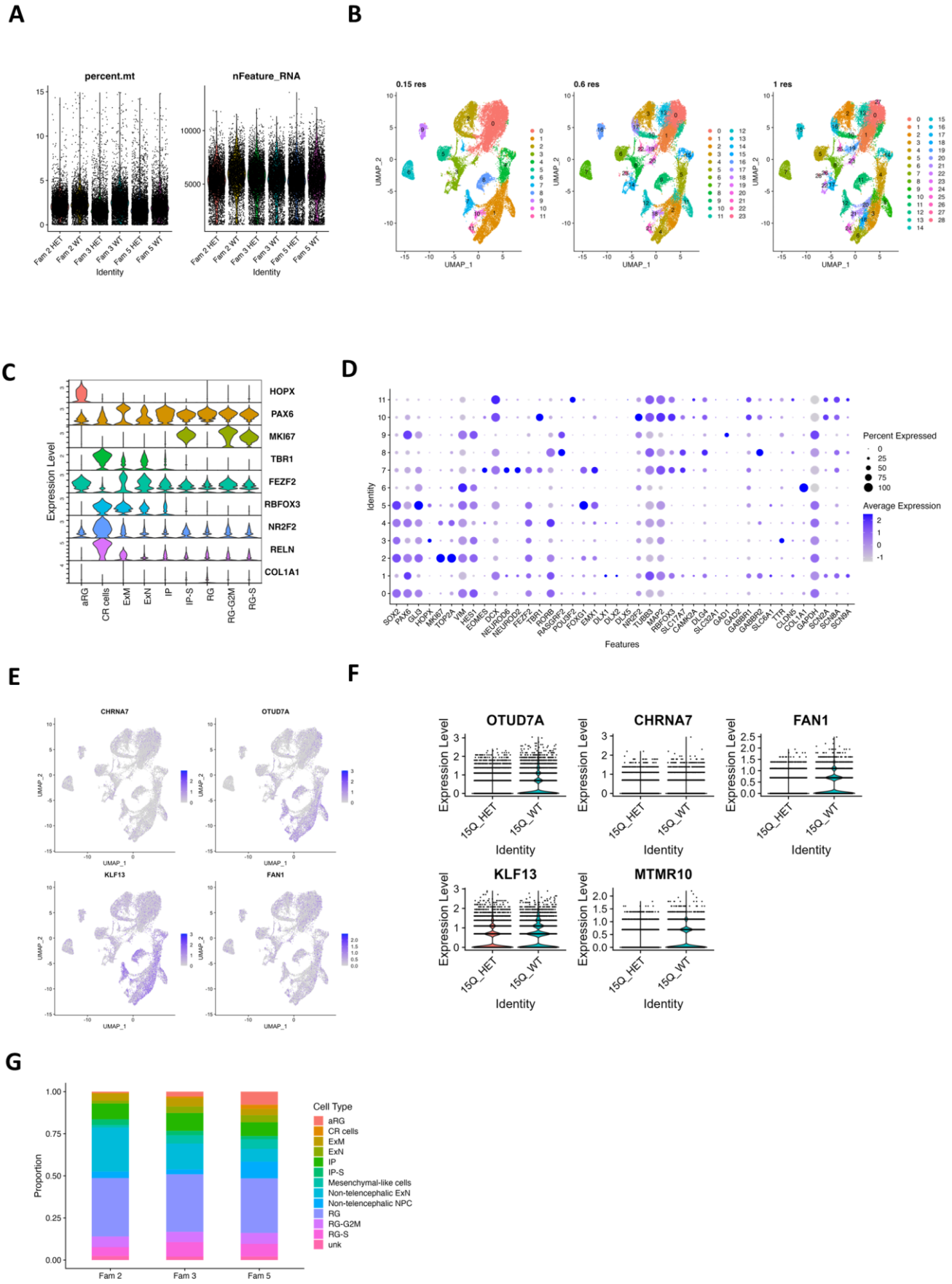
Sample	CN State	Type	CHR	Cytoband Start	Cytoband End	Size (kbp)
KS01_Fam 1 WT 89E	1	Loss	16	p11.2	p11.2	1486.116
KS02_Fam 1 HET 90N	1	Loss	2	q13	q13	861.678
KS02_Fam 1 HET 90N	3	Gain	7	q11.22	q11.22	815.515
KS02_Fam 1 HET 90N	3	Gain	14	q32.33	q32.33	407.886
KS02_Fam 1 HET 90N	**1**	**Loss**	**15**	**q13.2**	**q13.3**	**1530.69**
KS02_Fam 1 HET 90N	1	Loss	16	p11.2	p11.2	1486.116
KS02_Fam 1 HET 90N	3	Gain	18	q22.1	q22.1	439.799
KS03_Fam 2 WT 88T	3	Gain	14	q32.33	q32.33	417.131
KS03_Fam 2 WT 88T	3	Gain	17	q21.31	q21.31	600.249
KS04_Fam 2 CAR 86S	**1**	**Loss**	**15**	**q13.2**	**q13.3**	**1533.259**
KS04_Fam 2 CAR 86S	4	Gain	17	q21.31	q21.31	594.061
KS05_Fam 2 HET 87V	3	Gain	14	q32.33	q32.33	427.145
KS05_Fam 2 HET 87V	**1**	**Loss**	**15**	**q13.2**	**q13.3**	**2000.669**
KS05_Fam 2 HET 87V	1	Loss	16	p11.2	p11.2	1980.47
KS05_Fam 2 HET 87V	3	Gain	17	q21.31	q21.31	596.672
KS05_Fam 2 HET 87V	3	Gain	20	q11.21	q11.21	684.601
KS05_Fam 2 HET 87V	3	Gain	X	q23	q23	524.094
KS06_Fam 3 WT 0205-002-2	3	Gain	14	q32.33	q32.33	407.847
KS07_Fam 3 HET 0205-003-1	3	Gain	14	q32.33	q32.33	427.145
KS07_Fam 3 HET 0205-003-1	**1**	**Loss**	**15**	**q13.2**	**q13.3**	**1348.142**
KS07_Fam 3 HET 0205-003-1	3	Gain	17	q21.31	q21.31	559.66

E



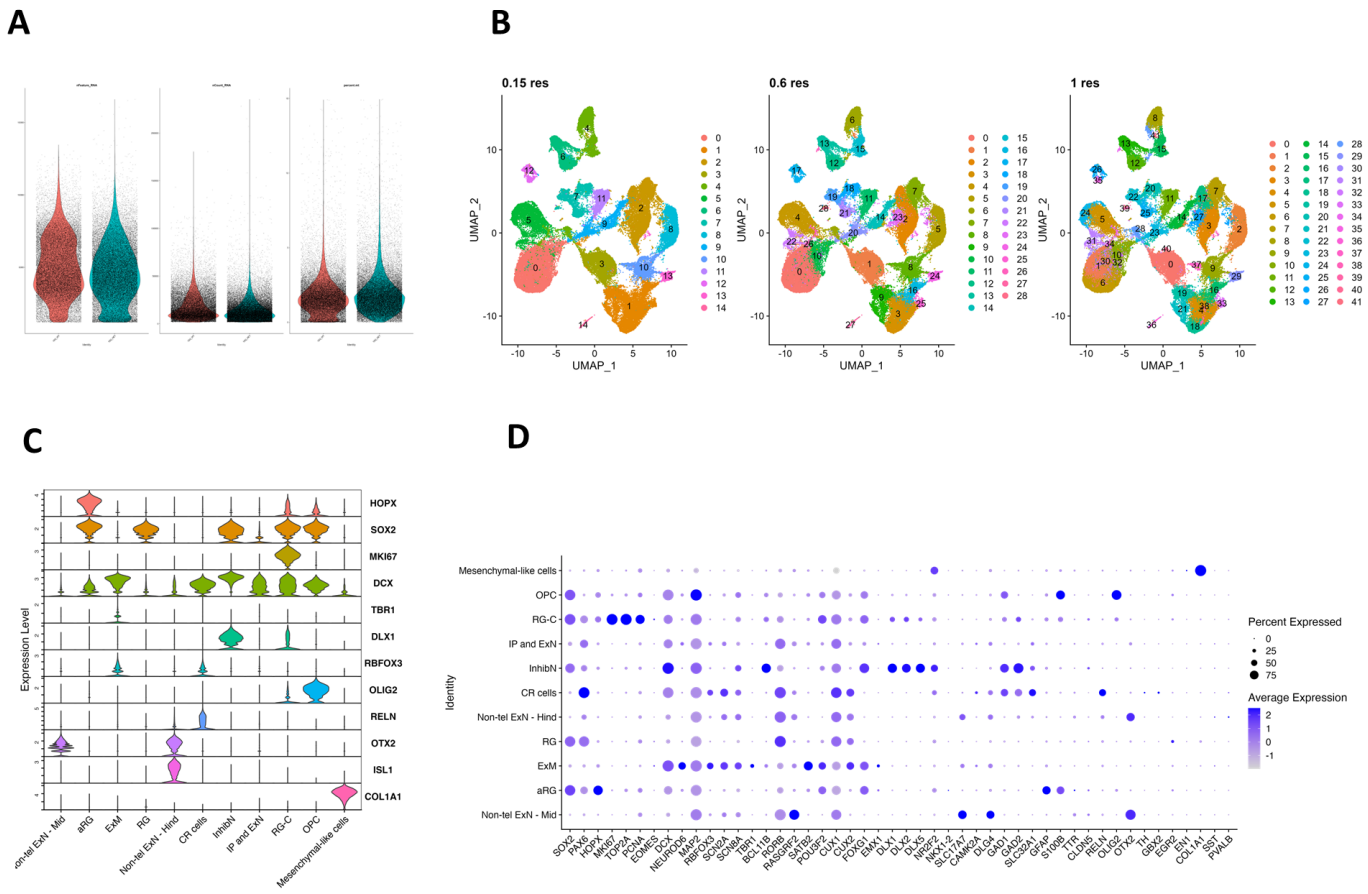
Supplementary Figure 1: Validation of iPSC reprogramming and cellular pluripotency.

(A) Representative pluripotency immunocytochemistry on reprogrammed iPSCs co-cultured with mouse embryonic fibroblasts. DAPI, OCT4-Cy3, NANOG-488. Images taken at 20x on an epifluorescent microscope. Scalebars represent 430 µm.
(B) Representative G-Band analysis on reprogrammed iPSCs show no karyotypic abnormalities.
(C) Microarray genotyping of three 15q13.3del families through affymetrix's cytoscan HD array.
(D) Flow cytometry in P17 reprogrammed iPSCs in 5 unique lines for pluripotency markers Tra-1-60 (SE labeled, y axis) and SSEA4 (FITC labeled, x axis). Data generated by Annie Cheng.
(E) 15q family 3 iPSCs are negative for mycoplasma detection (lanes 1-9: NEB 100bp ladder, mycoplasma positive control, negative control, WT #1, HET#1, WT#2, HET#2, HET#3,). 50 ng of DNA was loaded on a 2% agarose gel and imaged on a BioRad ChemiDoc imaging system. Data generated by Yi Pan.


A
percent.mt
nFeature_RNA
Fam 2 HET
Fam 2 WT
Fam 3 HET
Fam 3 WT
Fam 5 HET
Fam 5 WT
Identity
B
0.15 res
0.6 res
1 res
UMAP_1
UMAP_2
C
Expression Level
HOPX
PAX6
MKI67
TBR1
FEZF2
RBFOX3
NR2F2
RELN
COL1A1
aRG
CR cells
ExM
ExN
IP
IP-S
RG
RG-G2M
RG-S
D
Identity
Features
Percent Expressed
Average Expression
E
CHRNA7
OTUD7A
KLF13
FAN1
F
OTUD7A
CHRNA7
FAN1
KLF13
MTMR10
15Q_HET
15Q_WT
G
Proportion
Cell Type
aRG
CR cells
ExM
ExN
IP
IP-S
Mesenchymal-like cells
Non-telencephalic ExN
Non-telencephalic NPC
RG
RG-G2M
RG-S
unk
Fam 2
Fam 3
Fam 5

Supplementary Figure 2: Unguided neural organoid scRNA sequencing at the day 40 timepoint passes in house quality control measures.

(**A**): Percent mitochondrial read content (left) and read quality (right) in day 40 raw dataset.
(**B**) UMAP Plot of clustering resolution (res) ranges.
(C) Violin plot of canonical cell type marker expression across cell cluster types.
(D) Dot plot of canonical cell type marker expression across cell cluster types.
(E-F) 15q13.3 gene UMAPs (**E**) and violin plots (**F**) highlight cell type specific expression and patient heterozygosity.
(G) Population dynamics remain consistent between 15q13.3 families.


A
B
0.15 res
0.6 res
1 res
UMAP_1
UMAP_2
C
Expression Level
HOPX
SOX2
MKI67
DCX
TBR1
DLX1
RBFOX3
OLIG2
RELN
OTX2
ISL1
COL1A1
aRG
ExM
RG
CR cells
InhibN
IP and ExN
RG-C
OPC
Mesenchymal-like cells
D
Identity
Percent Expressed
Average Expression
E
Proportion
Fam 2 HET D120
Fam 2 WT D120
Fam 3 HET D120
Fam 3 WT D120
Fam 5 HET D120
Fam 5 WT D120
Fam 2
Fam 3
Fam 5
Cell Type
Non-tel ExN - Hind
Non-tel ExN - Mid
F
OTUD7A
CHRNA7
FAN1
KLF13
MTMR10
15Q_HET
15Q_WT
Identity

Supplementary Figure 3: Unguided neural organoid scRNA sequencing at the day 120 timepoint passes in house quality control measures.

(**A**): Percent mitochondrial read content (left) and read quality (right) in day 120 raw dataset.

(**B**) UMAP Plot of clustering resolution (res) ranges.

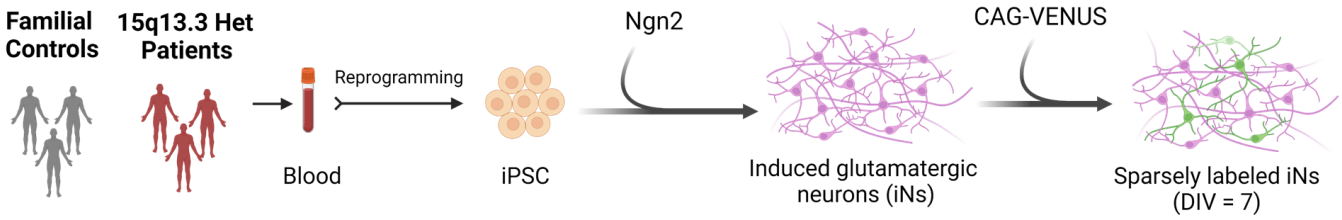
(C) Violin plot of canonical cell type marker expression across cell cluster types.

(D) Dot plot of canonical cell type marker expression across cell cluster types.

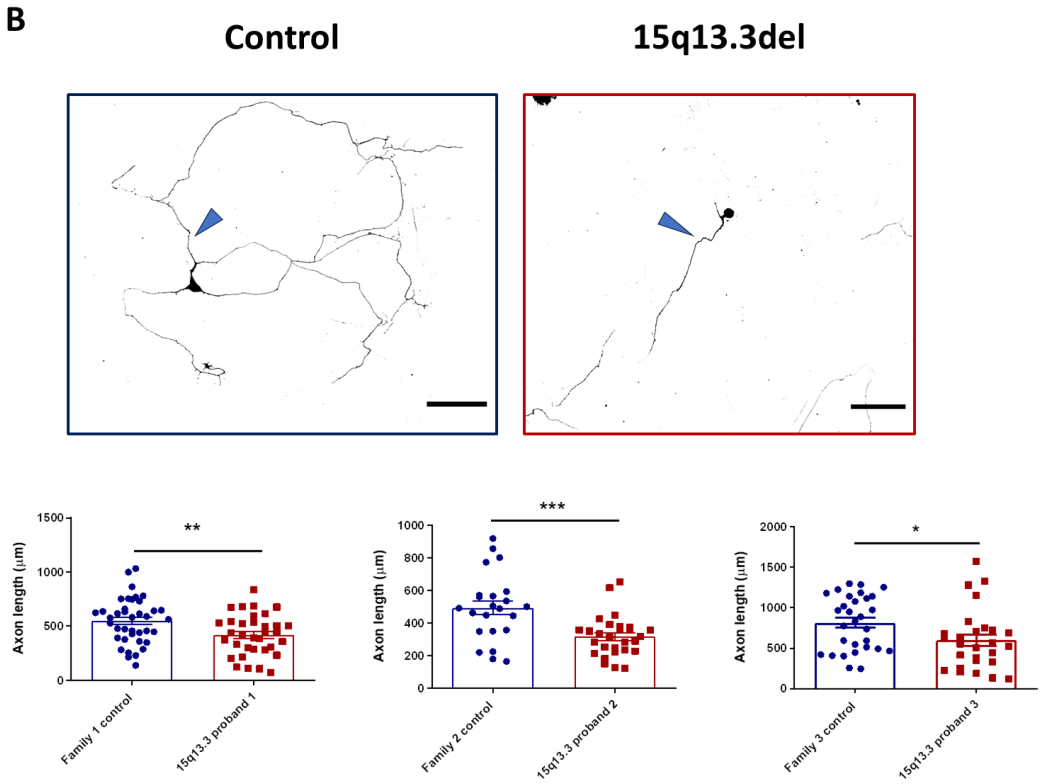
(E) Population dynamics are largely influenced by familial background

(F) 15q13.3 gene violin plots highlight patient heterozygosity.

A

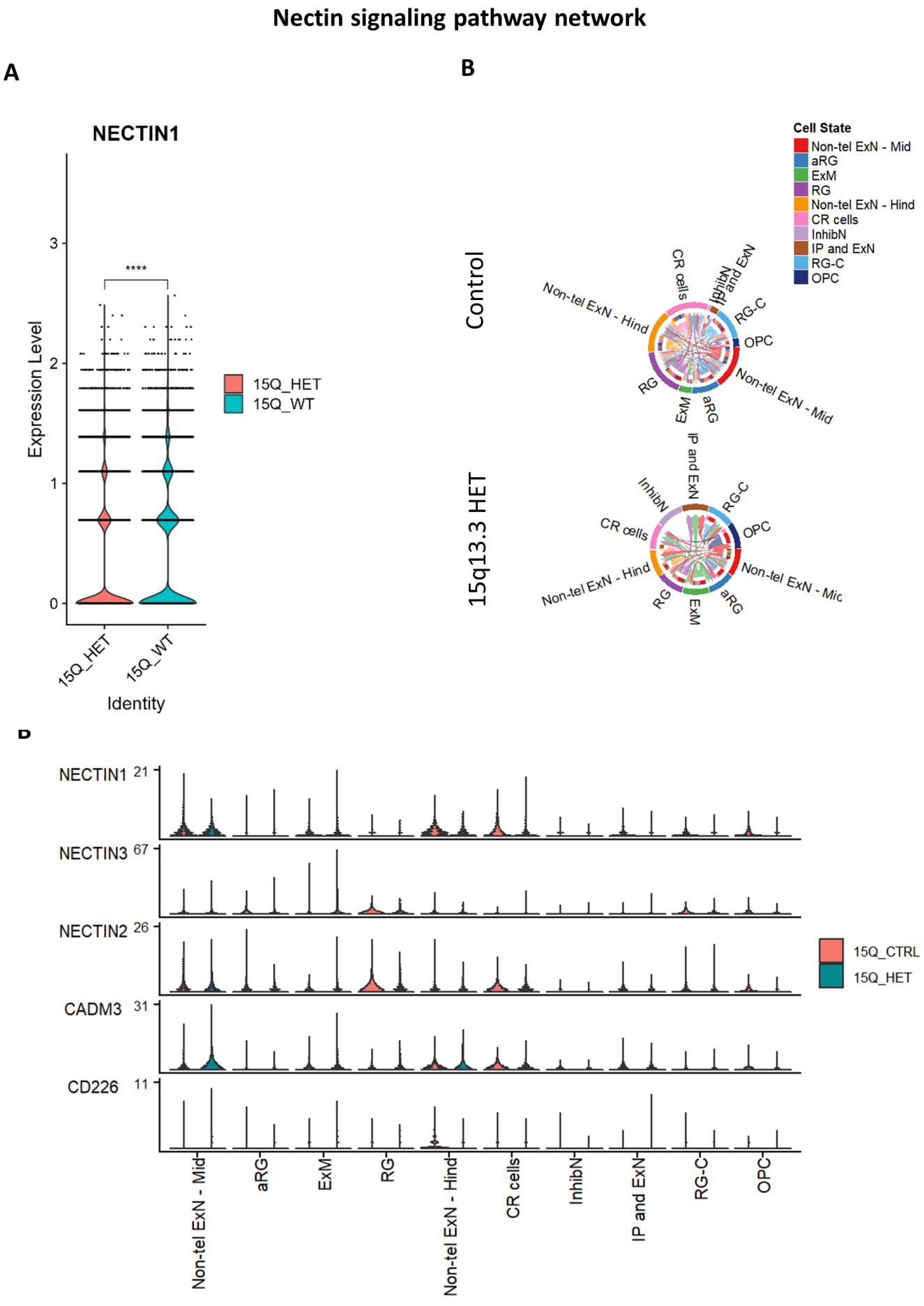

Familial Controls
15q13.3 Het Patients
Reprogramming
Blood
iPSC
Ngn2
Induced glutamatergic neurons (iNs)
CAG-VENUS
Sparsely labeled iNs (DIV = 7)

B


Control
15q13.3del
Axon length (µm)
**
Family 1 control
15q13.3 proband 1
***
Family 2 control
15q13.3 proband 2
*
Family 3 control
15q13.3 proband 3

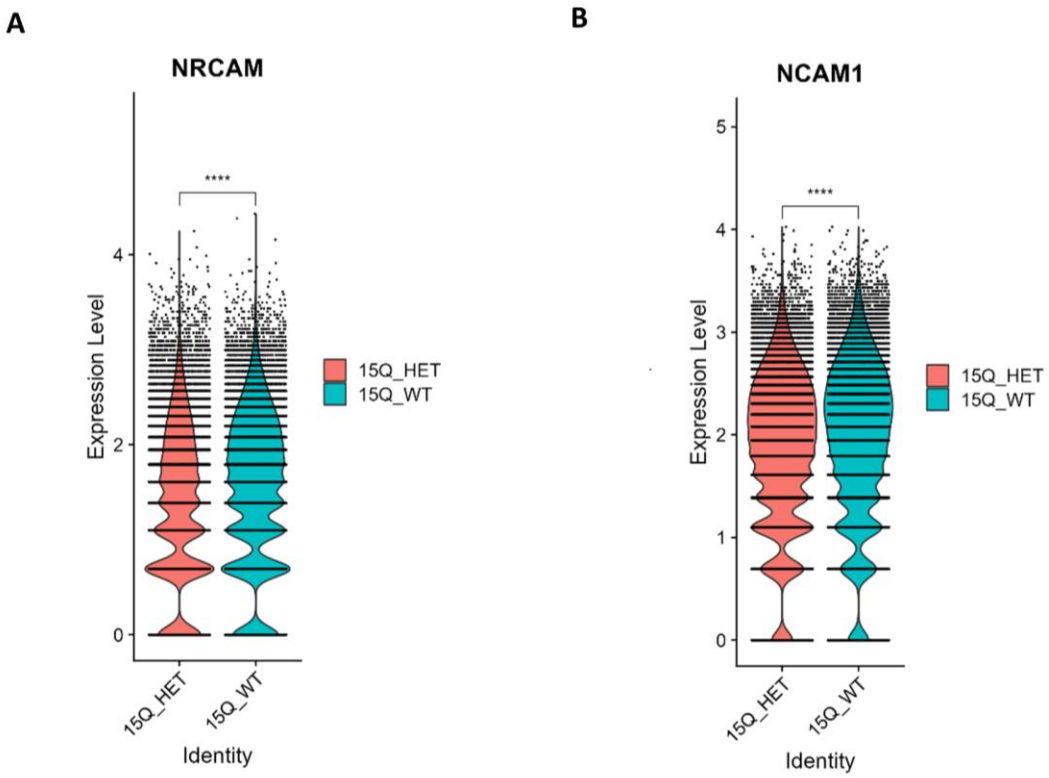
Supplementary Figure 4: 15q13.3del 2D neurogenin-2 (NGN2) neurons show reduced axon projection capacity.

(A) Transfection schematic to sparsely label early NGN2 neurons. *Day in vitro* (DIV) **(B)** (Top) Reconstructed tracings and (B) quantified axon length of Day 7 NGN2 VENUS transfected neurons with arrows denoting the axon (blue familial control, red 15q13.3 proband, and black arrows denoting the axon). Scalebars represent 100 μm. Data analyzed by student's *t-test*, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, values represent mean ± S.E.M.


Nectin signaling pathway network
A
NECTIN1
Expression Level
Identity
15Q_HET
15Q_WT
****
B
Control
15q13.3 HET
Cell State
Non-tel ExN - Mid
aRG
ExM
RG
Non-tel ExN - Hind
CR cells
InhibN
IP and ExN
RG-C
OPC
NECTIN1
NECTIN3
NECTIN2
CADM3
CD226
15Q_CTRL
15Q_HET

Supplementary Figure 5: Nectin signaling is disrupted on a cell-type specific basis in 15q13.3del UNOs.

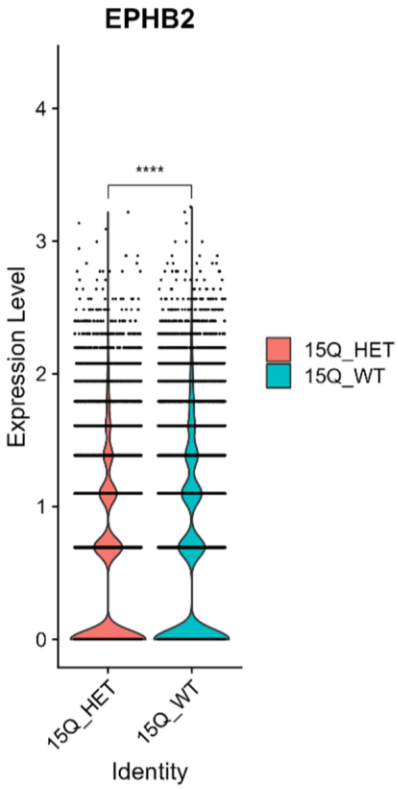
(A) Violin plots of normalized Nectin1 mRNA expression in aggregate D120 15q13.3del UNO tissue. ****p<0.0001, non-parametric Wilcoxon rank sum test.
(B) Chord plots of control (top) and 15q13.3del (bottom) nectin signaling.
(C) Violin plot of the expression distribution of signaling genes involved in the inferred nectin signaling network.



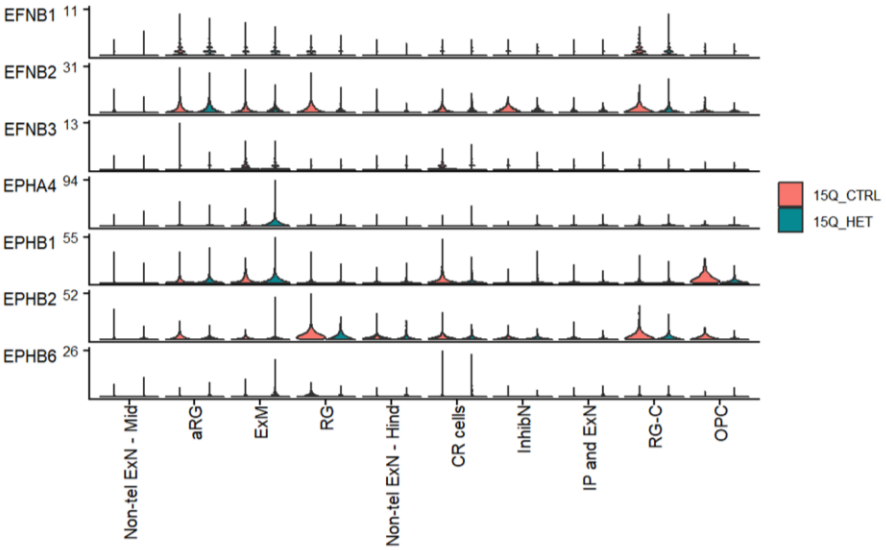
Supplementary Figure 6: Cell adhesion molecules NRCAM and NCAM1 are significantly reduced in D120 15q13.3del UNOs.

Violin plots of normalized (**A**) NRCAM and **(B)** NCAM1 expression in aggregate D120 15q13.3del UNO tissue. ****p<0.0001, non-parametric Wilcoxon rank sum test.

A

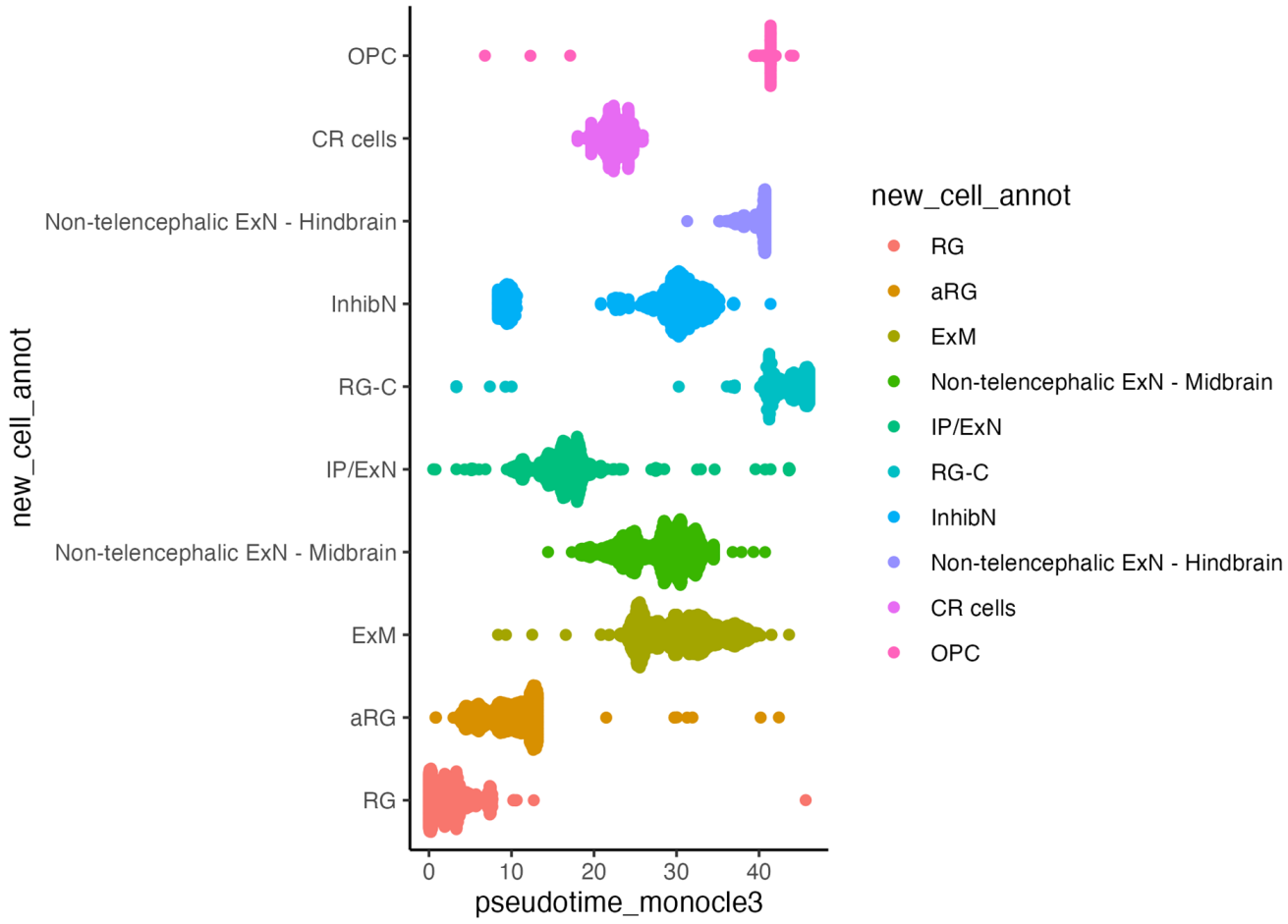

EPHB2
****
Expression Level
0
1
2
3
4
15Q_HET
15Q_WT
15Q_HET
15Q_WT
Identity

B


EFNB1 11
EFNB2 31
EFNB3 13
EPHA4 94
EPHB1 55
EPHB2 52
EPHB6 26
15Q_CTRL
15Q_HET
Non-tel ExN - Mid
aRG
ExM
RG
Non-tel ExN - Hind
CR cells
InhibN
IP and ExN
RG-C
OPC

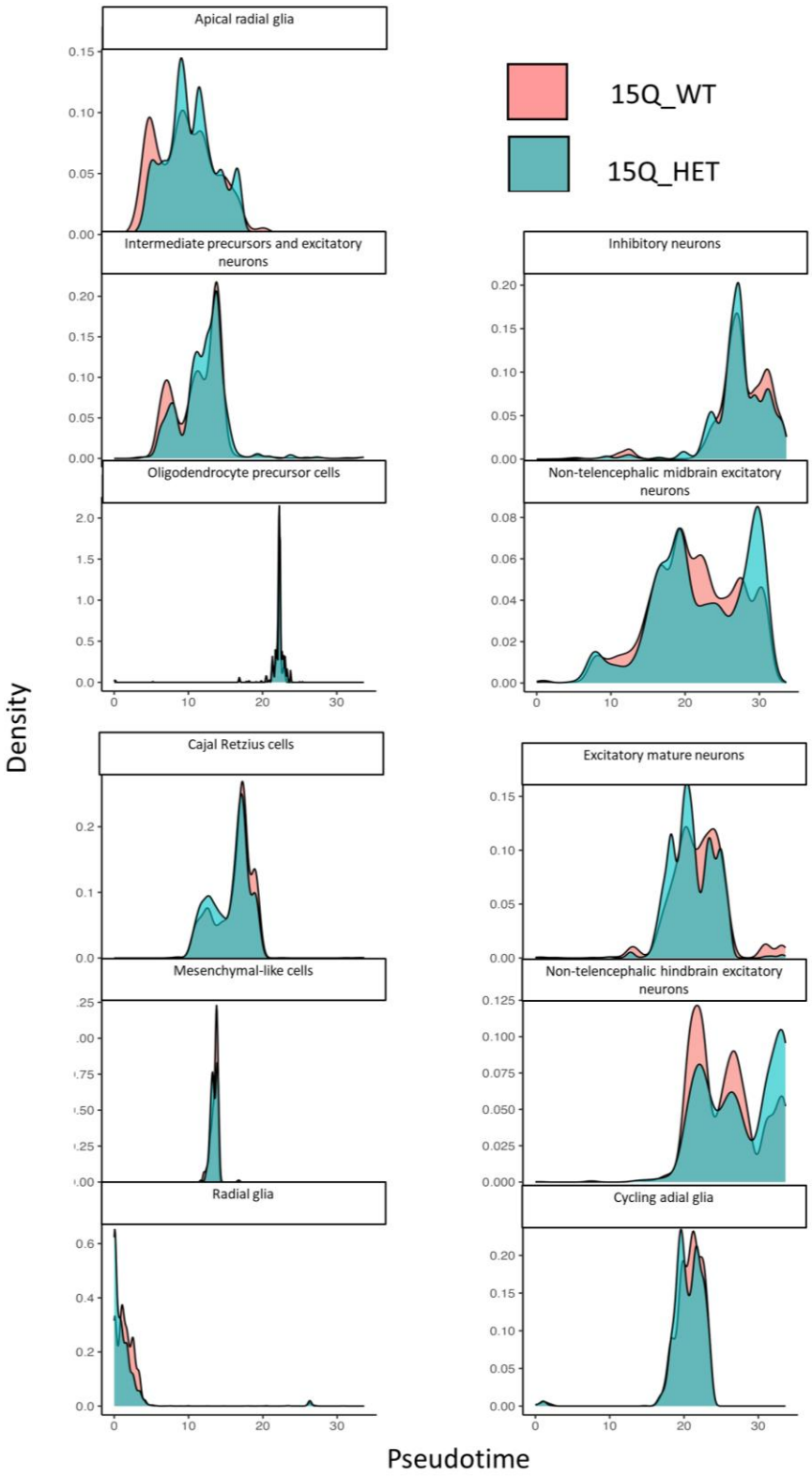
Supplementary Figure 7: EphrinB signaling is disrupted on a cell-type specific basis in 15q13.3del UNOs.

(A**)** Violin plots of normalized EphrinB2 mRNA expression in aggregate D120 15q13.3del UNO tissue. ****p<0.0001, non-parametric Wilcoxon rank sum test. **(B)** Violin plot of the expression distribution of signaling genes involved in the inferred EphrinB signaling network.



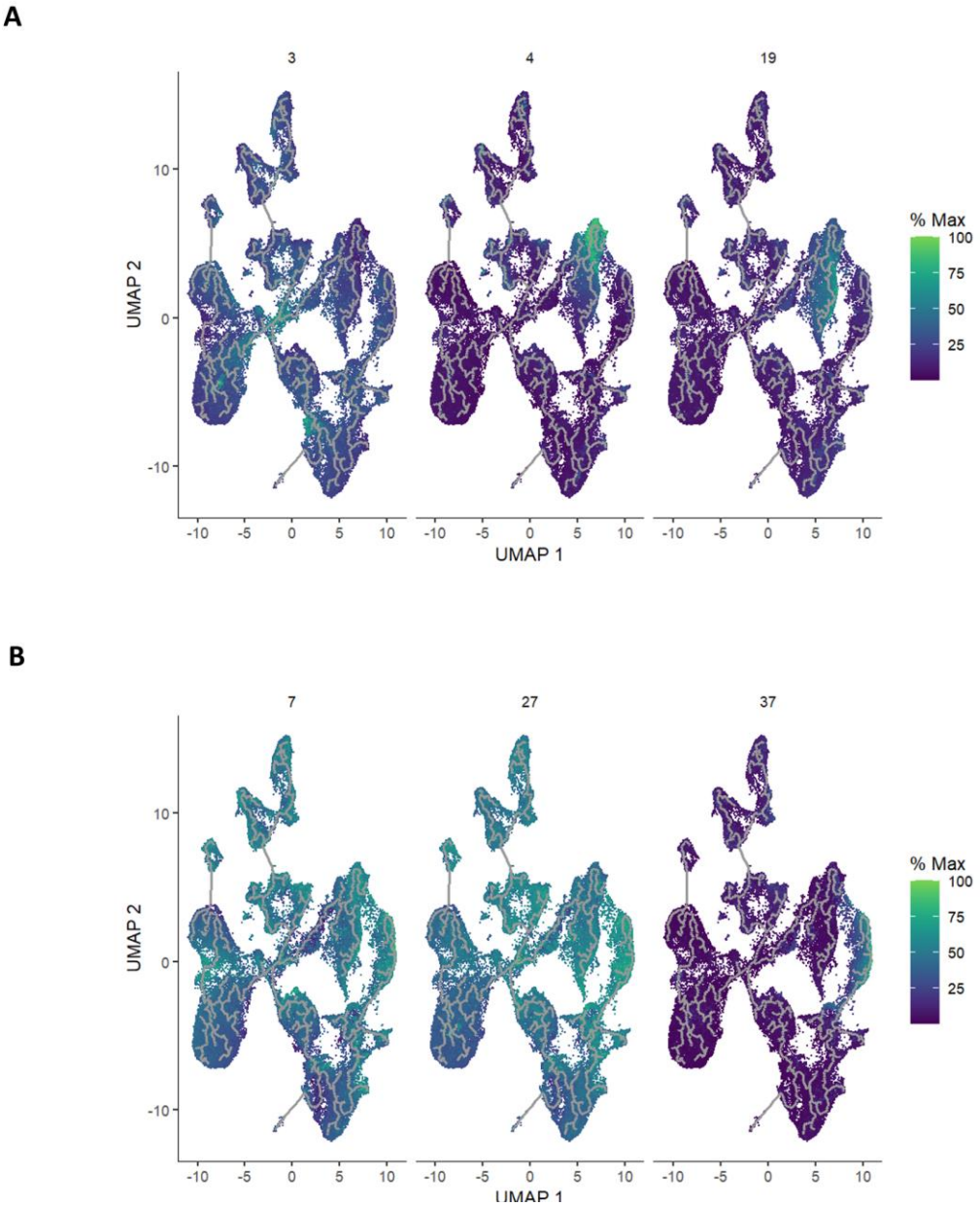
Supplementary Figure 8: Visualization of predicted cell type emergence shows progenitor to mature neuron trajectory (Beeswarm plot).

Beeswarm plot of log-fold change (x-axis) in cell abundance across developmental stages in neighborhoods of radial glial cells.


Apical radial glia
15Q_WT
15Q_HET
Intermediate precursors and excitatory neurons
Inhibitory neurons
Oligodendrocyte precursor cells
Non-telencephalic midbrain excitatory neurons
Cajal Retzius cells
Excitatory mature neurons
Mesenchymal-like cells
Non-telencephalic hindbrain excitatory neurons
Radial glia
Cycling adial glia
Density
Pseudotime

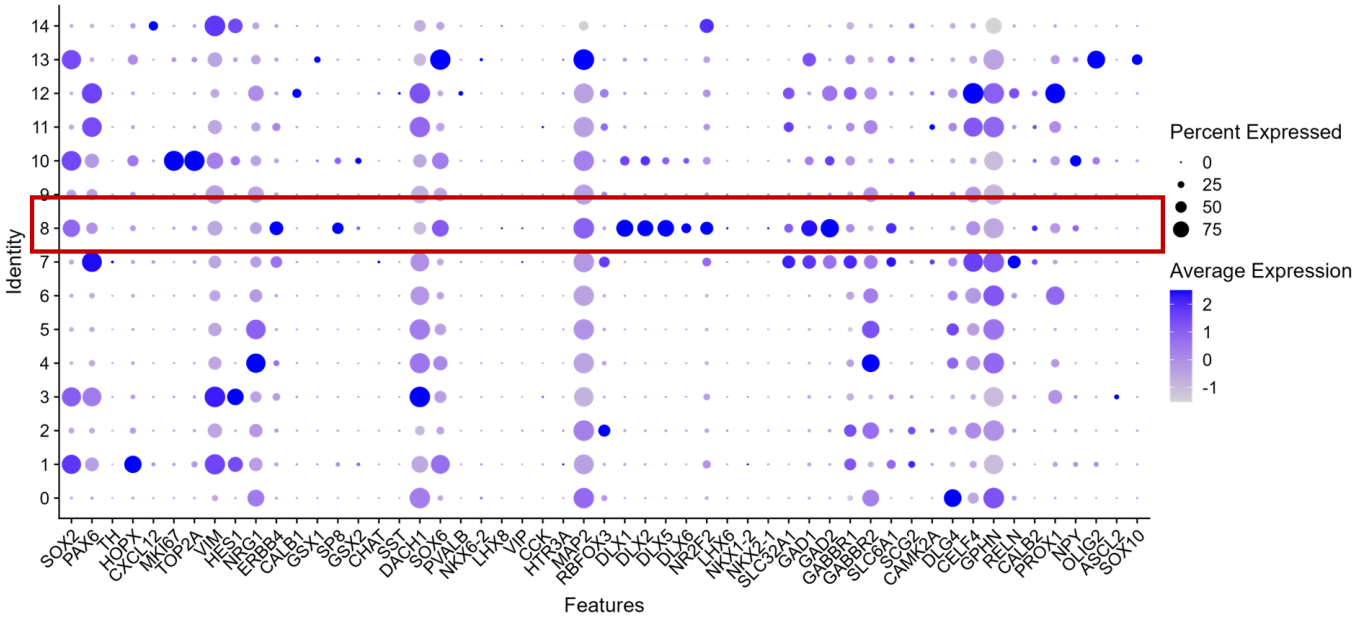
Supplementary Figure 9: Pseudotime trajectory analysis predicts disruptions in multiple cell type emergence.

Pseudotime ridge density plots demonstrate cell-type specific changes in inferred trajectory following Asymptotic two-sample Kolmogorov-Smirnov test. Complete statistics are listed in **Table 3.**



Supplementary Figure 10: Pseudotime modules 19 and 37 have high specificity to excitatory mature neurons and inhibitory neuron populations, respectively.

UMAP plot of pseudotime module gene expression in (**A**) modules enriched in excitatory mature neurons and (**B**) inhibitory neuron populations.



Supplementary Figure 11: UNO inhibitory neurons express canonical progenitor markers.

Dot plot of canonical inhibitory neuron markers with highlighted inhibitory neuron cluster 8.

4 **CHAPTER 4:** Abnormal developmental trajectories in guided 15q13.3del dorsal and ventral forebrain organoids converge on human-specific transcriptional dysregulation and abnormal circuitry.

4.1 **15q13.3del dorsal forebrain organoids have broad transcriptional dysregulation and unique enrichment for pathways of DNA repair.**

The 15q13.3 microdeletion syndrome is a highly heterogeneous disorder, and like many NDDs requires multiple patient backgrounds to get a representative scope of the pathobiology. Bulk RNA Sequencing (bulk RNA Seq) offers a more cost-effective alternative to scRNA Seq and enables the inclusion of multiple patient samples across time. We chose to perform bulk RNA Seq on dorsal and ventral organoids derived from our entire 15q13.3del cohort, and at two timepoints: day 50 (D50) and day 100 (D100) (**Fig. 19**), to capture the timepoint for assembloid profiling and a period of corticogenesis (D50, see methods) as well as a more mature, cellularly diverse, and electrically active tissue types (D100).[139] As the tissue was prepared from each cohort, we noticed an increase in size amongst the ventral organoids by D50 amongst multiple 15q13.3del families that was not present in the dorsal organoid, suggesting that the UNO phenotypes may be present in a tissue-specific manner (**Fig. 20**).

In total, 66 samples were processed for each tissue type. Inclusion criteria for samples included the following: RIN scores > 8, RNA concentrations meeting the required library preparations (5- ng/µL minimum), zero adapter contamination, read depth > 25 million

reads/sample, and Phred scores> 35 (QC in **Supplementary Figures 10-11**). In total, 54/66 samples passed for the ventral bulk seq, require exclusion of family 4 due to low RNA yield and RIN values. The resulting ~16 000 genes were then processed for DGE analysis with DESeq2 (**Fig. 20**). Once we were confident in the differentiation of our organoids, we probed the differentially expressed gene (DEG) list within each dataset for downregulation of the 15q13.3 genes *KLF13, FAN1, OTUD7A*, and *CHRNA7*, (**Supplementary Figures 12-13, 16-17**), providing validation for the heterozygosity of the 15q13.3 locus in our patient lines.

Next, we assessed the DEGs more broadly to examine the affected biological pathways. Amongst the dorsal forebrain dataset, 1 534 DEGs were identified at the day 50 timepoint (644 upregulated, 890 downregulated), and 3 694 DEGs at the day 100 timepoint (1 761 upregulated, 1 933 downregulated) with adjusted p-value (pAdj) to account for false discovery rate (FDR/q-value < 0.05, top 25 DEGs reported in **Tables 4-5**). The gene sets were sorted by fold change and the top 100 up- and downregulated genes were used for GO term analysis using ShinyGO (V0.8).[155]

We first examined the upregulated gene list, which yielded very few GO enrichment pathways; amongst the entire gene list, only a small fraction of genes (as few as 7-8) were significantly enriched for molecular function pathways regulated by oxidoreductases (data not shown), whereas all other pathway databases (biological processes, cellular component, KEGG) yielded no significant enrichment. It was immediately apparent that the downregulated genes, however, had a strong enrichment for transcription factors; amongst the top 25 downregulated genes (**Table 4**), 48% were transcription factors, specifically zing-finger proteins. We examined the implicated

pathways in the top 100 downregulated list, and found a strong enrichment in transcriptional regulation, specifically by RNA polymerase II (**Fig. 22B**). In addition, more forebrain-specific pathways such as neurogenesis and neuron differentiation had > 2-fold enrichment from the same dataset.

We profiled putative neurogenesis by targeting the $SOX2^+$ rosette structures that, similar to UNOs, are present in dorsal forebrain GNOs and produce newborn $TBR1^+$ neurons that later migrate outside of the rosette structure. We chose a single family and grew organoids to the day 50 timepoint and quantified the proportion of $TBR1^+$ cells within the rosette structures (**Fig. 23**). We found a significant reduction in $TBR1^+$ cells, indicating possible disruptions to neuron generation or differentiation.

The mature Day 100 dorsal organoid dataset yielded similar results as the D50 dataset, sharing half of the DEGs in the respective up and downregulated gene lists. GO term analysis of the top 100 upregulated genes produced minimal enrichment in all database categories, whereas the top 100 downregulated genes produced enrichment for transcriptional regulation, neurogenesis, and neuron differentiation, in a nearly identical fashion to the D50 dataset. We performed gene set enrichment analysis (GSEA) to examine the gene lists more broadly (beyond the top 100 DEGs), and produced a ranked gene list that incorporated both the significance and fold change by applying the -log10(pAdj)xSIGN(Fold change) function.[169] The ranked list was used for GSEA, which assesses the ranked value and position of a gene to calculate an enrichment score (ES) related to a pre-defined list of biological pathways.[143] Amongst the upregulated genes that were ranked at the top of the list, we found, to our surprise, significant enrichment (FDR<0.05) in multiple chromosomal and DNA-associated pathways for genes ranked at

the top of our list (**Fig. 22D**). These pathways included the *Activation of ATR in response to replication stress* (Day 50 dataset ES = 0.66, FDR <0.05), *Replication fork processing* (Day 100 dataset, ES = 0.71) and *Diseases of DNA repair* (Day 100 dataset, ES = 0.61).

Due to the similarities of the Day 50 and Day 100 datasets, we combined the raw data into a single data matrix and repeated the DESeq2 analysis with the inclusion of "Timepoint" as a covariate. We found several similar DEGs and pathways (including neurogenesis and transcriptional regulation, **Fig. 24**. Of note, when we plotted the distribution of the top 75 downregulated genes on the human genome, we found significant enrichment at four distinct loci on chromosomes X and 19 (**Fig. 24B, FDR <0.001**). Due to the sex differences in patient lines, the X enriched genes were filtered out, leaving 11 genes on the long arm of chromosome 19. We noticed that 10/11 of these genes encoded zinc finger (ZNF) proteins, including the human-specific transcription factor and regulator of early neocortical expansion, *ZNF558* (**Fig. 24C**). We questioned if this chromosomal enrichment could be due to regulatory activity of the 15q13.3 gene and transcription factor, *KLF13*,[170] and searched for evidence of KLF13 binding sites at this region. We accessed the Encyclopedia of DNA Elements (ENCODE) database for publicly available chromatin immunoprecipitation sequencing (ChIP Seq) bed files of KLF13 ChIP Seq data. We found two datasets: one using a 3xFLAG-tagged KLF13 in HepG2 cells (a hepatoblastoma cell line used to study the liver) and a GFP-tagged KLF13 dataset in K562 (cancer cell line). We accessed bed files from each dataset, set the visualization to fold change over control, and after searching for the *ZNF558* gene found consistent enrichment at its promoter site between both datasets (**Fig. 25)**, suggesting putative binding of KLF13. We also probed for the only known

target of *ZNF558* in NPCs, the mitophagy gene *SPATA18*. Given that ZNF558 represses this SPATA18, its upregulation suggests that that ZNF558 is reduced on a functional level.

We were unsuccessful in characterizing the remaining ZNF proteins, however. We plotted the 50+ ZNF genes downregulated in the dorsal gene set using STRING (Search Tool for the Retrieval of Interacting Genes), a database of known protein-protein interactions, to see if the genes converged on any known biological pathway (**Supplementary Fig. 15**). The only enriched pathway was for generic transcriptional activity, suggesting that the downstream targets to these ZNFs are presently uncharacterized. This arm of the project has therefore identified novel and human-specific disruptions of global and ZNF-enriched transcription in a forebrain model of the 15q13.3 microdeletion syndrome.

4.2 Migratory and cell adhesion pathways are disrupted in 15q13.3del ventral forebrain organoids.

We questioned if inhibitory neurons would have similar impairments in global transcription and neurogenesis, given the abnormal developmental trajectory of the inhibitory neurons in UNOs from our previous project. We repeated the bulk RNA Seq experiment using ventral forebrain GNO tissue derived from the same cohort. As the organoids developed, we observed an increase in size amongst multiple 15q13.3 patient-derived ventral GNOs at the D50 timepoint (**Fig. 20**), somewhat reminiscent of the D40 15q13.3 UNO growth abnormalities. After processing the raw data for 66 total samples, a

PCA produced similar clustering of the raw data by timepoint and familial (**Fig. 27**). DEGs were identified using DESeq2, yielding 644 upregulated and 890 downregulated at the D50 timepoint, and 1761 upregulated, 1933 downregulated at the D100 timepoint. Importantly, 15q13.3 genes *KLF13, FAN1, OTUD7A*, and *CHRNA7* were all significantly downregulated at each timepoint (**Supplementary Figures 15-16**), providing validation for the heterozygosity of the 15q13.3 locus in our patient lines amongst multiple tissue types.

The upregulated genes at the D50 timepoint represents the smallest DEG list within the dataset. We performed GO enrichment analysis on the top 100 genes sorted by fold change and found enrichment in KEGG pathways relating to synaptic signaling such as cyclic adenosine monophosphate (cAMP) signaling (relative to all neural cell types), cholinergic (inhibitory neurons) and glutamatergic (excitatory neurons) synapses, and known regulating pathways of neuron migration, Rap1 and chemokine signaling.[171,172] We screened the corresponding downregulated genes for transcription factors to compare to the dorsal organoid dataset, and found that amongst the top 25 downregulated genes, 28% represented transcription factors, with over half comprised of zinc finger (ZNF) proteins. This proportion of transcription factors is roughly half of what was reported in the dorsal organoid datasets, and so we questioned the function of the remaining 82% of genes. We performed GO analysis on the top 100 downregulated genes (**Fig. 27**), which yielded enrichment for multiple biological pathways including cell adhesion (with the highest gene count of 27), neurogenesis, and cell migration, the latter including genes such as *S100A11* and Sonic Hedgehog protein (*SHH*). We probed for the predominant cell type produced in ventral GNOs at this timepoint, somatostatin (SST) interneurons to

determine whether deficits in neurogenesis are associated with the observed reduced neural populations.[173] SST transcripts were indeed amongst the top 200 downregulated genes in this dataset, suggesting reductions in this cell type. To see if reduced transcripts translated to protein, we measured SST fluorescent intensity from three 15q13.3 families at the D50 timepoint. In doing so we quantified a reduction in both SST protein expression (when normalized to MAP2 expression) as well as the number of SST^+ nuclei in this tissue type (**Fig. 27**).

To examine how disruptions in neuron generation may affect ventral forebrain tissue at a later timepoint, we performed the same GO analysis on the up and downregulated gene sets for the Day 100 tissue. We found amongst the upregulated genes enrichment for neuroactive ligand receptor interactions and other synaptic signaling pathways (**Fig. 28**), which was unexpected given the reduction of SST protein at the D50 timepoint. Similarly, the downregulated genes were also enriched for synaptic signaling pathways, but in a G couple protein receptor (GPCR)-specific manner. Of the non-neuronal pathways enriched amongst the downregulated genes, transcriptional regulation was amongst the highest by fold change, suggesting convergence of a disrupted transcriptional pathway.

One way that the ventral datasets differed from the dorsal organoids was the directionality in the enriched GSEA set, showing enrichment exclusively from the lowest ranked (downregulated) genes (**Fig. 27-28**). Amongst the D50 dataset, there was strong enrichment for notch signaling (ES = -0.81), a pathway required for neuron maturation, migration, and survival, as well as the well-established axon guidance and neuron migration signaling pathway, Slit/Robo (ES = -0.63).[174] Both the D50 and D100 pathway

showed enrichment for integrin binding, which can act as adhesion molecules and therefore supports the previous adhesion protein GO enrichment in the D50 dataset. To assess shared genes between the two timepoints further, a single matrix file was uploaded and processed through DESeq2 to achieve a total of 4713 DEGs. We plotted the top 100 downregulated genes on the human chromosome and found significant enrichment at three loci: one on chromosome X, and three on chromosome 19 (**Fig. 29**).

4.3 Summary

Transcriptional dysregulation is a clear and persistent phenotype amongst all 15q13.3del forebrain organoid tissues in this study, although the consequence of this dysregulation is unclear. By contrast, GO analyses revealed ventral organoid-specific enrichment to cell adhesion and migration deficits at multiple timepoints, with additional enrichment for G-PCR signaling impairments at the D100 timepoint, suggesting a functional consequence to deficits that occur at early (D50) timepoints. To examine the convergent and lineage-specific changes of dorsal and ventral tissues in a more biologically relevant model, we combined the two organoids into assembloids and examined circuit dynamics using light sheet microscopy.

4.4 Comparison of common and divergent ventral forebrain genes across time shows the presence of unique migratory-associated genes at a later timepoint.

In order to identify temporally relevant changes in ventral forebrain gene expression, we examined the genes that were unique to each timepoint in the ventral organoid dataset (**Fig 30**). The gene sets shared considerable overlap, with over two thirds of the D50 genes intersecting with the D100 timepoint, and roughly half of the D100 genes overlapping the D50 genes. From the D50 unique gene set, we were surprised to find

translation as a highly enriched biological pathway. Conversely, both the shared and unique D100 genes showed impairments in cell migration and cell adhesion, suggesting retention and exacerbation of these pathways over time, possibly indicating more severe disruptions to the distribution and physical connections between neurons of a circuit. To further characterize changes in gene expression across time, we examined genes that had inverse expression patterns between the two timepoints (going up to down or down to up). We identified 33 genes that were originally upregulated at D50 and then downregulated at D100, and following GO enrichment analysis and STRING visualization, found a small subset to be associated with axonal myelination, a process crucial for action potential propagation (**Fig. 31**).

4.5 15q13.3 forebrain organoids converge on dysregulated cell adhesion pathways

Lastly, we wanted to determine if there was any convergence of DEGs or pathways present in the dorsal and ventral datasets at both timepoints. We compared the shared DEGs between all four datasets and identified 223 genes (**Fig. 32**). We used the entire gene set for GO enrichment analysis, and found enrichment exclusively in cell adhesion related pathways, suggesting that this abnormality persists across multiple tissue types. Furthermore, when plotting the chromosomal distributions of these DEGs, there was once again spatial enrichment at four loci: one each on chromosomes X and 19, with two additional enrichment loci on chromosome 5. When examining the Chr19 genes further, the enrichment for ZNFs persisted, including 10/13 of the genes in these regions (whereas the remaining three were imprinting genes (*PEG3*, *ZIM2*) and a myosin superfamily

gene, *MYH14)*. Interestingly, the chromosome 5 enrichment represented a cluster of 11 protocadherin genes at position 5q31.3.

4.6 Inhibitory neurons in the 15q13.3 microdeletion fail to migrate in multiple genotype-lineage dorsal-ventral assembloid combinations.

The bulk RNA sequencing data heavily suggested impairments in both dorsal and ventral tissues harboring the 15q13.3 microdeletion. Despite clear enrichment for adhesions proteins in both tissue types, as well as tissue-specific disruptions in interneuron migration, the analyses are from an isolated organoid system that lacks circuit and synaptic input between excitatory and inhibitory neurons. To predict how interneuron dynamics may be affected in isolation as well as in environments that contain excitatory neurons, we examined the overlapping DEGs between D50/D100 dorsal and the interneuron population from the D120 UNO scRNA Seq (**Fig. 33**). Of the biological pathway GO terms, one of the top pathways by fold change was neuron migration, including several well-known interneuron migration genes including *DAB1*, *DCX*, and the human specific gene, *SRGAP2*.[175,176] There was also considerable enrichment for morphological processes such as cellular projection and axonal pathways, suggesting that the 15q13.3del inhibitory neurons may have impairments in migration as well as axonal/neurite outgrowth.

Imbalances in inhibitory and excitatory neurons are characteristic of NDD models, including ASD and epilepsy,[177–179] and may result from deficits in neuronal maturation or a failure of the migrating neurons to reach their final destination in the developing cortex. One neural population that is particularly vulnerable to migratory deficits are GABAergic

interneurons, which perform saltatory migrations across the subpallium of the developing telencephalon, traversing tangentially along extensive paths while responding to chemical cues from the extracellular environment.[180,181] This prenatal phenomenon is rarely captured in the clinic, but can be modeled faithfully by forebrain assembloids, which are made by fusing guided forebrain organoids of dorsal and ventral lineage. The assembloid model captures the microenvironmental cues necessary for the migration of GABAergic interneuron across the dorsal-ventral axis to produce functionally relevant circuits reminiscent of the developing forebrain.[135,140,182]

To examine interneuron migration in a 15q13.3 microdeletion background, we sparsely labeled three 15q13.3del family lines by infecting iPSCs with lentiviral constructs harbouring hSYN-rTTA and either pTET-CMV-eGFP or pTET-CMV-mCherry (**Supplementary Fig. 17**) The infected lines are inducible with the addition of doxycycline and importantly, are under a pan-neuronal-specific promoter. Induction can occur in a subset of samples and allows us to probe the remaining assembloids from the same batch for validation experiments. After infection, the hiPSC lines were differentiated into dorsal and ventral forebrain assembloids using a combination of genotypes and forebrain lineages to evaluate how different genotype microenvironments in the dorsal organoid would influence the migration of the ventral interneurons. The individual organoids were then merged into assembloids on day 25 of differentiation, which is the earliest time that the organoids have identical patterning media. The resulting four combination of assembloids (**Fig. 33**) were left to innervate for a month, doxycycline-induced for five days to reach maximum fluorophore expression (assessed

visually), and then subjected to optical tissue clearing (CUBIC protocol, see **Methods**) followed by light sheet microscopy.

Using a combined dataset from three unique families, we found a significant decrease in interneuron migration amongst multiple D-V assembloid combinations: the largest decrease of 13% occurred between the WT-WT and HET-HET combinations ($p<0.001$), which was also present when comparing the HET-HET samples to the HET-WT combination (12%, $p<0.01$). Interestingly, there were also reductions in the migration of the WT-HET compared to the WT-WT combination (9.9% decrease, $p<0.05$) (**Fig. 34C**). To assess the putative migration path of the migrating neurons, we measured the distance of the migrating cell to the closest surface of the ventral organoid and found no difference between any of the combinations. We then assessed the cellular morphology of the migrating interneurons to see if there was a maturation impact following migration and measured the most prominent neurites of the migrating cells using the NeuriteTracer package in IMARIS. We found a significant reduction in neurite length in all but the WT-WT combinations, with the largest reduction in the HET-HET combinations (**Fig. 35**), suggesting impairments in neurite outgrowth or maturation.

4.7 Summary

This study is the first characterization of human forebrain circuitry in a 15q13.3del background. Using guided dorsal and ventral forebrain assembloids from three unique families, we were able to characterize impairments in interneuron migration and morphology using an optimized tissue clearing technique. We also added nuance to the assembloid system through the novel introduction of genotype- and lineage-combinations to parse out cell autonomous *vs.* non-autonomous contributions to interneuron circuitry.

4.8 Chapter 4 Figures

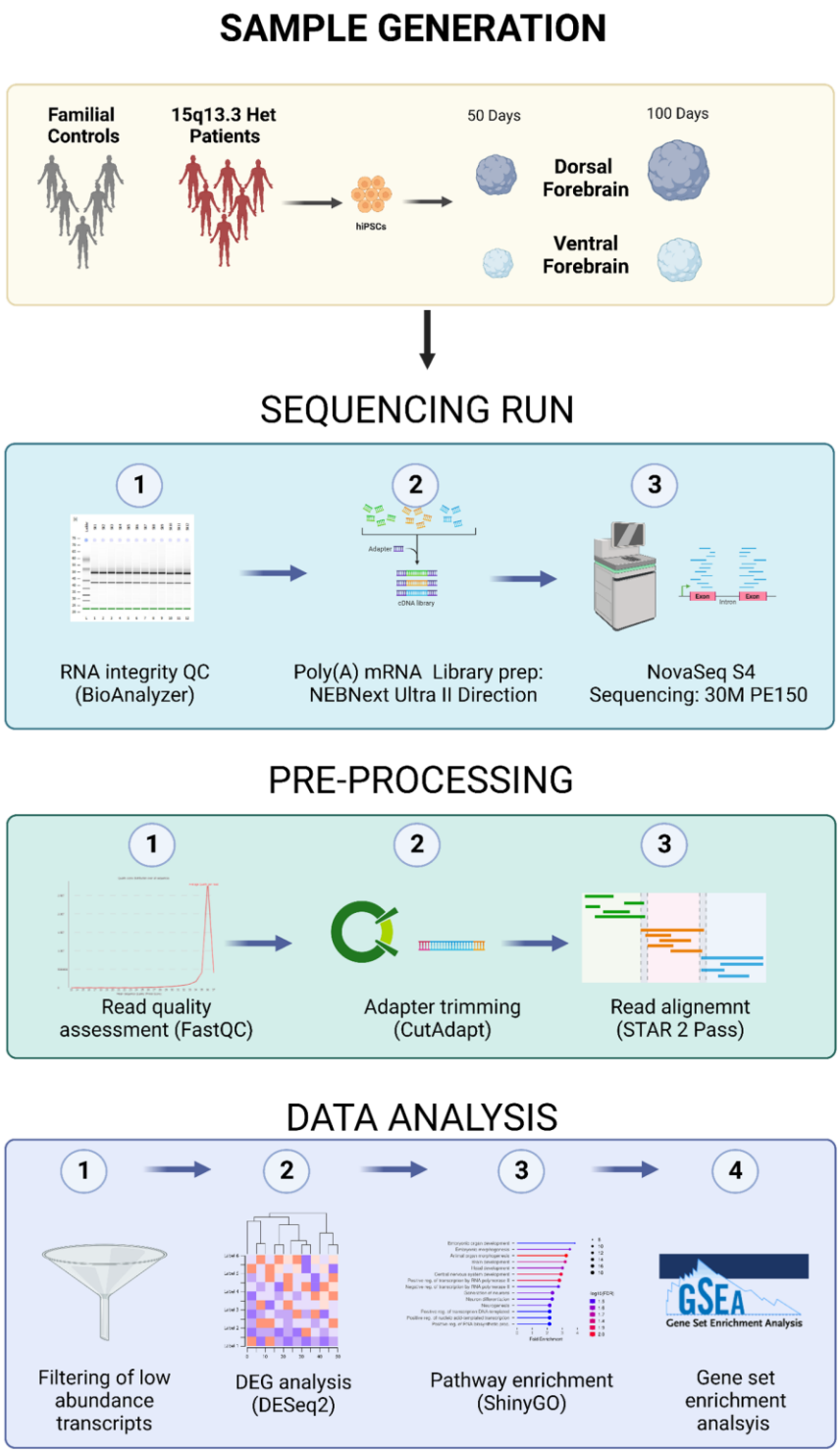
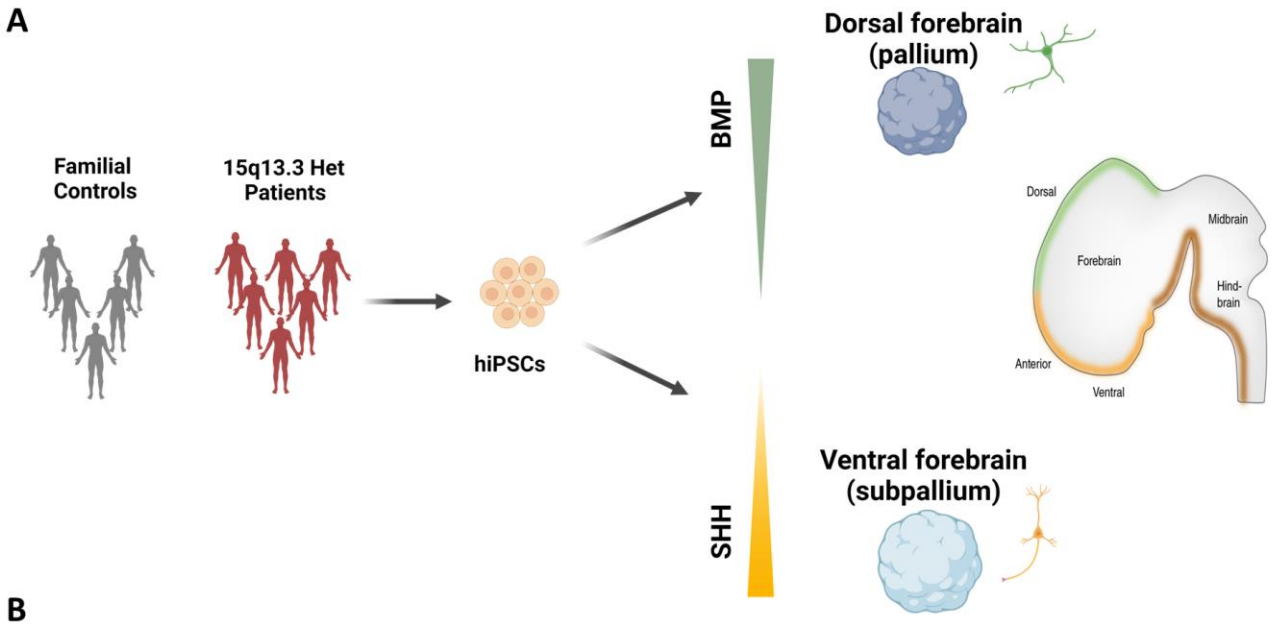


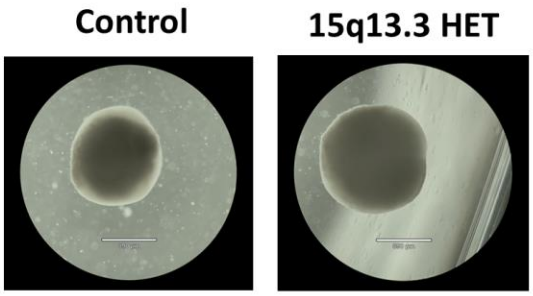
Figure 19: Pipeline schematic for processing and analyzing bulk RNA sequencing from forebrain guided neural organoids.

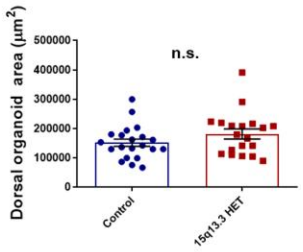
Bulk RNA sequencing (bulk RNA Seq) of entire 15q13.3del cohort includes tissue dissociation at two timepoints, followed by cDNA library generation and mRNA sequencing. Raw read quality is assessed and then trimmed of adapter content, following alignment to the human genome using an established read processing pipeline from Dr. Kristen Hope's lab. Downstream data analysis includes filtering of lowly abundant transcripts, differential gene expression (DGE) analysis, pathway enrichment, and gene set enrichment analysis.



B

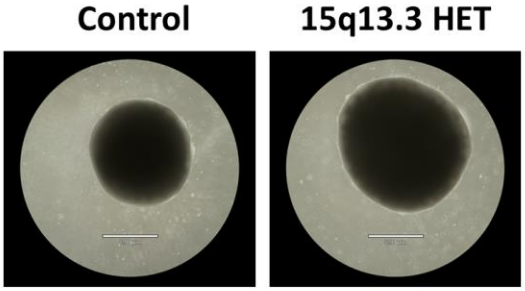
DORSAL FOREBRAIN ORGANOIDS





C

VENTRAL FOREBRAIN ORGANOID



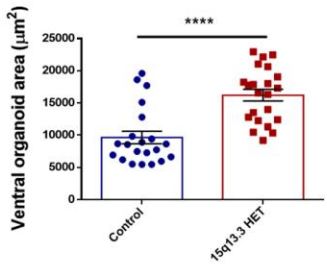


Figure 20: 15q13.3del ventral but not dorsal forebrain organoids recapitulate size increase seen in unguided neural organoids.

(A) Schematic of forebrain organoid generation from human induced pluripotent stem cells (hiPSCs). BMP (bone morphogenic protein) SHH (Sonic hedgehog). Adapted from Miura and Pașca, 2019. *Nature Biotechnology.*
(B) Day 55 15q13.3del dorsal forebrain organoid area (N= 21-24 and HET (N=22-27) from two families. Data represent mean ±SEM in organoid slices n.s = nonsignificant.
(C) Day 55 15q13.3del ventral forebrain organoids are larger at the day 55 timepoints (N= 21-24 and HET (N=22-27) from two families. Data represent mean ±SEM , **** $p < 0.001$, data analyzed by a two-sided student's t-test,

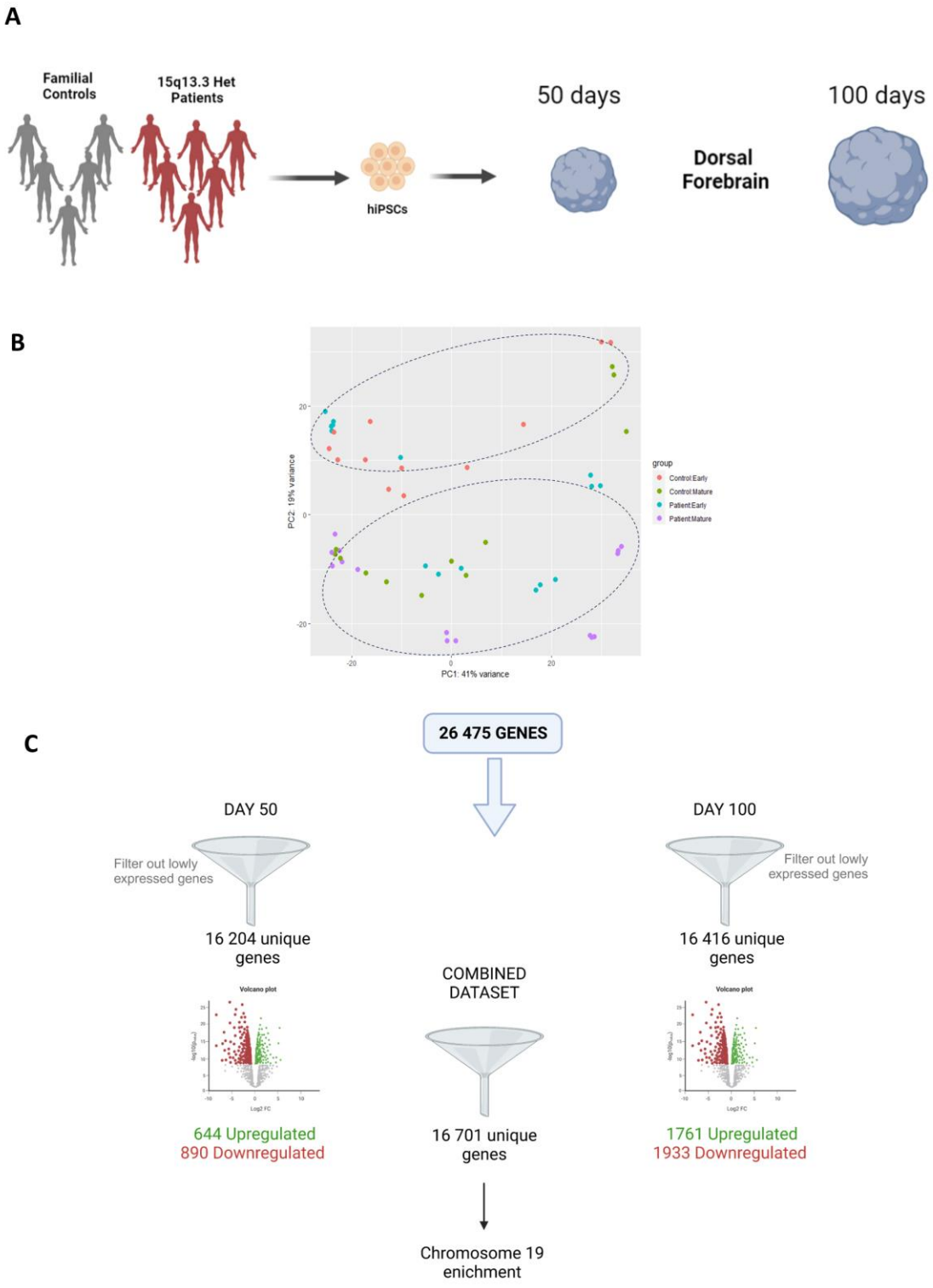

A
Familial Controls
15q13.3 Het Patients
hiPSCs
50 days
Dorsal Forebrain
100 days
B
PC2: 19% variance
PC1: 41% variance
group
Control:Early
Control:Mature
Patient:Early
Patient:Mature
C
26 475 GENES
DAY 50
Filter out lowly expressed genes
16 204 unique genes
Volcano plot
644 Upregulated
890 Downregulated
COMBINED DATASET
16 701 unique genes
Chromosome 19 enichment
DAY 100
Filter out lowly expressed genes
16 416 unique genes
Volcano plot
1761 Upregulated
1933 Downregulated

Figure 21: Summary of guided dorsal forebrain organoid bulk RNA sequencing results from the entire 15q13.3del cohort.

(**A**) Schematic of bulk seq experiment at day 50 and day 100 timepoints.
(**B)** Principal component analysis (PCA) with highlighted timepoints and genotype. Dashed line represents spatial separation of samples by timepoint.
(C) Gene counts pre- and post-filtering for lowly expressed transcripts.

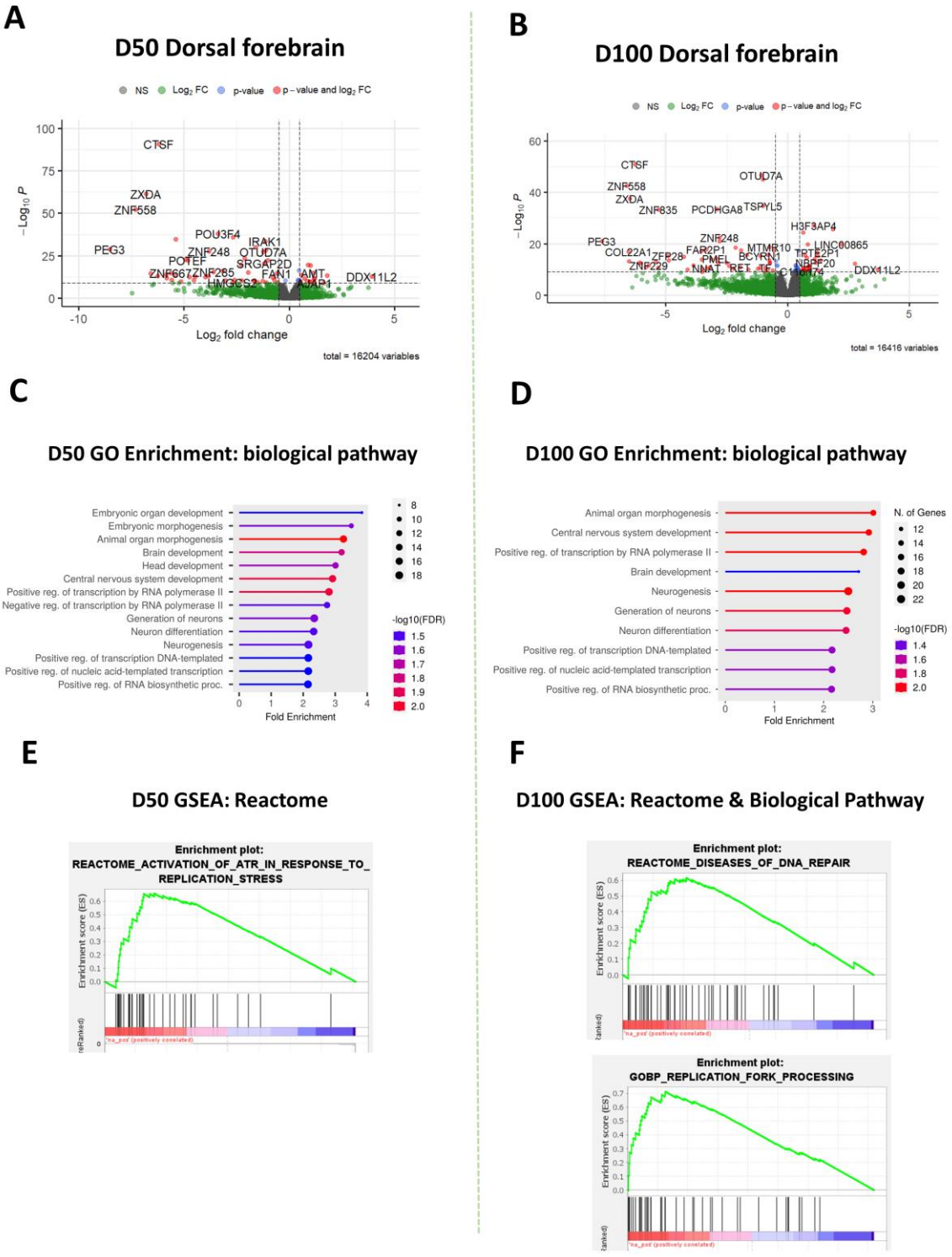

A
D50 Dorsal forebrain
NS
Log2 FC
p-value
p-value and log2 FC
−Log10 P
Log2 fold change
total = 16204 variables
CTSF
ZXDA
ZNF558
POU3F4
PEG3
IRAK1
ZNF248
OTUD7A
POTEF
SRGAP2D
ZNF667
ZNF285
FAN1
AMT
HMGCS2
ADAP1
DDX11L2
B
D100 Dorsal forebrain
total = 16416 variables
C
D50 GO Enrichment: biological pathway
Embryonic organ development
Embryonic morphogenesis
Animal organ morphogenesis
Brain development
Head development
Central nervous system development
Positive reg. of transcription by RNA polymerase II
Negative reg. of transcription by RNA polymerase II
Generation of neurons
Neuron differentiation
Neurogenesis
Positive reg. of transcription DNA-templated
Positive reg. of nucleic acid-templated transcription
Positive reg. of RNA biosynthetic proc.
Fold Enrichment
-log10(FDR)
D
D100 GO Enrichment: biological pathway
N. of Genes
E
D50 GSEA: Reactome
Enrichment plot:
REACTOME_ACTIVATION_OF_ATR_IN_RESPONSE_TO_REPLICATION_STRESS
Enrichment score (ES)
F
D100 GSEA: Reactome & Biological Pathway
Enrichment plot:
REACTOME_DISEASES_OF_DNA_REPAIR
Enrichment plot:
GOBP_REPLICATION_FORK_PROCESSING

Figure 22: 15q13.3 dorsal forebrain organoids show persistent transcriptional deficits and gene signatures for DNA repair.

Volcano plot of DEGs at **(A)** day 50 (644 upregulated, 890 downregulated) and **(B)** day 100 (1761 upregulated, 1933 downregulated). Blue: pAdj<0.01, red: Log2FC > 1. (**C-D**) GO enrichment analysis for biological pathways show enrichment in transcriptional regulation and neurogenesis.
(E-F) Gene set enrichment analysis (GSEA) of day ranked dataset shows alterations in **(E)** *D50 ATR- in response to replication stress* (ES = 0.66, FDR <0.05), and (F) disruptions in DNA repair and enrichment for replication fork processing (BP = biological pathway, ES = 0.71, ES = 0.61, FDR <0.05).

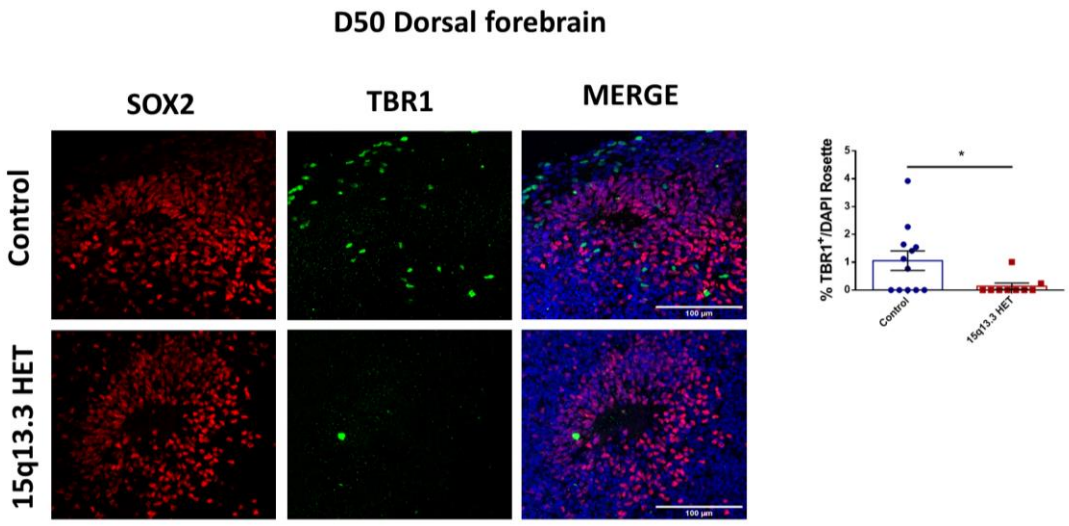


Figure 23: Immature 15q13.3del dorsal forebrain organoids have reduced TBR1$^+$ newborn neurons.

Reduction in TBR1$^+$ newborn neurons in dorsal organoid rosettes (N Control = 11, N 15q13.3del = 13) from 15q13.3del family 3, *p <0.05; student's unpaired t-test.

	baseMean	log2FoldChange	padj
PEG3	3610.648712	-8.516014533	2.36E-26
ZNF208	36.84528711	-7.470165441	0.016600324
ZNF558	652.7056727	-7.317517954	3.21E-49
ZXDA	285.496561	-6.798460033	2.04E-58
ZNF667.AS1	756.8087481	-6.578518485	1.39E-12
ZNF300P1	10.61232334	-6.511393264	3.94E-06
LDOC1	2747.142848	-6.435870732	9.95E-07
ZFP28	975.4877638	-6.257274146	3.28E-11
CTSF	606.4764821	-6.206274698	2.18E-87
ZNF471	376.4355206	-5.853063986	6.27E-12
ZNF229	525.7069278	-5.81984268	5.00E-10
ZNF667	509.1718689	-5.639166807	1.03E-12
ZNF728	29.09074034	-5.585602728	7.44E-09
FAM135B	66.87225757	-5.526964835	3.18E-07
ZNF736	453.743189	-5.420873993	1.13E-11
EN1	372.3782509	-5.399793232	0.001271871
ZNF835	240.7103476	-5.394037208	4.22E-32
PRDM13	864.1743009	-5.262596632	0.001156584
PCDHGA3	462.5804729	-5.239306122	2.47E-07
FNDC7	13.79038139	-5.155664453	3.73E-08
FOLH1	12.851132	-5.012683404	8.78E-05
SKOR2	3594.423387	-4.980089778	0.005347248
FAR2P1	36.96228983	-4.946596953	2.11E-20
POTEF	35.53171792	-4.823799142	7.15E-20
COL22A1	129.252971	-4.771346296	4.48E-07

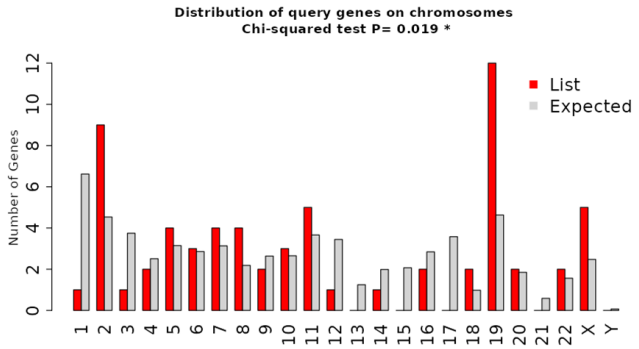
Table 4: Top 25 downregulated genes in dorsal forebrain bulk RNA sequencing (Day 100 Timepoint).

List of top 25 downregulated genes in day 100 dorsal forebrain dataset. Genes filtered pAdj < 0.05 and ranked by log2FoldChange.

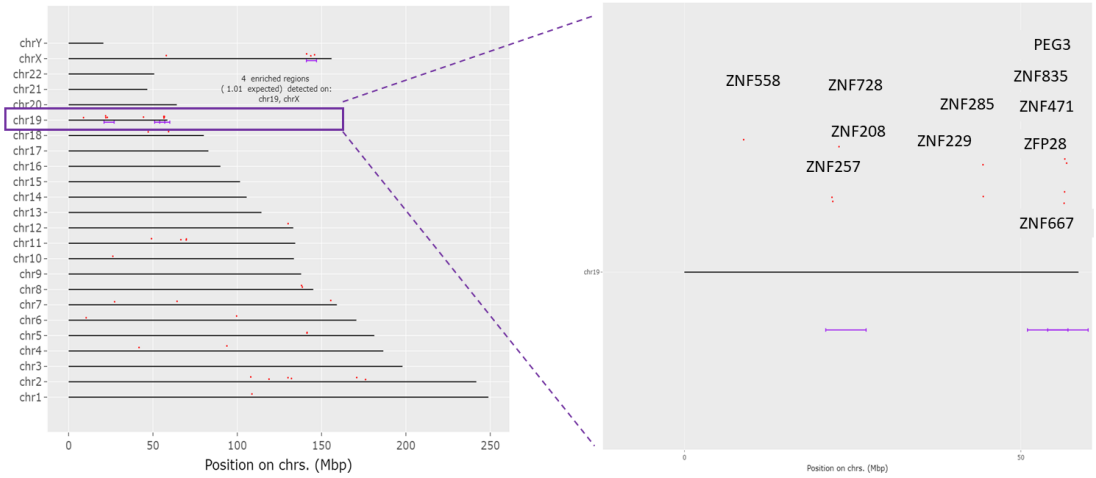
	baseMean	log2FoldChange	pAdj
PEG3	1948.743	-7.61327	9.64E-19
ZNF558	516.8062	-6.60405	8.28E-40
COL22A1	254.7218	-6.53585	7.54E-15
FAM135B	138.0919	-6.51476	2.13E-11
ZXDA	344.1521	-6.47769	5.79E-35
ZNF208	35.90901	-6.37236	0.008225
CTSF	890.7547	-6.27425	8.79E-48
TFAP2B	1724.643	-6.20064	3.86E-06
NAP1L6	9.027446	-6.08388	7.4E-11
ZNF667.AS1	755.4572	-6.01773	1.33E-10
SKOR2	1369.439	-5.97662	4.88E-05
LDOC1	2774.015	-5.94154	3.82E-07
ZNF736	337.5212	-5.81799	3.8E-09
ZNF300P1	10.65236	-5.75262	0.00027
ZNF229	514.0741	-5.70148	9.11E-10
FOLH1	18.19634	-5.43663	3.96E-06
CPLX4	373.1362	-5.42397	6.34E-05
ZNF471	363.4777	-5.40261	4.08E-11
ZNF728	27.99268	-5.3115	1.57E-06
ZNF835	239.3233	-5.30822	1.04E-30
PCDHGA7	267.351	-5.30134	4.41E-09
CECR7	13.63934	-5.28672	2.82E-07
PCDHGA3	686.401	-5.17394	2.1E-05
MED15P9	6.952533	-5.0675	2.61E-06
GAD2	1849.734	-4.9524	3.84E-05

A

D50/100 Combined Dorsal forebrain organoid



B



C

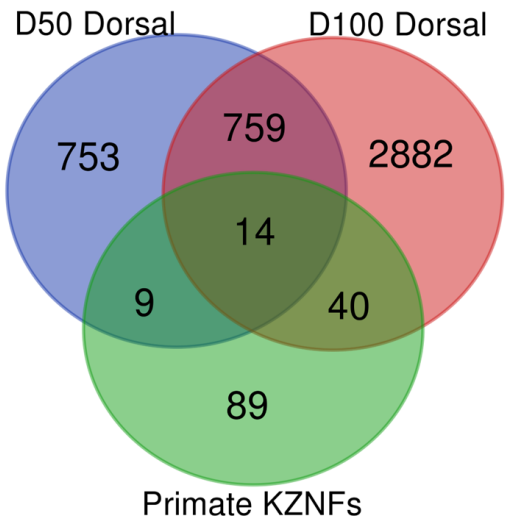


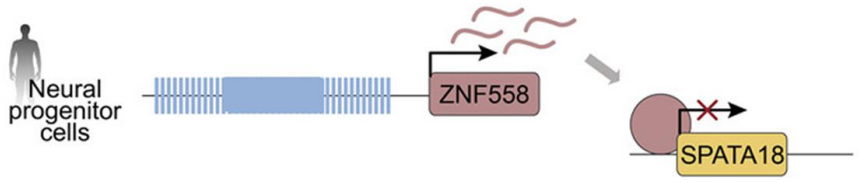
Figure 24: 15q13.3del dorsal forebrain organoids converge on a downregulated gene set of Krüppel-domain zinc finger proteins (KZNFs) on chromosome 19.

(**A-B)**: DEG distribution of combined D50/D100 DeSEQ2 analysis shows enrichment for chromosome 19, Chi squared test, p =0.019.

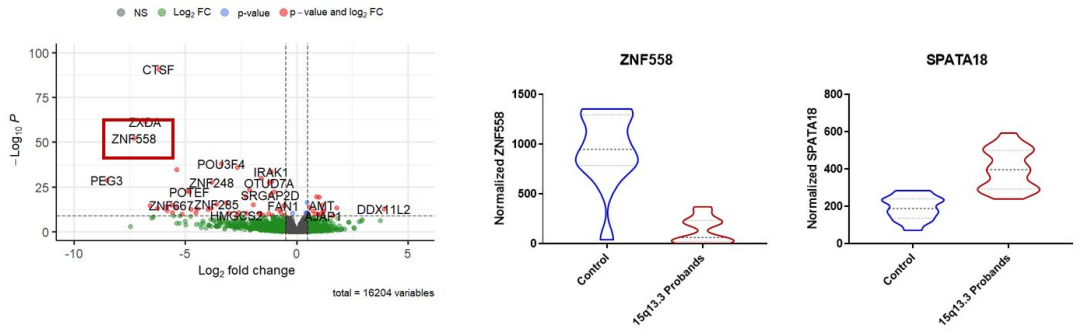
(**B):** Genome mapping reveals cchromosomal 19 enrichment loci containing multiple KZNFs, FDR<0.001.

(**C)**: Putative binding domains of KLF13. Bed tracks K562 and HepG2 cell line ChIP-Seq experiments visualized on IGV (UC San Diego, https://igv.org/app)

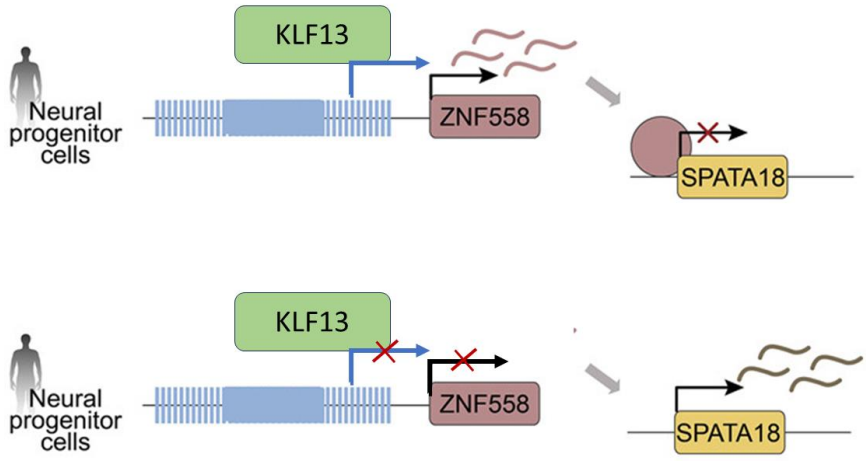
A


Neural progenitor cells
ZNF558
SPATA18

B


NS
Log2 FC
p-value
p-value and log2 FC
CTSF
ZXDA
ZNF558
POU3F4
IRAK1
PEG3
ZNF248
OTUD7A
POTEF
SRGAP2D
ZNF667ZNF285
FAN1
AMT
DDX11L2
HMGCS2
PAP1
-Log10 P
Log2 fold change
total = 16204 variables
ZNF558
Normalized ZNF558
Control
15q13.3 Probands
SPATA18
Normalized SPATA18
Control
15q13.3 Probands

C


KLF13
Neural progenitor cells
ZNF558
SPATA18
KLF13
Neural progenitor cells
ZNF558
SPATA18

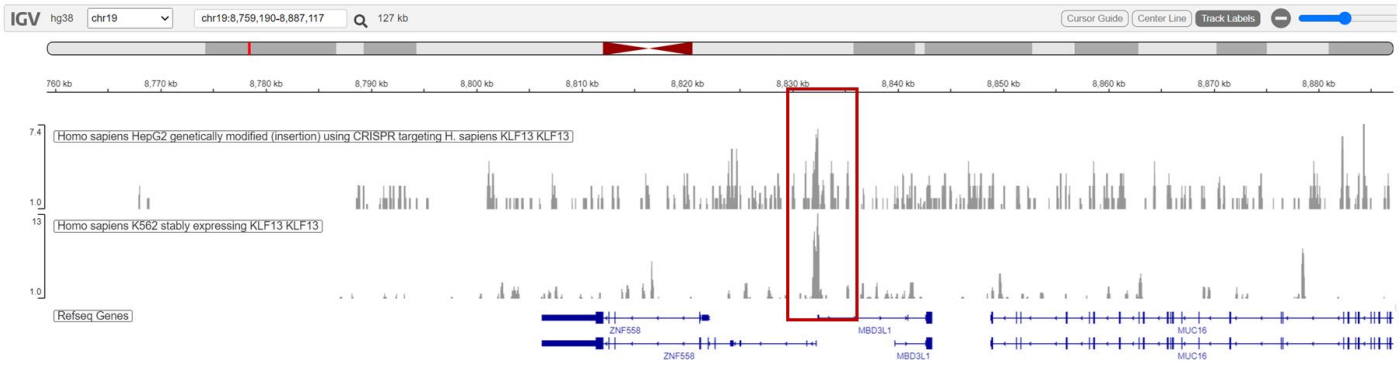
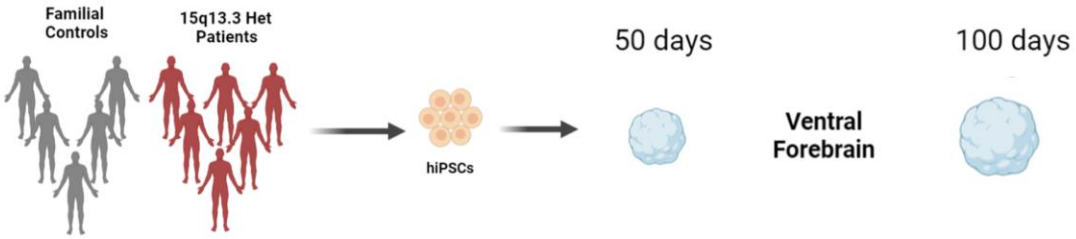

IGV
hg38
chr19
chr19:8,759,190-8,887,117
127 kb
Cursor Guide
Center Line
Track Labels
Homo sapiens HepG2 genetically modified (insertion) using CRISPR targeting H. sapiens KLF13 KLF13
Homo sapiens K562 stably expressing KLF13 KLF13
Refseq Genes
ZNF558
MBD3L1
MUC16

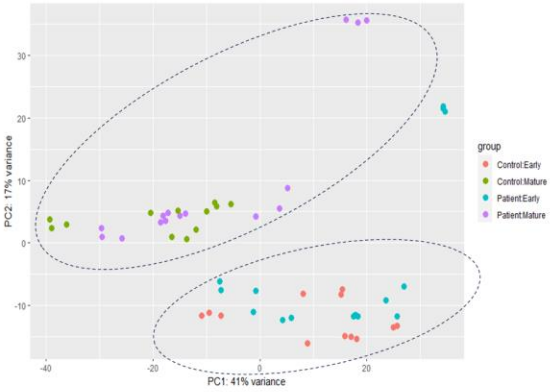
Figure 25: KZNF *ZNF558* is a top downregulated gene in 15q13.3del dorsal forebrain organoids and has putative binding sites for *KLF13*.

(A) Regulatory mechanism of *ZNF558* to *SPATA18* in human NPCs.
(B) Volcano plot and normalized mRNA expression of top downregulated KZNF, ZNF558, and its primary target, SPATA18 in day 100 dorsal organoid dataset.
(C) Proposed regulatory mechanism of *ZNF558* to *SPATA18* in human NPCs.
Images adapted from *Johannson et al., Cell Stem Cell, 2022.*[170]
https://doi.org/10.1016/j.stem.2021.09.008
(D) Putative binding domains of KLF13. Bed tracks K562 and HepG2 cell line ChIP-Seq experiments visualized on IGV (UC San Diego, https://igv.org/app)

A


Familial Controls
15q13.3 Het Patients
hiPSCs
50 days
Ventral Forebrain
100 days

B


PC2: 17% variance
PC1: 41% variance
group
Control:Early
Control:Mature
Patient:Early
Patient:Mature

C

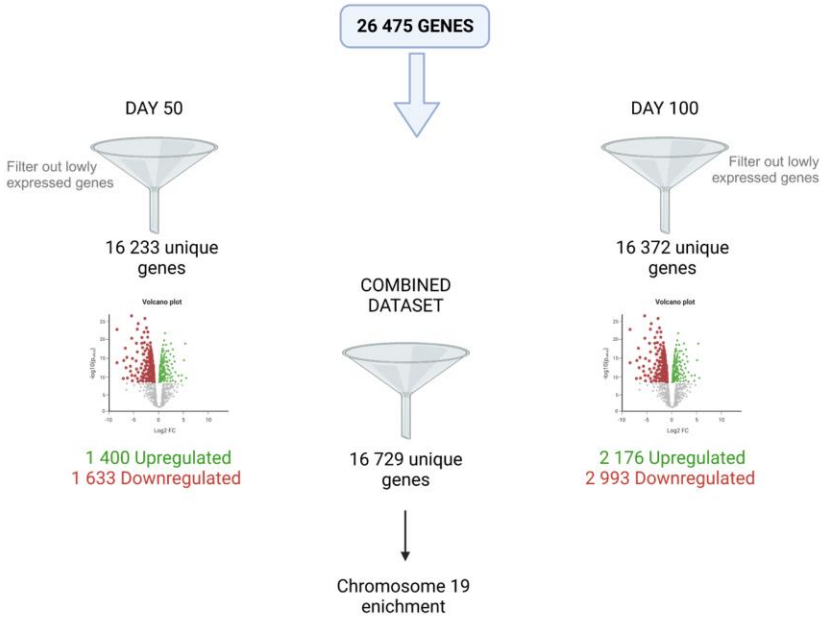

26 475 GENES
DAY 50
Filter out lowly expressed genes
16 233 unique genes
1 400 Upregulated
1 633 Downregulated
COMBINED DATASET
16 729 unique genes
Chromosome 19 enichment
DAY 100
Filter out lowly expressed genes
16 372 unique genes
2 176 Upregulated
2 993 Downregulated

Figure 26: Summary of guided ventral forebrain organoid bulk RNA sequencing results from the entire 15q13.3del cohort.

(**A)** Schematic of bulk seq experiment at day 50 and day 100 timepoints.
(**B**) Principal component analysis (PCA) with highlighted timepoints and genotype. Dashed line represents spatial separation of samples by timepoint.
(C) Gene counts pre- and post-filtering for lowly expressed transcripts.

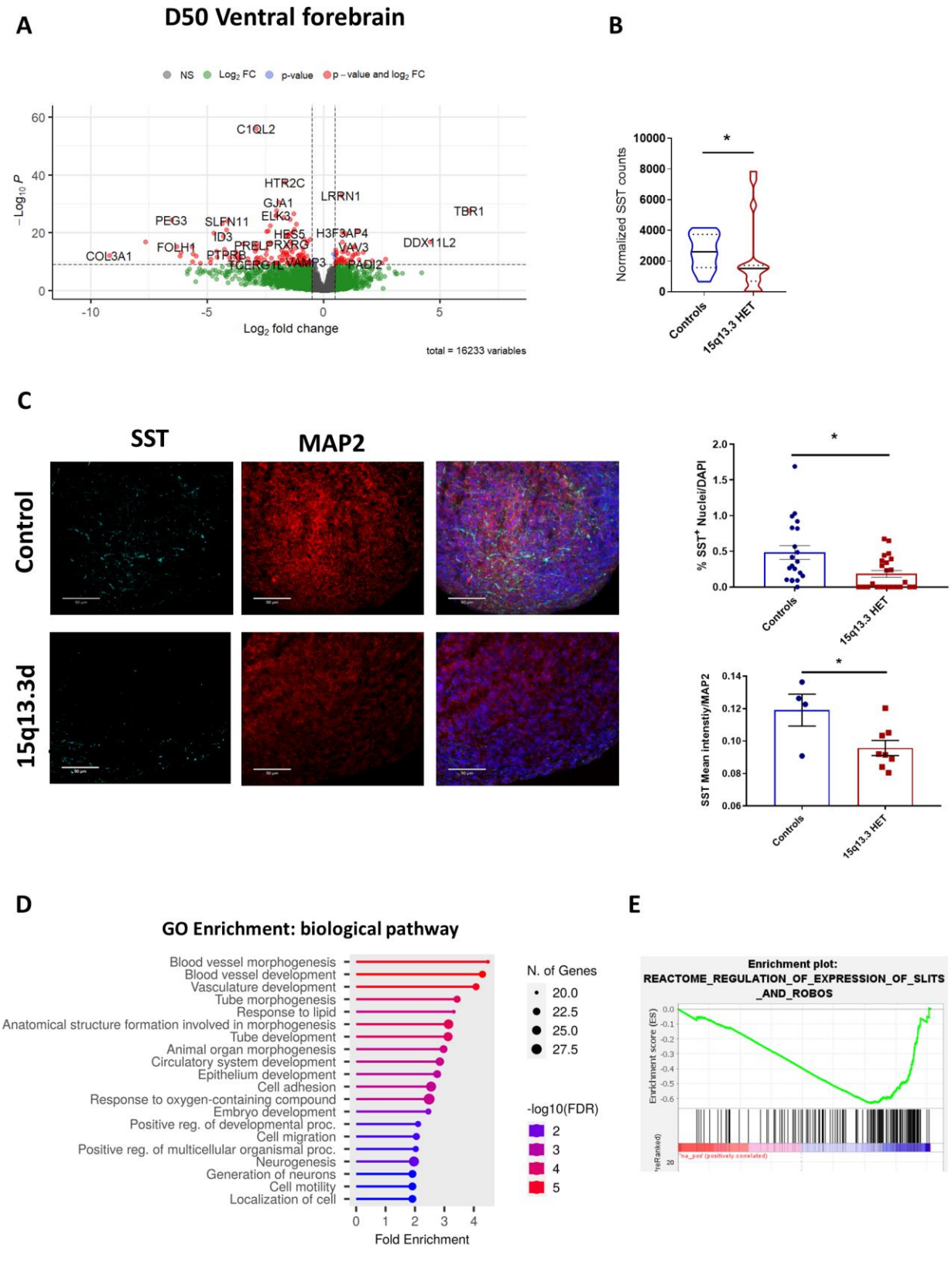

A
D50 Ventral forebrain
NS
Log2 FC
p-value
p – value and log2 FC
C1QL2
HTR2C
GJA1
LRRN1
TBR1
PEG3
SLFN11
ELK3
HES5
H3F3AP4
ID3
PRELP
RXRG
VAV3
DDX11L2
FOLH1
COL3A1
PTPRB
TGFBR1
VAMP3
PADI2
−Log10 P
Log2 fold change
total = 16233 variables
B
Normalized SST counts
Controls
15q13.3 HET
C
SST
MAP2
Control
15q13.3d
% SST+ Nuclei/DAPI
SST Mean intenstiy/MAP2
Controls
15q13.3 HET
D
GO Enrichment: biological pathway
Blood vessel morphogenesis
Blood vessel development
Vasculature development
Tube morphogenesis
Response to lipid
Anatomical structure formation involved in morphogenesis
Tube development
Animal organ morphogenesis
Circulatory system development
Epithelium development
Cell adhesion
Response to oxygen-containing compound
Embryo development
Positive reg. of developmental proc.
Cell migration
Positive reg. of multicellular organismal proc.
Neurogenesis
Generation of neurons
Cell motility
Localization of cell
Fold Enrichment
N. of Genes
20.0
22.5
25.0
27.5
-log10(FDR)
2
3
4
5
E
Enrichment plot:
REACTOME_REGULATION_OF_EXPRESSION_OF_SLITS_AND_ROBOS
Enrichment score (ES)

Figure 27: 15q13.3 immature ventral forebrain organoids show disruptions in migratory signaling pathways and reductions in interneuron marker SST.

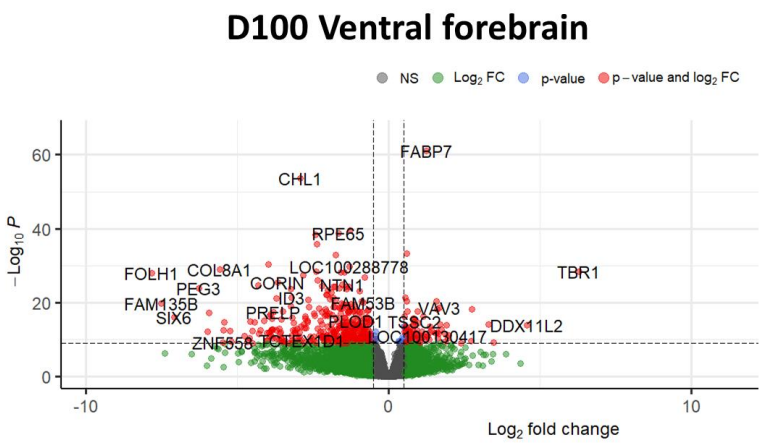
(A) Volcano plot of DEGs at day 50 (644 upregulated and 890 downregulated). Blue: pAdj <0.01, red: Log2FC > 1.
(B) Normalized D50 ventral organoid read count of somatostatin (SST) mRNA. Data represent mean ±SEM of raw SST read counts. * pAdj < 0.05; Wald test.
(C) SST proportion and mean expression is reduced in Day 50 15q13.3del ventral forebrain organoids. (Left) representative 20x images, right quantification of N= 12 WT and N= 12 HET ventral organoids across three families, 3 sections per organoid. Data represent mean ±SEM in organoid slices; *p <0.05; student's unpaired t-test.
(D) GO analysis for biological pathways show enrichment for neurogenesis and neuronal migration pathways at the day 50 timepoint.
(E) GSEA reveals enrichment for axonal pathfinding and migration pathway, Slit/Robo (ES = -0.63)

Table 5: Top 25 downregulated genes in ventral forebrain bulk RNA sequencing (Day 50 Timepoint).

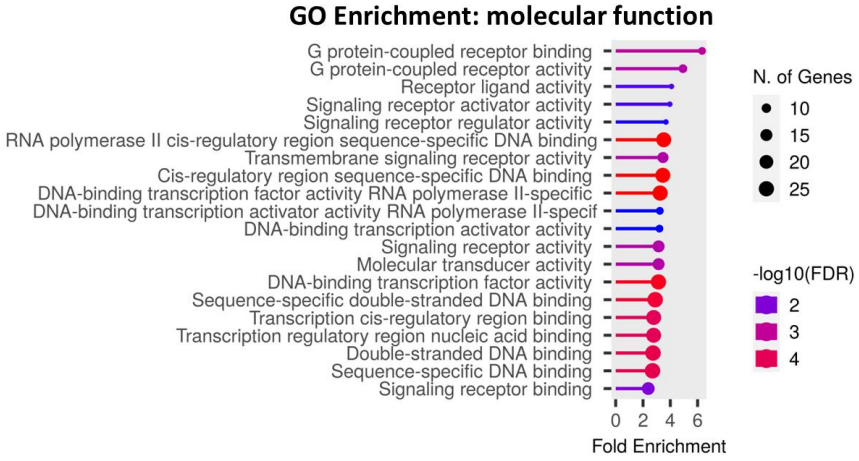
List of top 25 downregulated genes in day 50 ventral forebrain dataset. Genes filtered by pAdj < 0.05 and ranked by log2FoldChange.

	baseMean	log2FoldChange	pAdj
COL3A1	**1179.49**	**-9.20127**	**1.45E-10**
COL22A1	**111.1972**	**-7.65273**	**5.8E-15**
PEG3	**7232.308**	**-6.52512**	**6.11E-22**
FOLH1	**54.63103**	**-6.29906**	**2.03E-13**
TWIST1	**28.07798**	**-6.16304**	**1.7E-10**
PCDHGA7	**211.7828**	**-6.12121**	**1.66E-11**
LUM	**93.25936**	**-5.84693**	**7.56E-06**
ZNF300P1	**39.8875**	**-5.84276**	**5.4E-05**
LDOC1	**2894.513**	**-5.76367**	**2.15E-06**
FAM135B	**164.2121**	**-5.65609**	**2.2E-13**
ZNF558	**525.2187**	**-5.62053**	**1.37E-08**
SIX6	**179.2279**	**-5.61259**	**4.93E-11**
LINC00261	**170.8218**	**-5.49776**	**0.008725**
SLC30A8	**112.1498**	**-5.46381**	**2.28E-10**
DCN	**236.3414**	**-5.43434**	**2.11E-05**
COL6A3	**61.92872**	**-5.34214**	**1.61E-06**
FOXA1	**275.0448**	**-5.31634**	**5.59E-05**
LRRC61	**223.5458**	**-5.30044**	**3.85E-08**
AGT	**159.2011**	**-5.24702**	**7.07E-06**
ZNF736	**406.3007**	**-5.17735**	**1.52E-05**
ZNF667.AS1	**1343.099**	**-5.11138**	**6.28E-06**
SLITRK4	**1475.801**	**-4.92694**	**1.71E-05**
FERD3L	**89.10706**	**-4.89566**	**3.01E-05**
KCNMB1	**30.99346**	**-4.87673**	**2.29E-08**
PROCR	**20.19262**	**-4.83446**	**1.72E-09**
CECR7	**7.966717**	**-4.7715**	**0.000204**

A


D100 Ventral forebrain
NS
Log2 FC
p-value
p – value and log2 FC
FABP7
CHL1
RPE65
FOLH1
COL8A1
LOC100288778
TBR1
PEG3
GORIN
NTN1
ID3
FAM135B
FAM53B
VAV3
SIX6
PRELP
PLOD1
TSSC2
DDX11L2
ZNF558
TCTEX1D1
LOC100130417
−Log10 P
Log2 fold change

B


GO Enrichment: molecular function
G protein-coupled receptor binding
G protein-coupled receptor activity
Receptor ligand activity
Signaling receptor activator activity
Signaling receptor regulator activity
RNA polymerase II cis-regulatory region sequence-specific DNA binding
Transmembrane signaling receptor activity
Cis-regulatory region sequence-specific DNA binding
DNA-binding transcription factor activity RNA polymerase II-specific
DNA-binding transcription activator activity RNA polymerase II-specif
DNA-binding transcription activator activity
Signaling receptor activity
Molecular transducer activity
DNA-binding transcription factor activity
Sequence-specific double-stranded DNA binding
Transcription cis-regulatory region binding
Transcription regulatory region nucleic acid binding
Double-stranded DNA binding
Sequence-specific DNA binding
Signaling receptor binding
Fold Enrichment
N. of Genes
-log10(FDR)

C

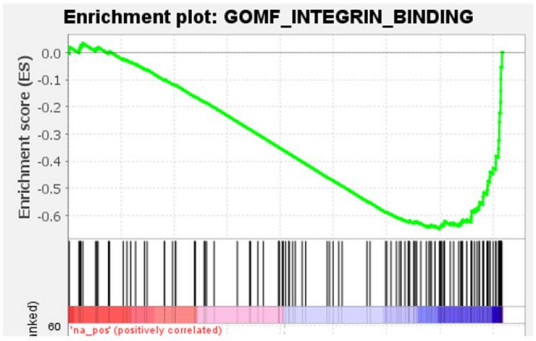

Enrichment plot: GOMF_INTEGRIN_BINDING
Enrichment score (ES)
'na_pos' (positively correlated)

Figure 28: Neuron-specific signaling pathways are reduced in 15q13.3del mature ventral forebrain organoids.

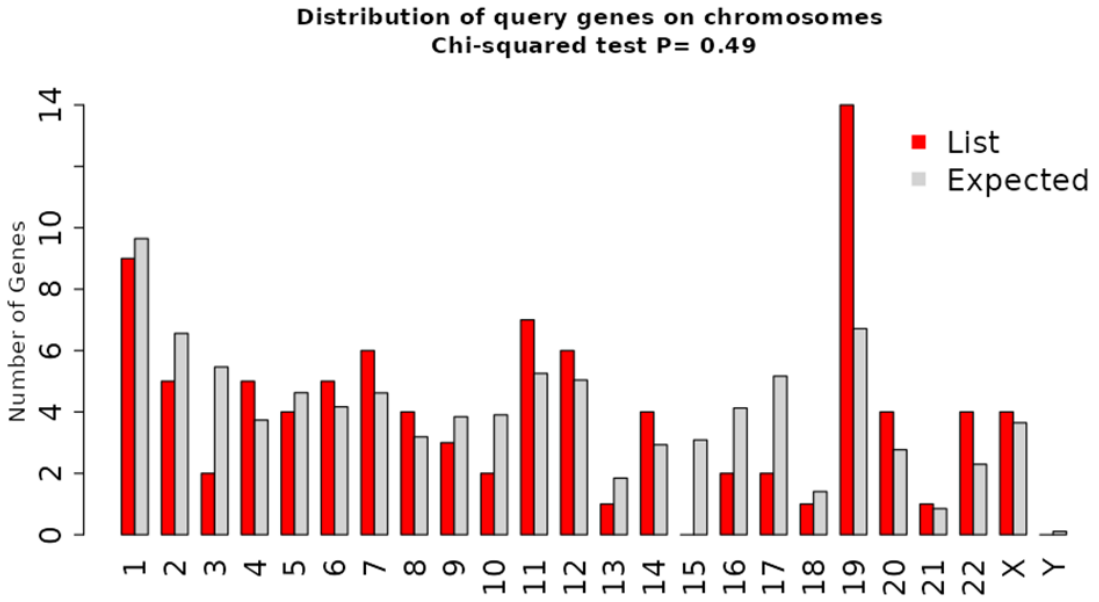
(A) Volcano plot of DEGs at D100 (1 761 upregulated, 1 933). Blue: pAdj <0.01, red: Log2FC > 1.
(B) Gene ontology for biological pathways show enrichment for synaptic signaling pathways at the day 100 timepoint.
(C) GSEA of ranked D100 genes show disruptions in integrin cell adhesion pathway, ES = -0.65, FDR q value = 0.007.

Table 6: Top 25 downregulated genes in ventral forebrain bulk RNA sequencing (Day 100 Timepoint).

List of top 25 downregulated genes in day 100 ventral forebrain dataset. Genes filtered by pAdj < 0.05 and ranked by log2FoldChange.

	baseMean	log2FoldChange	pAdj
FOLH1	101.2478	-7.83901	1.24E-25
FAM135B	360.1774	-7.52242	4.43E-18
COL3A1	1494.2	-7.41097	8.29E-06
SIX6	186.57	-7.09227	1.57E-14
F13A1	54.44859	-6.51084	1.25E-05
PEG3	6001.094	-6.28824	7.86E-22
ZNF257	11.80996	-6.01378	0.004992
GGT5	40.00518	-5.99019	5.71E-11
ETNPPL	331.1101	-5.9417	9.06E-16
TWIST1	35.20769	-5.86834	0.00011
MGP	32.313	-5.74005	1.88E-07
FOXA1	293.8018	-5.67381	4.53E-07
LDOC1	1647	-5.63985	7.79E-07
COL8A1	878.785	-5.58496	1.34E-26
ZNF558	574.8074	-5.49471	2.34E-08
ZNF208	66.1937	-5.45779	0.0103
PI16	416.5701	-5.45302	2.71E-11
COL22A1	168.4798	-5.43698	2.85E-13
SLITRK4	2919.432	-5.42347	6.84E-06
CTSF	564.3847	-5.41222	5.05E-08
LOC403323	24.92178	-5.23568	3.78E-11
PCDHGA7	216.3414	-5.18884	1.51E-08
FOXA2	61.15958	-4.97349	1.01E-05
ZNF728	93.04832	-4.95356	0.000751
ZNF667.AS1	1097.887	-4.8637	3.47E-06

D50/D100 Combined Ventral forebrain analysis



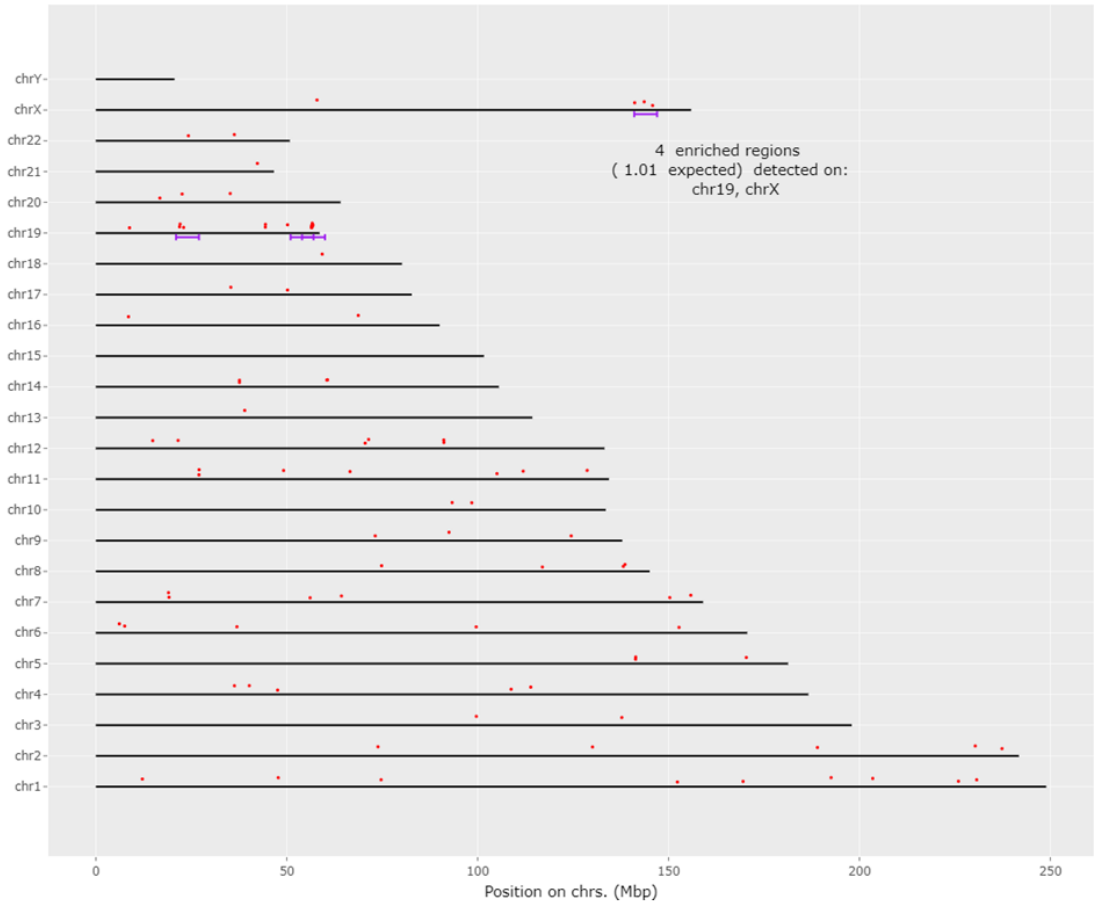


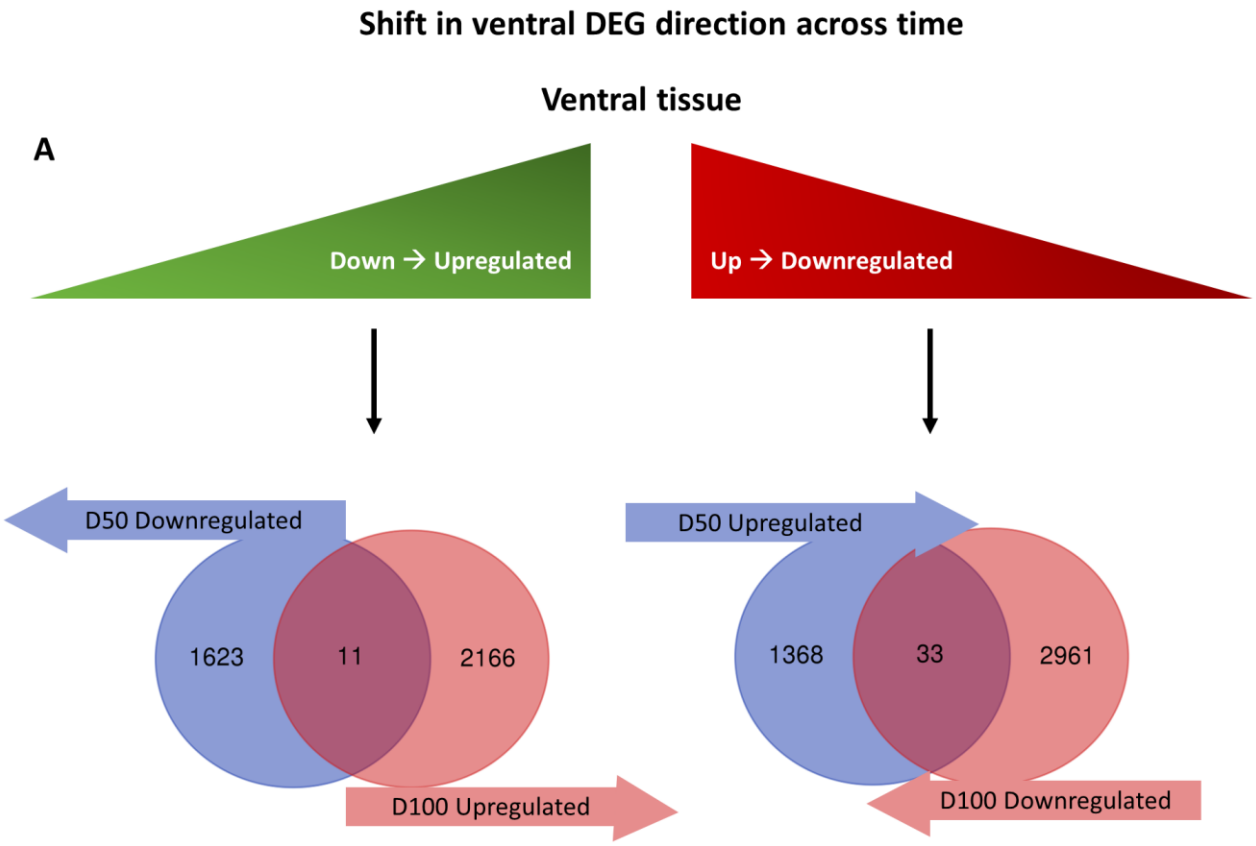
Figure 29: 15q13.3del ventral forebrain organoids share enrichment for Krüppel-domain zinc finger proteins on chromosome 19.

Chromosomal mapping of common DEGs between forebrain tissues at each timepoints share enrichment for chromosome 19p13.2 and 19q13.43.

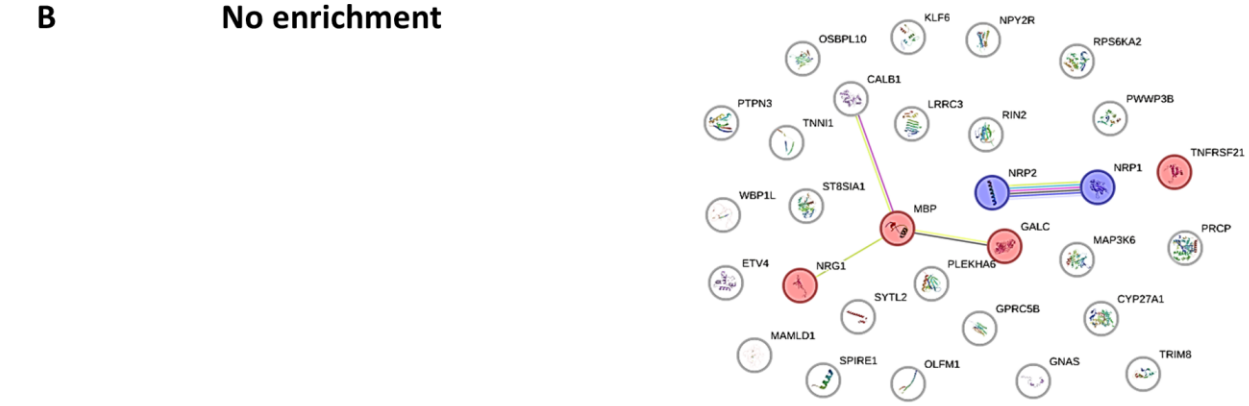

A
D50/D100 Combined Ventral forebrain analysis:
Divergent pathways in downregulated gene sets
D50 Ventral downregulated
D100 Ventral downregulated
542
1092
1902
B
Ribosomal large subunit assembly
Cytoplasmic translation
Ribosomal large subunit biogenesis
Ribosome biogenesis
Translation
Peptide biosynthetic proc.
Amide biosynthetic proc.
Peptide metabolic proc.
Cellular amide metabolic proc.
Cellular macromolecule biosynthetic proc.
Organonitrogen compound biosynthetic proc.
0.0 2.5 5.0 7.5
Fold Enrichment
N. of Genes
20
40
60
-log10(FDR)
5
10
C
Response to oxygen-containing compound
Cell adhesion
Cell migration
Organonitrogen compound catabolic proc.
Positive reg. of protein metabolic proc.
Vesicle-mediated transport
Response to endogenous stimulus
Protein transport
Intracellular transport
Neurogenesis
0.0 0.5 1.0 1.5
Fold Enrichment
N. of Genes
190
200
210
-log10(FDR)
2
3
4
5
6
D
Animal organ morphogenesis
Anatomical structure formation involved in morphogenesis
Tube development
Epithelium development
Cell adhesion
Cell migration
Reproductive proc.
Positive reg. of multicellular organismal proc.
Reproduction
Response to endogenous stimulus
Positive reg. of developmental proc.
Positive reg. of signal transduction
Negative reg. of cell communication
Response to oxygen-containing compound
Negative reg. of response to stimulus
Reg. of cell differentiation
Cell projection organization
Plasma membrane bounded cell projection organization
Neurogenesis
Cell-cell signaling
0.0 0.5 1.0 1.5 2.0
Fold Enrichment
N. of Genes
120
130
140
-log10(FDR)
2
4
6
8
10
12

Figure 30: Cell migration disruptions are enriched in shared 15q13.3del D50/D100 DEGs as well as D100-unique genes.

(A) Venn diagram of common and unique downregulated genes in 15q13.3del ventral forebrain tissue across time.
(**B)** GO analysis for biological pathways shows enrichment for ribosomal pathways amongst DEGs unique to the D50 timepoint.
(**C)** GO analysis for biological pathways reveals enrichment for cell migration and cell adhesion pathways amongst DEGs unique to the D100 timepoint.
(**D)** Shared D50 and D100 ventral downregulated genes show enrichment for cell migration and cell adhesion pathways.


Shift in ventral DEG direction across time
Ventral tissue
A
Down → Upregulated
Up → Downregulated
D50 Downregulated
1623
11
2166
D100 Upregulated
D50 Upregulated
1368
33
2961
D100 Downregulated

B


No enrichment
KLF6
NPY2R
OSBPL10
RPS6KA2
CALB1
PTPN3
LRRC3
PWWP3B
TNNI1
RIN2
TNFRSF21
NRP2
NRP1
WBP1L
ST8SIA1
MBP
GALC
PRCP
MAP3K6
ETV4
NRG1
PLEKHA6
SYTL2
GPRC5B
CYP27A1
MAMLD1
SPIRE1
OLFM1
GNAS
TRIM8

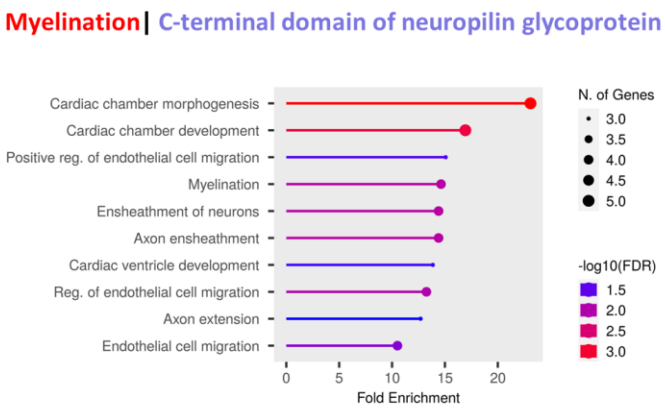

Myelination| C-terminal domain of neuropilin glycoprotein
Cardiac chamber morphogenesis
Cardiac chamber development
Positive reg. of endothelial cell migration
Myelination
Ensheathment of neurons
Axon ensheathment
Cardiac ventricle development
Reg. of endothelial cell migration
Axon extension
Endothelial cell migration
0
5
10
15
20
Fold Enrichment
N. of Genes
3.0
3.5
4.0
4.5
5.0
-log10(FDR)
1.5
2.0
2.5
3.0

Figure 31: Maturation of ventral 15q13.3del organoids promotes a shift in DEG direction of axon ensheathment genes.

(A) Venn diagram of common genes in left: D50 downregulated and D100 upregulated ventral forebrain tissue produces no statistical enrichment during GO analysis and (right) D50 upregulated and D100 downregulated ventral forebrain tissue.
(**B)** GO analysis for biological pathways show enrichment for myelinating pathways in common D50/D100 DEGs of opposite trajectory (up → downregulation).

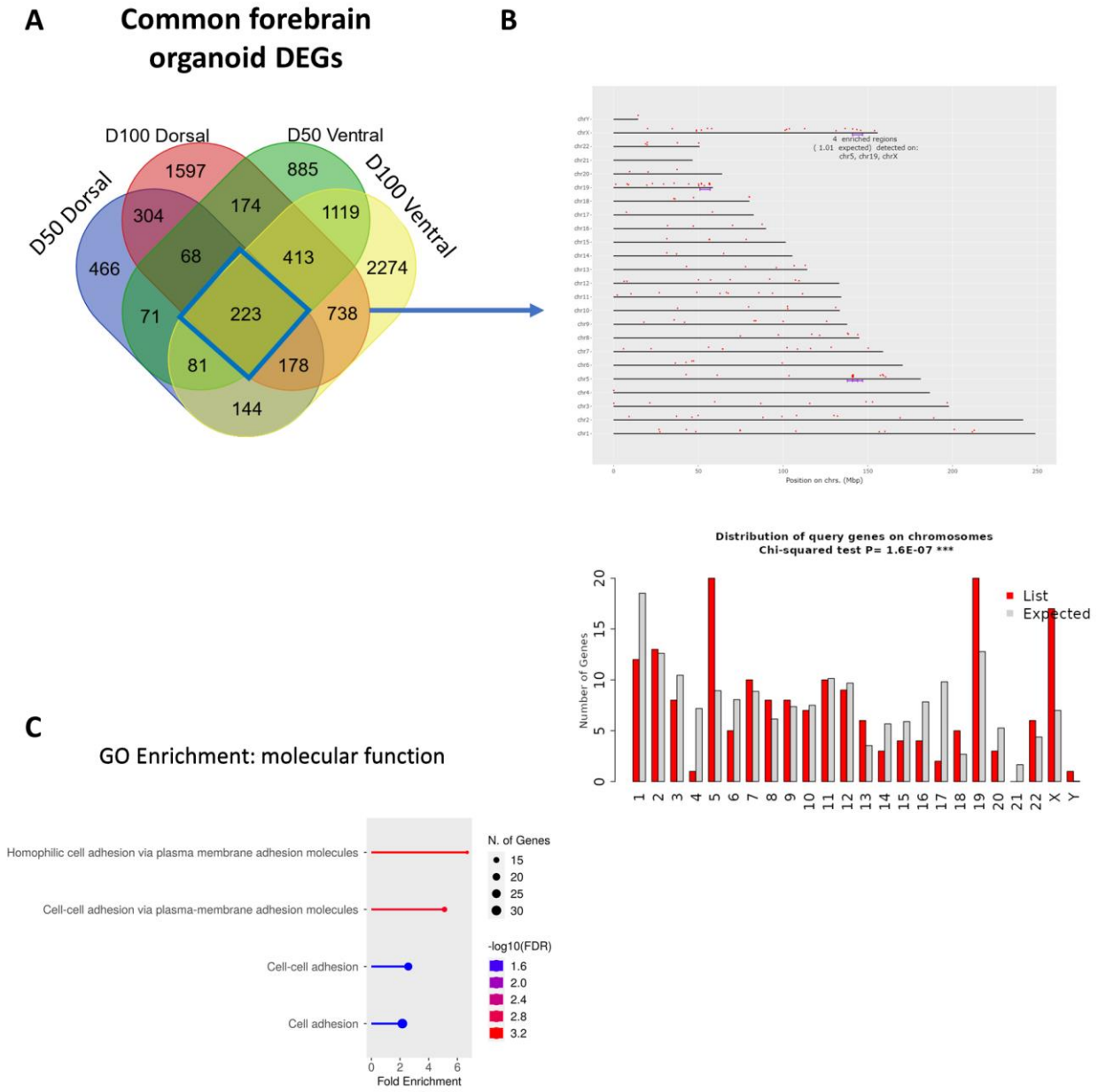


Figure 32: Shared forebrain organoid DEGs converge on genes relating to cell adhesion.

(A) Venn diagram of common and unique DEGs in 15q13.3del forebrain tissue.
(B) Chromosomal mapping of common DEGs reveals enrichment for chromosome 19 (FDR <0.05, Chi squared test p = 1.67e-07).
(C) Gene ontology for biological pathways show enrichment (day 100 timepoint only) for cell adhesion exclusively.

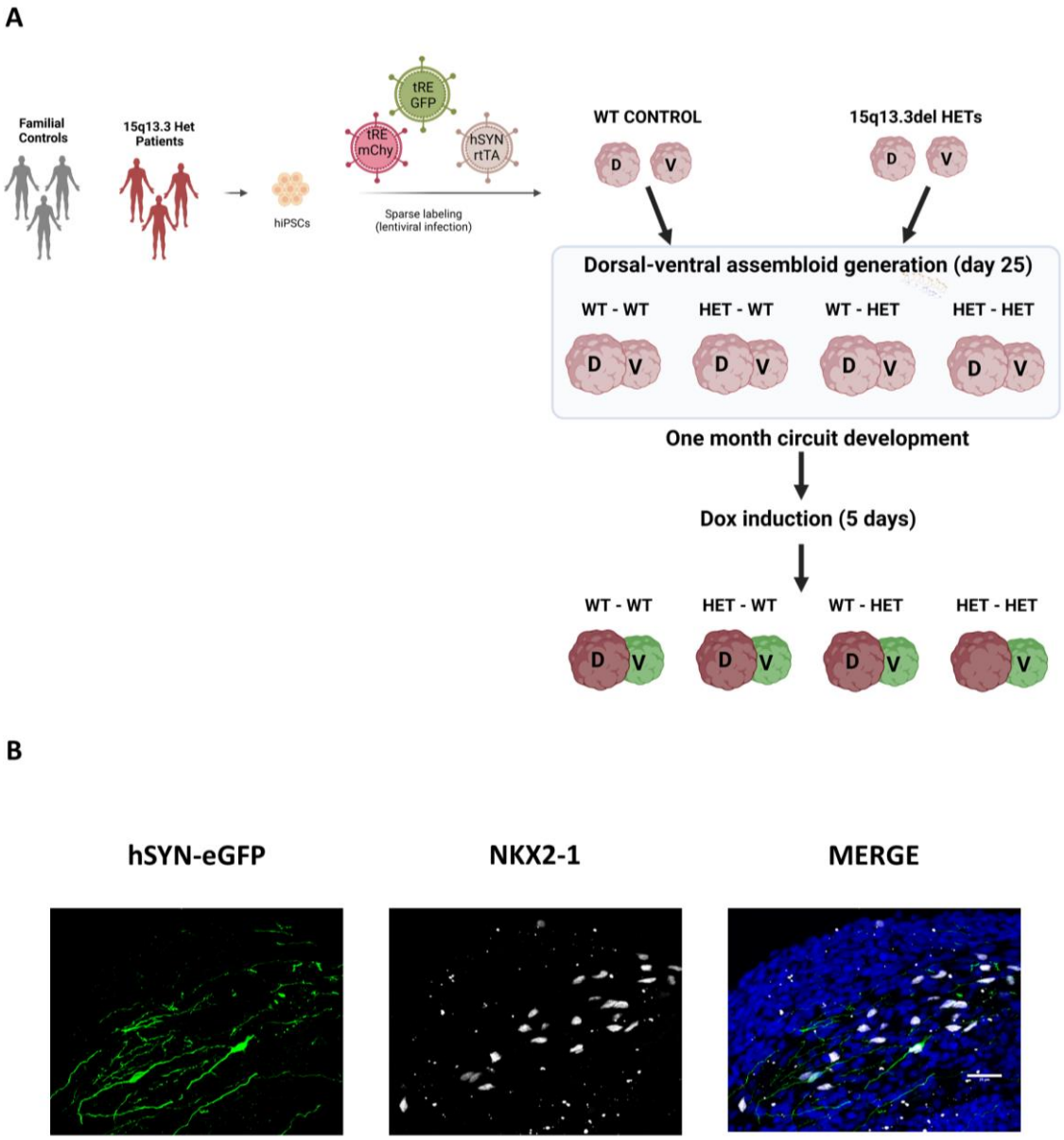
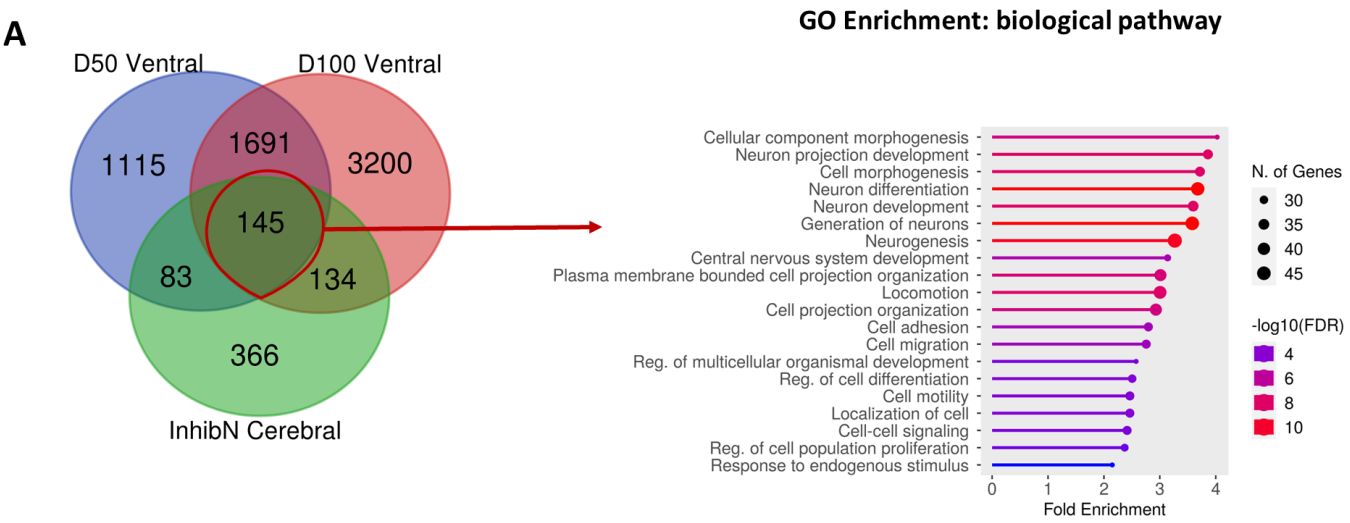
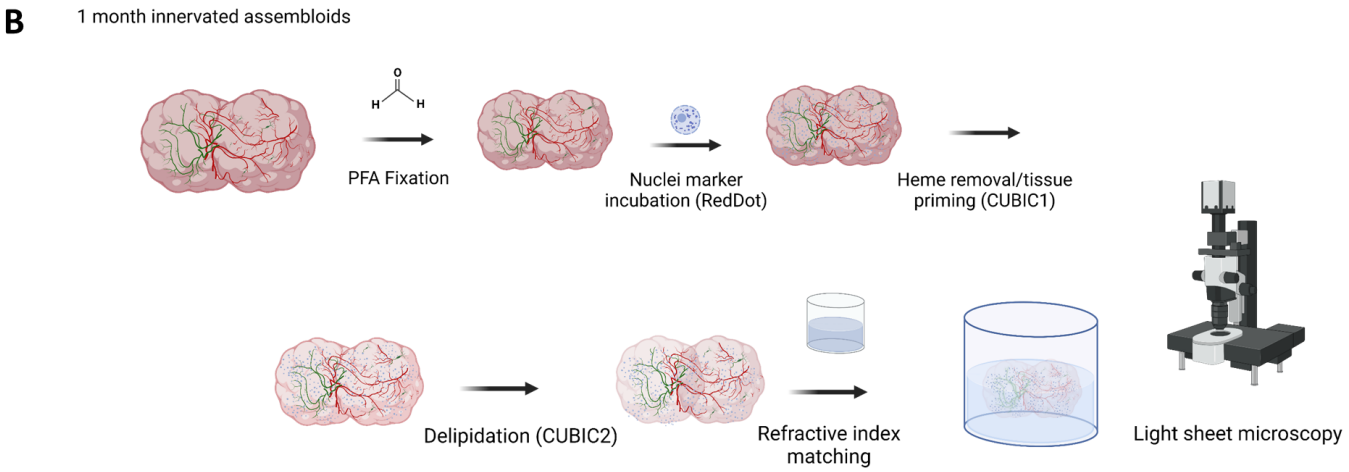


Figure 33: Dorsal-ventral forebrain assembloid pipeline.

(A) Experiment schematic and genotype/lineage combinations of dorsal-ventral forebrain assembloids.

(B) Migrating hSYN-eGFP^{+} interneurons are immunopositive for the inhibitory neuron marker NKX2-1.


A
D50 Ventral
D100 Ventral
InhibN Cerebral
1115
1691
3200
145
83
134
366
GO Enrichment: biological pathway
Cellular component morphogenesis
Neuron projection development
Cell morphogenesis
Neuron differentiation
Neuron development
Generation of neurons
Neurogenesis
Central nervous system development
Plasma membrane bounded cell projection organization
Locomotion
Cell projection organization
Cell adhesion
Cell migration
Reg. of multicellular organismal development
Reg. of cell differentiation
Cell motility
Localization of cell
Cell-cell signaling
Reg. of cell population proliferation
Response to endogenous stimulus
0
1
2
3
4
Fold Enrichment
N. of Genes
30
35
40
45
-log10(FDR)
4
6
8
10


B
1 month innervated assembloids
PFA Fixation
Nuclei marker incubation (RedDot)
Heme removal/tissue priming (CUBIC1)
Delipidation (CUBIC2)
Refractive index matching
Light sheet microscopy

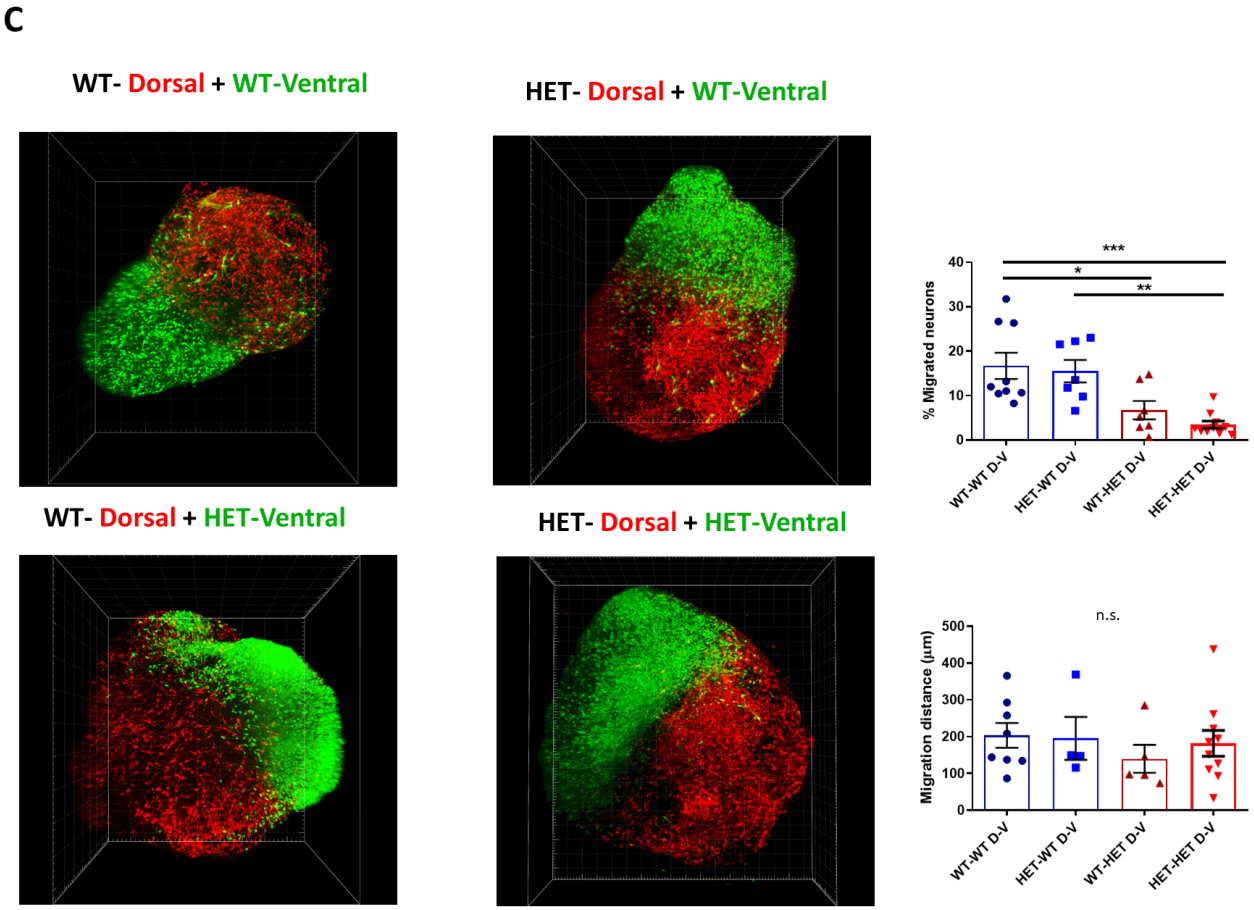
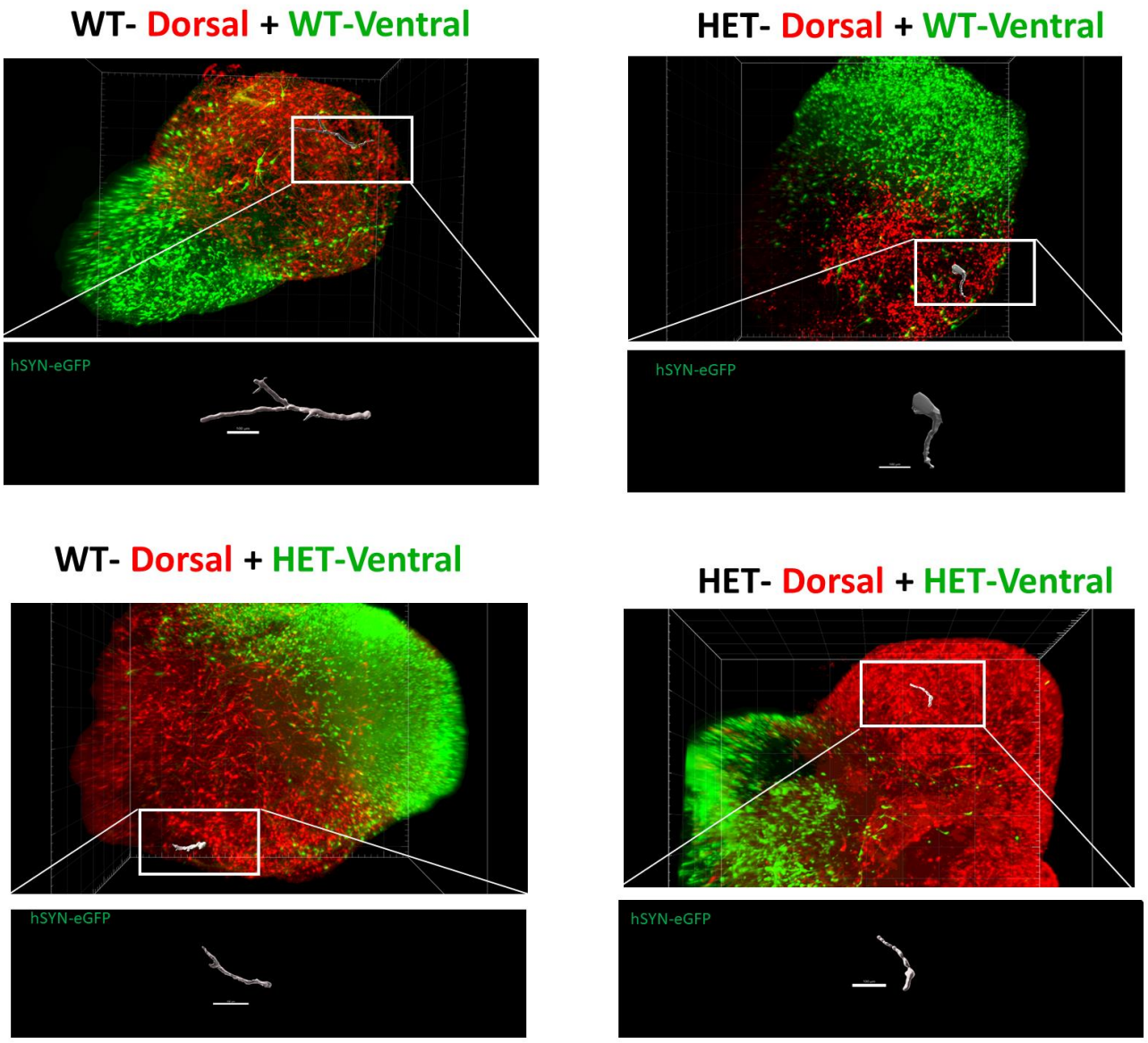

C
WT- Dorsal + WT-Ventral
HET- Dorsal + WT-Ventral
WT- Dorsal + HET-Ventral
HET- Dorsal + HET-Ventral
% Migrated neurons
***
*
**
WT-WT D-V
HET-WT D-V
WT-HET D-V
HET-HET D-V
Migration distance (µm)
n.s.

Figure 34: 15q13.3del forebrain interneurons show deficits in neuron migration.

(A) Common DEGs amongst ventral forebrain bulk transcriptomics and UNO inhibitory scRNA Seq converge on neuron migration pathways.
(B): CUBIC tissue clearing schematic for light sheet microscopy imaging.
(C) (Left) Representative 2x IMARIS image of all combinations for migration experiment with quantified migration proportion (N= 5-6 assembloids/combination) amongst three 15q13.3del families. (Right) quantification of proportion of migrating neurons and distance migrated. *Data points represent mean* $\pm SEM$ *in full assembloids; *n.s. = nonsignificant p<0.05, **p<0.01, ***p<0.001; One-Way ANOVA with Tukey's post-hoc.*

A


WT- Dorsal + WT-Ventral
hSYN-eGFP
HET- Dorsal + WT-Ventral
hSYN-eGFP
WT- Dorsal + HET-Ventral
hSYN-eGFP
HET- Dorsal + HET-Ventral
hSYN-eGFP

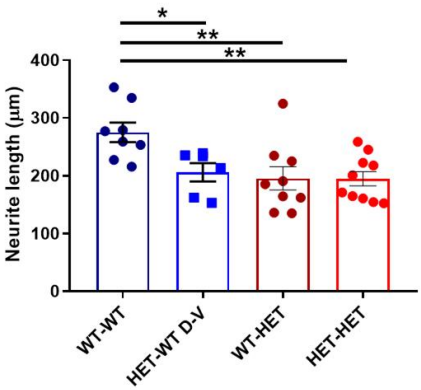

*
**
**
Neurite length (µm)
400
300
200
100
0
WT-WT
HET-WT D-V
WT-HET
HET-HET

Figure 35: Morphological abnormalities in 15q13.3 ventral forebrain interneurons.

IMARIS representative reconstruction of a migrating neuron with average neurite length per assembloid below. *Data points represent mean* $\pm SEM$ *in full assembloids; *n.s. = nonsignificant p<0.05, **p<0.01, ***p<0.001; One-Way ANOVA with Tukey's post-hoc.*

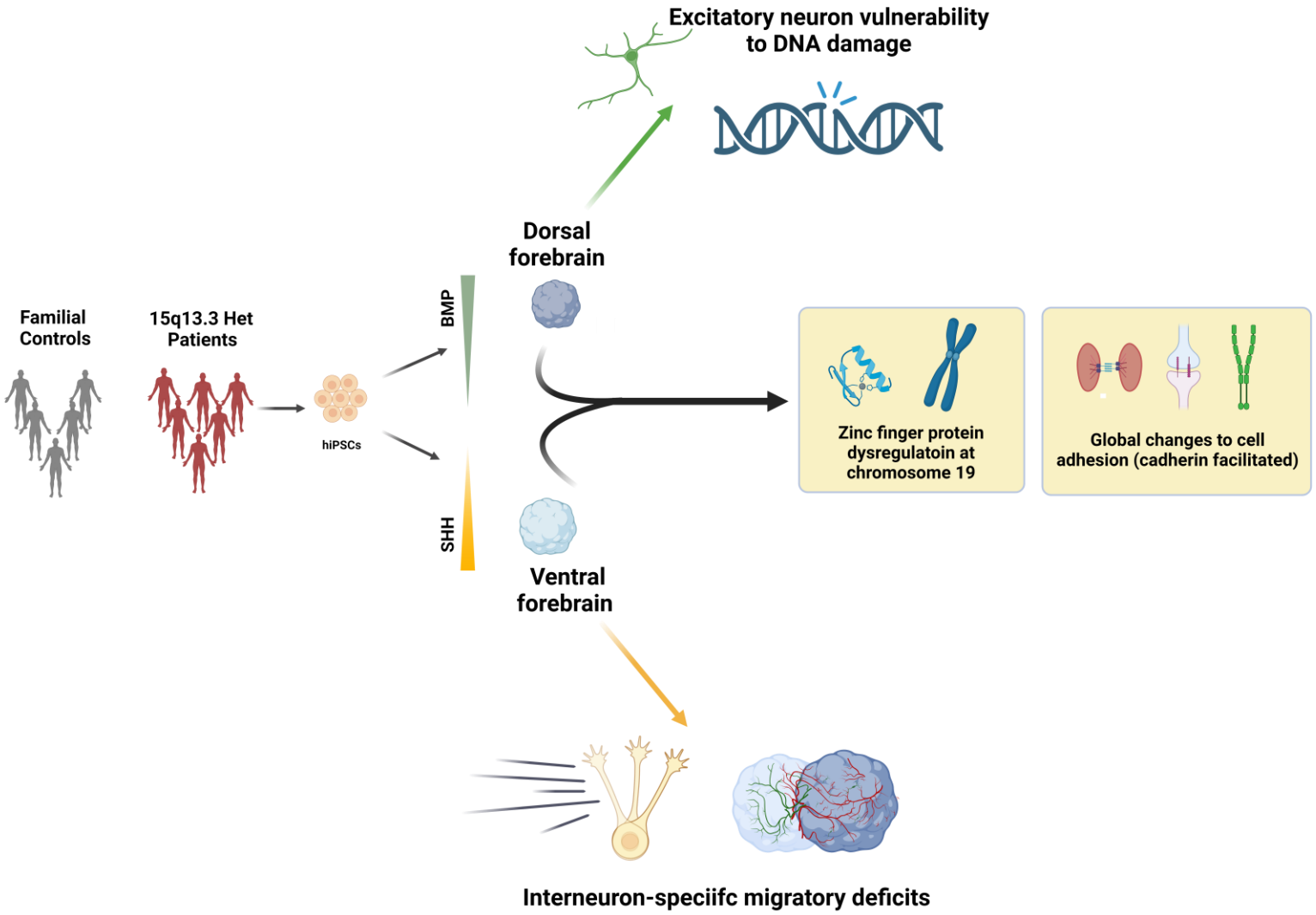
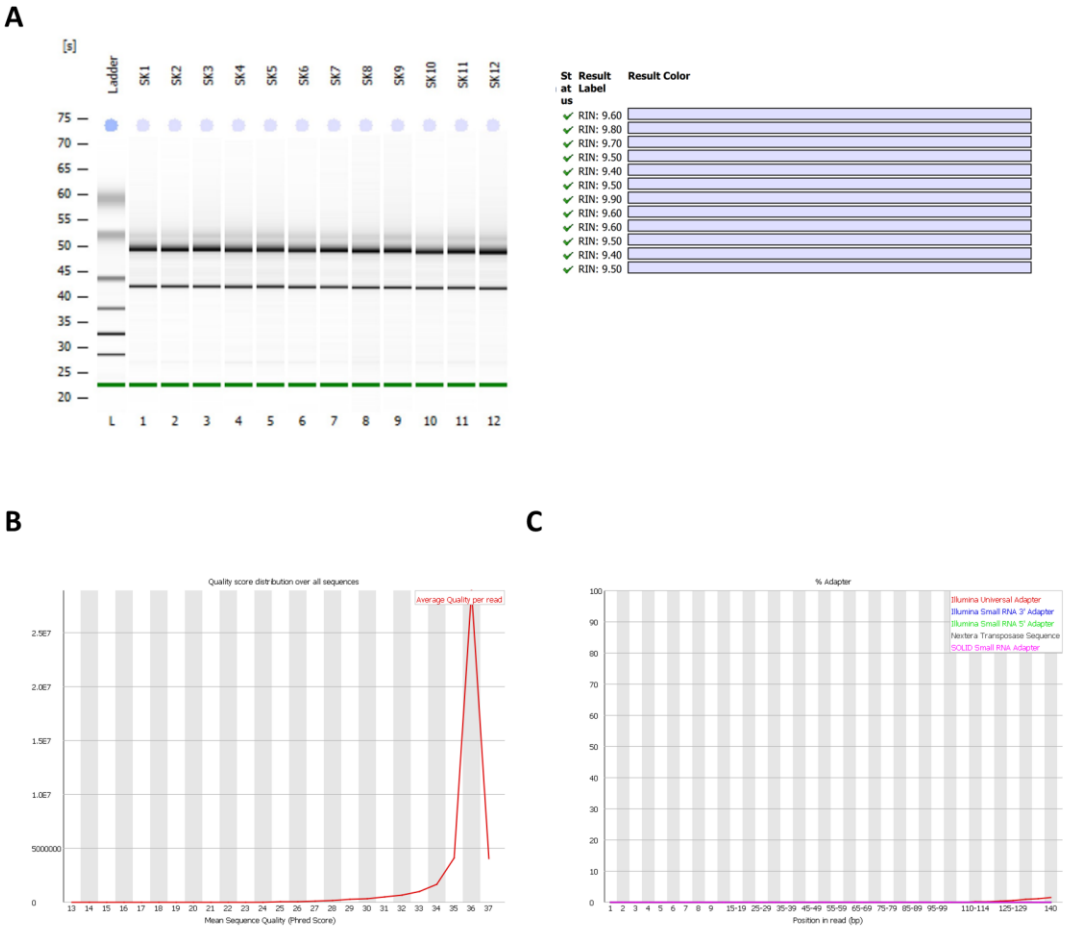


Figure 36: Summary of guided 15q13.3del organoid approaches.

Bulk transcriptomics on 15q13.3del dorsal (top) and ventral (bottom) forebrain organoids converges on disruptions to chromosome 19 zinc finger proteins and cell adhesion and demonstrated excitatory neuron-specific DNA damage and interneuron-specific migration deficits.

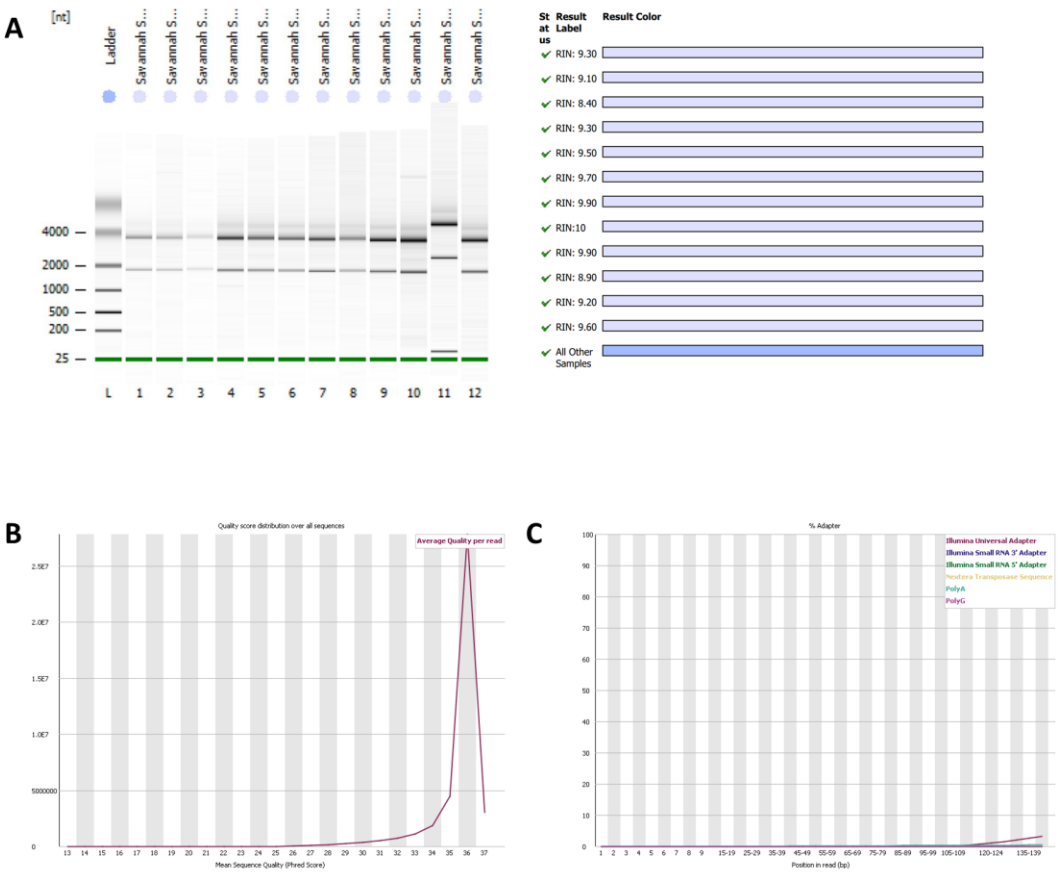
4.9 Chapter 4 Supplementary Data



Supplementary Figure 12: Dorsal forebrain organoid bulk RNA sequencing passes *in house* quality control measures.

(A) BioAnalyzer results of ribosomal subunits shows high RNA integrity number (RIN) of ribosomal ratios.

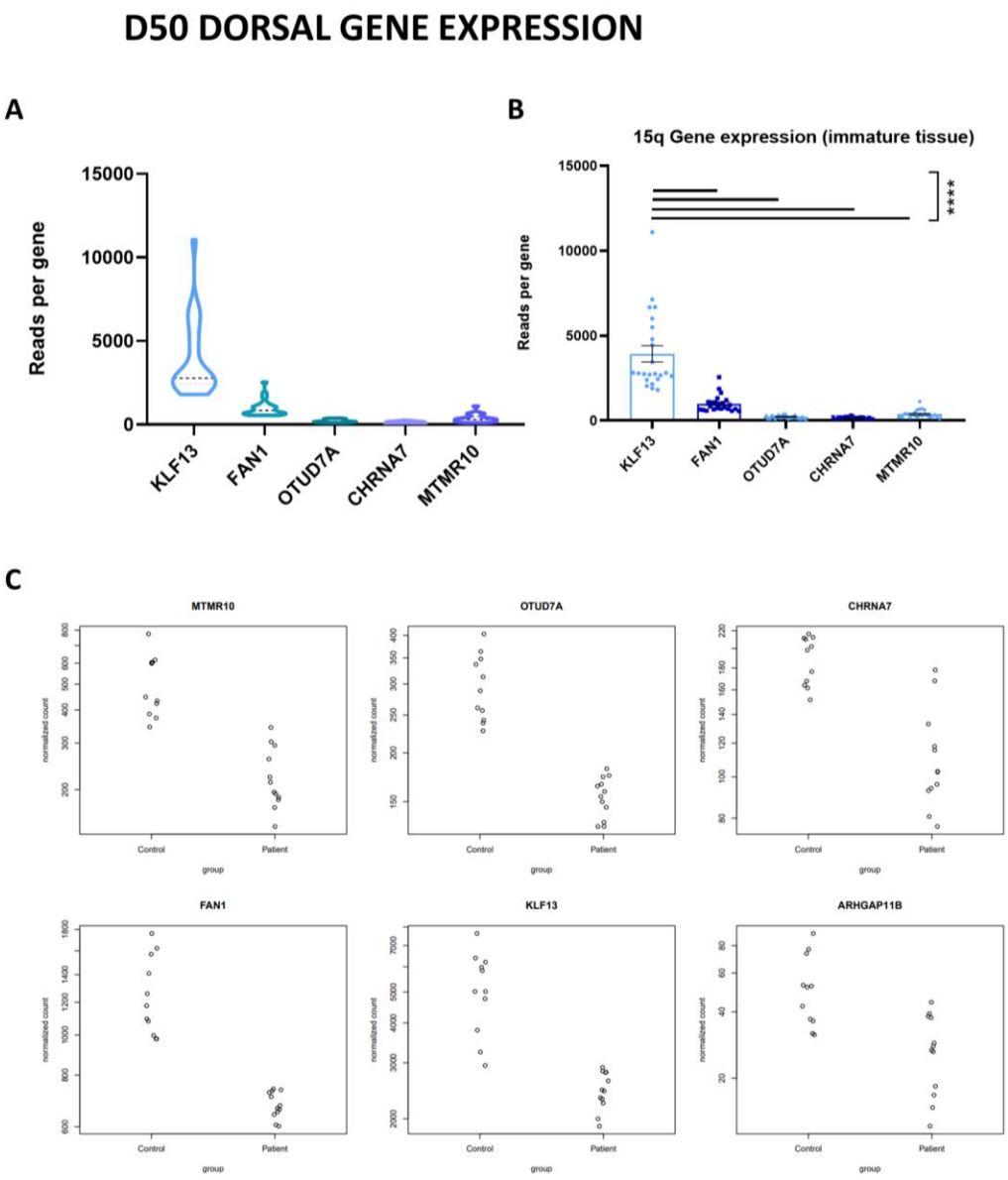
(B) FastQC assessment of read quality (left, Phred score) and **(C)** adapter contamination in raw reads from dorsal bulk RNA Seq.



Supplementary Figure 13: Ventral forebrain organoid bulk RNA sequencing passes in house quality control measures.

(A) BioAnalyzer results of ribosomal subunits shows high RNA integrity number (RIN) of ribosomal ratios.

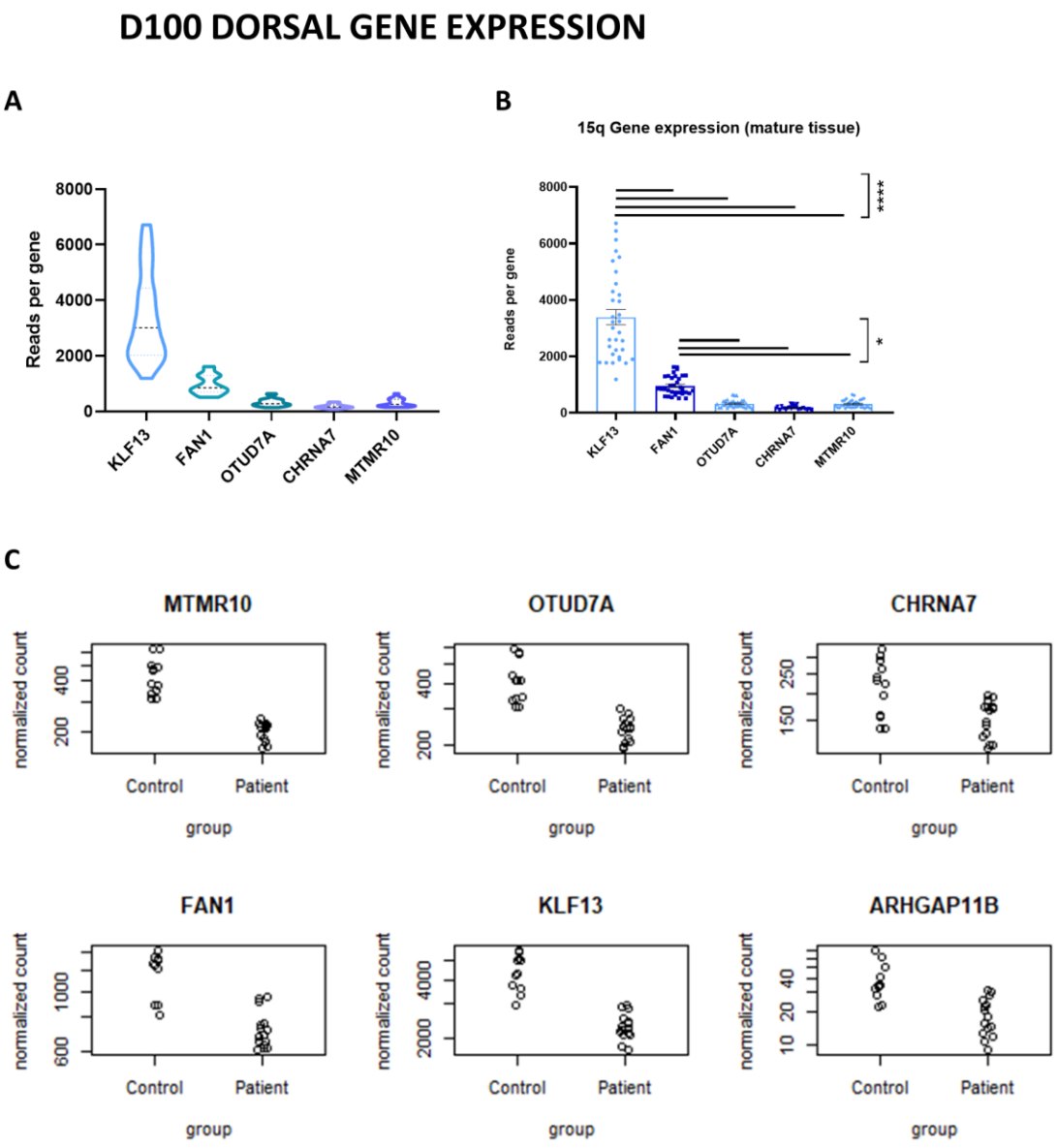
(B) FastQC assessment of read quality (left, Phred score) and **(C)** adapter contamination in raw reads from ventral bulk RNA Seq.



Supplementary Figure 14: 15q13.3 gene mRNA expression in D50 dorsal forebrain organoids.

(A-B) Raw D50 dorsal forebrain organoid read counts for 15q13.3 genes *KLF13, FAN1, OTUD7A, CHRNA7*, and *MTMR10* in (**A**) violin plots and (**B**) quantitative comparison between genes. *Data points represent individual sample counts, ****p<0.0001; One-Way ANOVA with Tukey's post-hoc.*

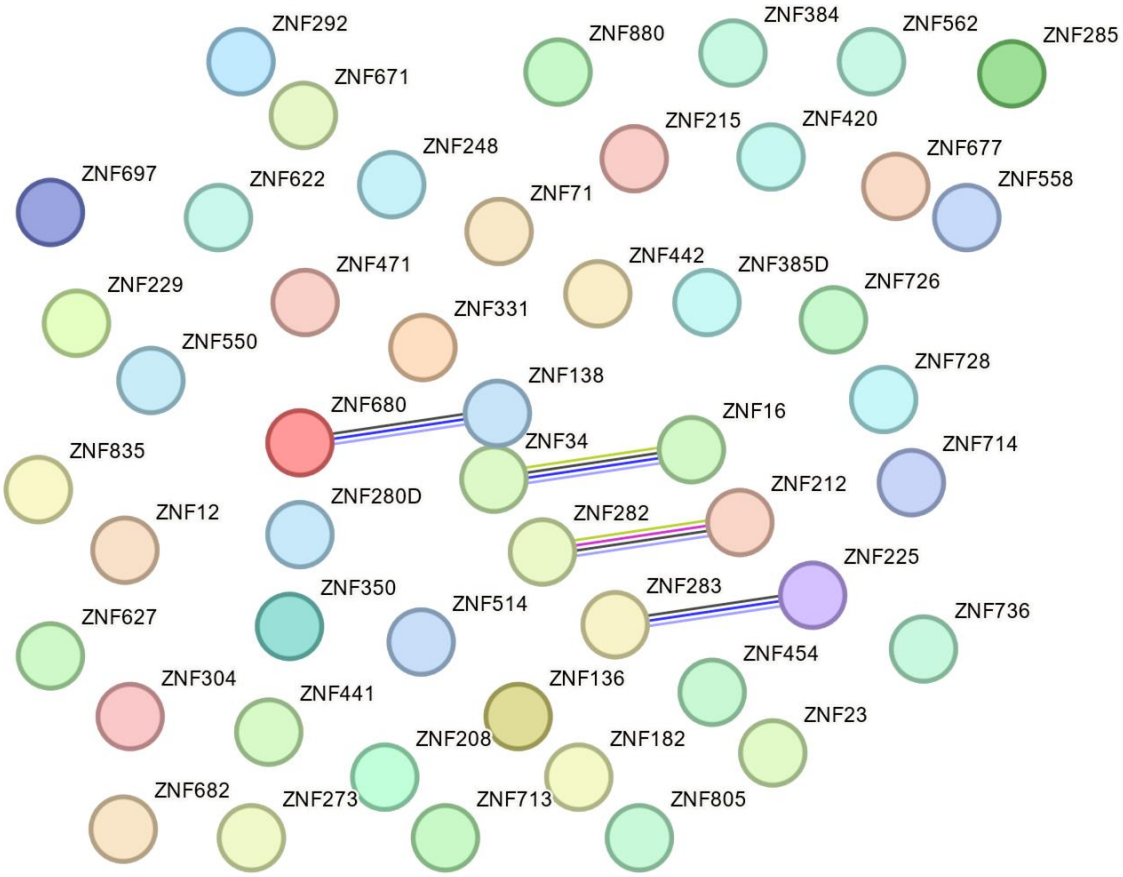
(**C)** Normalized D50 dorsal forebrain organoid read counts for 15q13.3 genes *KLF13, FAN1, OTUD7A, CHRNA7,* and *MTMR10.*



Supplementary Figure 15: 15q13.3 gene mRNA expression in D100 dorsal forebrain organoids.

(A-B) Raw D100 dorsal forebrain organoid read counts for 15q13.3 genes *KLF13, FAN1, OTUD7A, CHRNA7,* and *MTMR10* in (**A**) violin plots and (**B**) quantitative comparison between genes. *Data points represent individual sample counts, *p<0.05, ****p<0.0001; One-Way ANOVA with Tukey's post-hoc.*

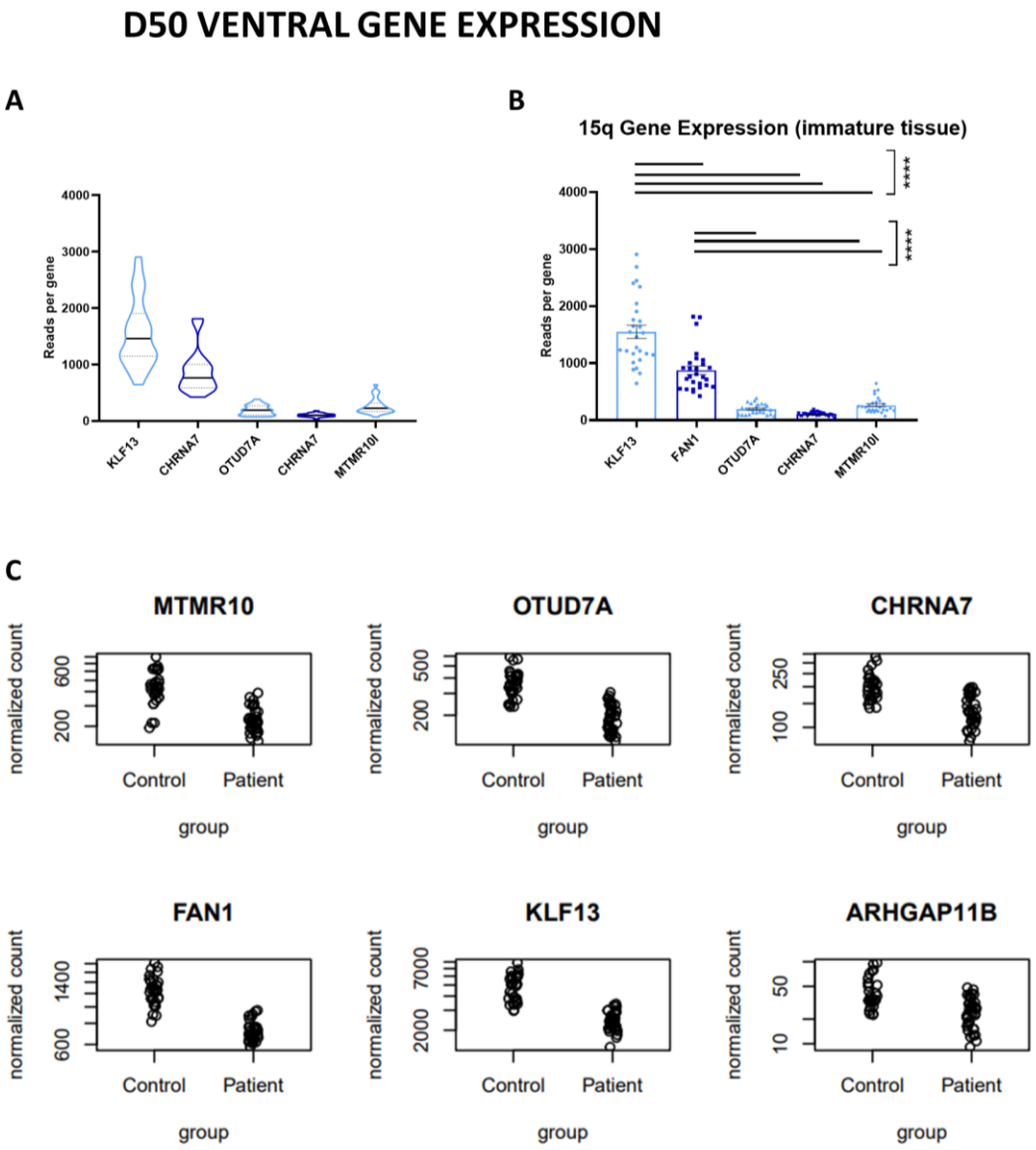
(**C**) Normalized D50 dorsal forebrain organoid read counts for 15q13.3 genes *KLF13, FAN1, OTUD7A, CHRNA7,* and *MTMR10.*



Reactome Pathways				
pathway	*description*	*count in network*	*strength*	*false discovery rate*
HSA-212436	Generic Transcription Pathway	35 of 1215	1.06	4.63e-28

Supplementary Figure 16: Enriched zing finger proteins do not converge on known biological pathways.

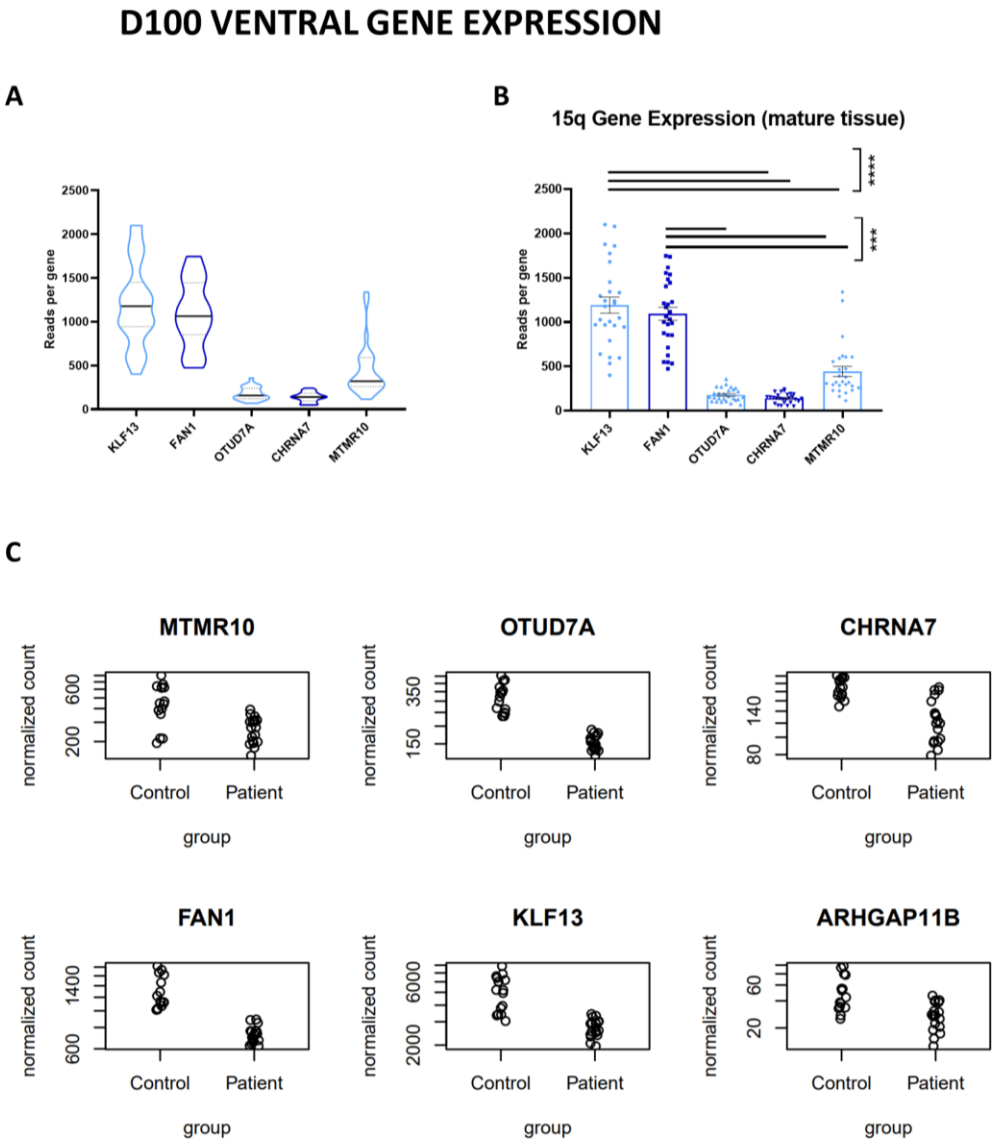
(Top) STRING visualization and (bottom) reactome pathway results of D50 dorsal forebrain downregulated ZNF genes shows unspecific enrichment for transcription pathways.



Supplementary Figure 17: 15q13.3 gene mRNA expression in D50 ventral forebrain organoids.

(A-B) Raw D50 ventral forebrain organoid read counts for 15q13.3 genes *KLF13*, *FAN1*, *OTUD7A*, *CHRNA7*, and *MTMR10* in (**A**) violin plots and (**B**) quantitative comparison between genes. *Data points represent individual sample counts, ****p<0.0001; One-Way ANOVA with Tukey's post-hoc.*

(**C**) Normalized D50 dorsal forebrain organoid read counts for 15q13.3 genes *KLF13*, *FAN1*, *OTUD7A*, *CHRNA7*, and *MTMR10*.

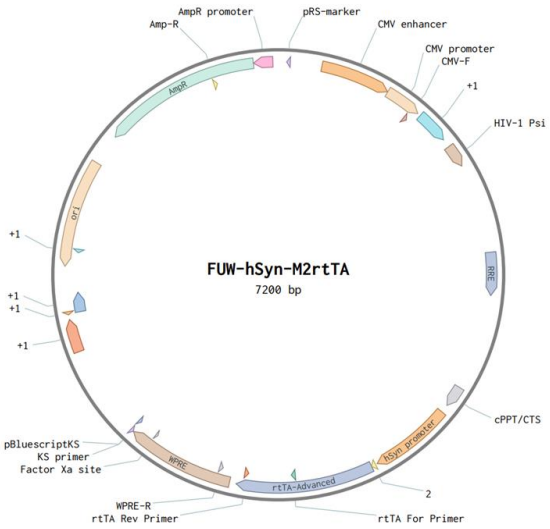


Supplementary Figure 18: 15q13.3 gene mRNA expression in D100 ventral forebrain organoids.

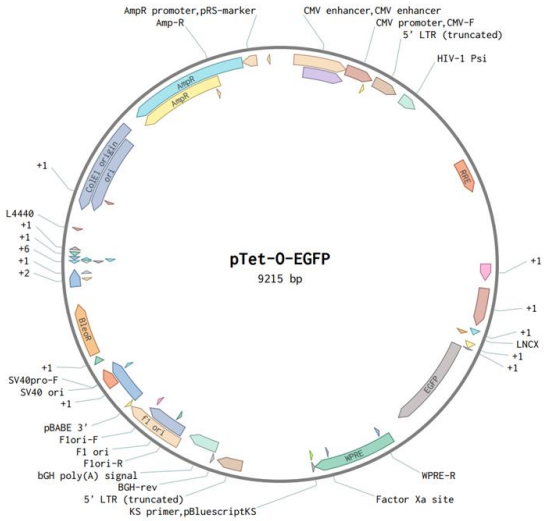
(A-B) Raw D100 ventral forebrain organoid read counts for 15q13.3 genes *KLF13*, *FAN1*, *OTUD7A*, *CHRNA7*, and *MTMR10* in (**A**) violin plots and (**B**) quantitative comparison between genes. *Data points represent individual sample counts, ***p<0.001, ****p<0.0001; One-Way ANOVA with Tukey's post-hoc.*

(**C)** Normalized D50 dorsal forebrain organoid read counts for 15q13.3 genes *KLF13*, *FAN1*, *OTUD7A*, *CHRNA7*, and *MTMR10*.

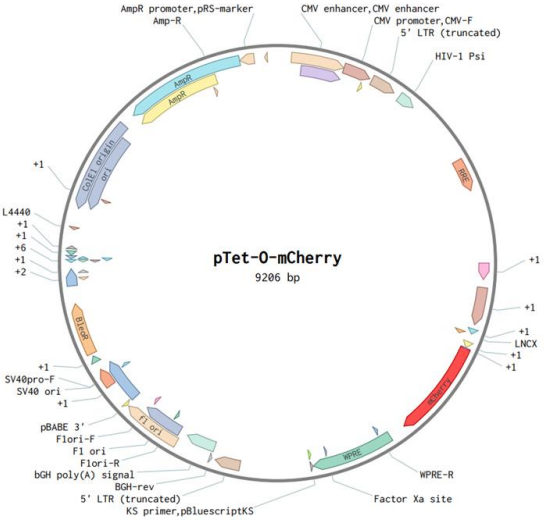
A


AmpR promoter
Amp-R
pRS-marker
CMV enhancer
CMV promoter
CMV-F
+1
HIV-1 Psi
RRE
cPPT/CTS
hSyn promoter
2
rtTA For Primer
rtTA-Advanced
WPRE
WPRE-R
rtTA Rev Primer
pBluescriptKS
KS primer
Factor Xa site
+1
+1
+1
+1
FUW-hSyn-M2rtTA
7200 bp

B

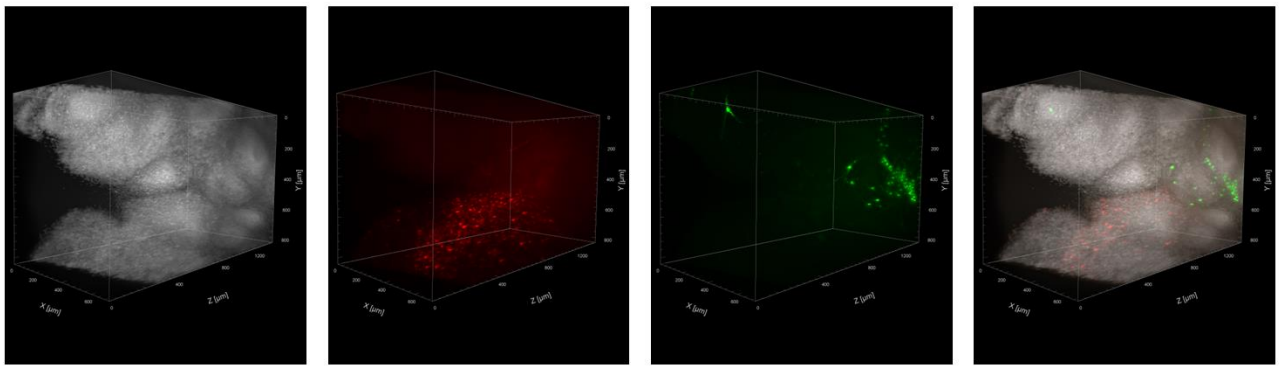

AmpR promoter,pRS-marker
Amp-R
CMV enhancer,CMV enhancer
CMV promoter,CMV-F
5' LTR (truncated)
HIV-1 Psi
RRE
LNCX
EGFP
WPRE
WPRE-R
Factor Xa site
5' LTR (truncated)
KS primer,pBluescriptKS
bGH poly(A) signal
BGH-rev
F1 ori
Flori-R
pBABE 3'
Flori-F
SV40 ori
SV40pro-F
Blast
AmpR
AmpR
ColE1 origin
ori
L4440
+1
+1
+6
+1
+2
pTet-O-EGFP
9215 bp

C


AmpR promoter,pRS-marker
Amp-R
CMV enhancer,CMV enhancer
CMV promoter,CMV-F
5' LTR (truncated)
HIV-1 Psi
RRE
LNCX
mCherry
WPRE
WPRE-R
Factor Xa site
5' LTR (truncated)
KS primer,pBluescriptKS
bGH poly(A) signal
BGH-rev
F1 ori
Flori-R
pBABE 3'
Flori-F
SV40 ori
SV40pro-F
Blast
AmpR
AmpR
ColE1 origin
ori
L4440
+1
+1
+6
+1
+2
pTet-O-mCherry
9206 bp

Supplementary Figure 19: Lentiviral constructs for doxycycline-inducible fluorophore expression.

Lentiviral constructs for sparse labeling, designed and generated by Dr. Nadeem Murtaza.



Supplementary Figure 20: Cerebral-cerebral assembloids do not produce migrating neurons detectable with light sheet microscopy.

Representative images of UNO-UNO assembloids tagged with CMV-eGFP and CMV-mCherry lentiviral constructs do not show detectable migration of either cell type. Images taken at 2x magnification using the Ultramicroscope II light sheet microscope (LaVision BioTec GmbH, Bielefeld, Germany).

5 **CHAPTER 5**: Discussion and future directions

Microdeletions at the 15q13.3 locus are one of the most common CNVs associated with neurodevelopmental disorders.[183,184] Individuals harboring the deletion display variable clinical presentations which have complicated therapeutic interventions and our understanding of disease pathology.[68,71,78] Mouse and 2D neural models have broadened our understanding of the late-stage impairments in synaptic activity and maturation but have not addressed the early developmental processes that may be disrupted and result in – or contribute to – these findings. Using both unguided and forebrain-directed dorsal and ventral organoids from a large 15q13.3 patient cohort, we identified early disruptions in cellular communication, transcription, and neural circuit development across multiple time points of development. We applied multiple transcriptomic approaches in organoids and assembloid 15q13.3del models to identify persistent in cell adhesion that suggests phenotypic convergence on a previously unexplored regulator to 15q13.3 NDD pathogenesis.

5.1 Persistent deficits in cell adhesion may underlie structural changes and alterations to cellular crosstalk.

Macrocephaly is not a common feature in individuals with the 15q13.3 microdeletion syndrome, and of the clinical reports within the literature, there are roughly equal proportions of patients with microcephaly.[79] A 15q13.3 heterozygous mouse model has been described to have an increase in brain and ventricle volume, however it contains the less common C57BL/6 genetic background[115] and has not been recapitulated since. The study used MRI to measure the total brain volume at postnatal days 10, 20 and 40, and

observed an increase in total brain and ventricle volume at day 20 and 40 without any gross structural abnormalities or laminar malformations (confirmed by Nissl staining). This suggests a uniform and unspecific size change that reflects our own data; however, they did not perform proliferation experiments to find the mechanism to the size increase.

The use of unguided neural organoids allowed us to unbiasedly profile the unique cell types that emerge in early fetal development. By applying scRNA Seq, we identified changes in the cellular composition of radial glia populations at an early timepoint where 15q13.3del UNOs showed an increase in organoid size. We confirmed the change in radial glia proportions through ICC of the VZ-like rosette structures, and further classified structural changes to the rosette itself. Despite the proliferative capacity of radial glia and the persistent size changes in the 15q13.3del UNOs, we were unable to find changes in cell proliferation nor cell cycle exit at the D40 timepoint (**Fig. 6**), suggesting other mechanisms may drive the increase in size. It is possible that a hyper-proliferating cell population is present earlier in 15q13.3del UNO development, and that the enlarged organoid and rosette size are a product of prior proliferation. We quantified a small but significant decrease in cell density amongst multiple 15q13.3del families at this timepoint, which may contribute to the size changes observed in the UNOs.

Due to the high proportion of radial glia and intermediate progenitors at the D40 timepoint and their capacity to differentiate into multiple cell types, we questioned whether increasing cellular diversity would exacerbate the disruptions in maturation and organoid size previously identified. By aging the UNOs to four months, we saw similar increases in organoid size, nuclei diameter, and a decrease in cell density that was consistent with the D40 timepoint. We were unable to characterize the rosette dynamics

at this timepoint, however, as the progenitor pool that occupies these structures reduces over time to mimic the thinning VZ of the aging brain. Using scRNA Seq we identified a variety of post mitotic cell types including excitatory neurons from multiple brain regions, inhibitory neurons, as well as their respective progenitors and radial glia. We also identified a considerable degree of variability between familial groups.

Initial comparisons between the six D120 scRNA Seq samples identified a considerable degree of heterogeneity that was mostly driven by genetic ancestry (familial background). Genetic variation is a consequence of using non-isogenic iPSC lines, the heterogeneity of which can be exacerbated in mature organoid models that contain diverse cell types. We regressed out the ancestral variable on the merged dataset to mitigate these differences but found the cell proportions to be drastically different due to residual variation. In particular, there were differences between family 2 *vs.* families 3 and 5 by cell proportion (**Supplementary Fig. 3**). With this in mind, we opted to perform analyses that focused on individual cell compositions rather than population dynamics (cell proportions), reasoning that any change that emerges despite the variability should be driven primarily by loss of the 15q13.3 genes.

We used CellChat to examine the intercellular communication at this later timepoint and quantified a decrease in the number and strength of statistically significant global interactions amongst the aggregate datasets. In this way, broad intercellular communication is predicted to be reduced in our UNO model. To identify causative pathways, we ranked the overrepresented signaling pathways in aggregate tissue and found many of the terms to be related to cell adhesion, including the top pathway, nectin signaling. This ranking considers the absolute value of the predicted signaling network,

and after probing for Nectin-1 mRNA in aggregate tissue (**Supplementary Fig. 5**) we concluded the change was likely a reduction in signaling.

Reduced cell adhesion pathways were then confirmed by REVIGO analysis, where cell-cell adhesion was a core term for the downregulated gene sets within (cycling) radial glia, intermediate progenitors, Cajal-Retzius cells, non-telencephalic hindbrain neurons, and excitatory mature neurons. Nectin signaling crosstalk showed a general reduction in the interaction strengths amongst multiple cell types, with an appreciable decrease in the interaction strength (thinning of the interaction ribbon) of the cycling radial glia.

Nectins are calcium-independent immunoglobin-like molecules involved in cellular adhesion. Their expression is not limited to neural tissue, but they have been well documented in the nervous system for their facilitation and maintenance of neural synapses.[185] Within the cortex, nectin expression is particularly enriched in radial glia cells,[186] where they are localized to the adherens junctions (AJs) that facilitate contact between adjacent cells. Nectins interact with the adaptor protein afadin as well as cadherin-catenins machinery to stabilize adherens junctions through interacting with their actin-rich cytoskeleton. Disruptions to either of these protein families cause disruptions to the AJs, the ventricular zones that they populate (including dispersion of the progenitor cells) and later cortical malformations in mouse models such as hydrocephaly and lissencephaly[187–189]. AJs and the adhesion proteins that populate them are therefore crucial for the structural integrity to cortical structures and cell localization[190]. It is possible that disruptions in nectin signaling between radial glia contribute towards their

change in cell density within the rosette structures, causing them to be spread out and larger in organoid size.

Nectin signaling was also disrupted in the inhibitory neuron population, in that both incoming and outcoming statistically significant (one-sided permutation test, p-adj < 0.05) predicted interactions were not detectable. This suggests that 15q13.3del interneurons may be particularly vulnerable to changes in nectin and possibly other adhesion signaling differences. Nectin proteins play a crucial role in early circuit development, where they localized to puncta adherentia junctions (PAJs) which connect pre and post synapses in neurons. Disruptions in nectin protein expression impacts the number and size of dendritic spines in cortical neurons[185] and can destabilize connections between axons and dendrites of neighbouring cells,[191,192] therefore affecting the architecture, connectivity, and functionality of neurons.

Nectin proteins have been shown to support inhibitory neuron maturation by establishing modes of plasticity in early development and are upregulated during periods of sensory deprivation.[193] In the primary visual cortex, reduction in the NECL2/CADM2 protein (which is significantly reduced in the 15q13.3del inhibitory neurons) disrupts synapse maturation between thalamocortical neurons and PV-interneurons through reduction of SST input onto the excitatory neurons, suggesting that disruptions to these genes may impact the capacity of interneurons to regulate microcircuitry.[193]

Amongst the cadherin-related signaling pathways, CellChat predicted dysregulation of both L1CAM and NCAM in the D120 aggregate data, and so it is possible that there are more neural-specific cell adhesions mechanisms in addition to

nectin signaling. Indeed, mRNA expression of adhesion proteins NRCAM and NCAM1 were altered in aggregate tissue and amongst the neurons and radial glia populations (**Supplementary Fig. 6**). Of note, *L1CAM* mutations have been associated with impairments in callosal axonal projections, which can result in irregular corpus callosum structures as seen in multiple 15q13.3 patients in our cohort[194] as well as in axon the projection capacity of 15q13.3del 2D NGN2 neurons (**Supplementary Fig. 4**). This suggests that a common pathophysiology could underlie structural changes both in organoids and in patient neural structures.

In addition to reductions in adhesion protein signaling, we noticed several axon guidance pathways affected in the over representation plot, including semaphorins 3-6 and the ephrin molecules ephrin-A (EphA) and ephrin-B (EphB), the last of which were the 2nd and 10th most over-represented pathways by rank similarity, respectively (**Supplementary Fig. 7**). Ephrins are ligands that bind to the Eph receptors, a membrane-bound family of receptor tyrosine kinases that regulate diverse cellular processes including cell migration, localization, axon guidance, synapse formation, and cell adhesion.

While most 15q13.3del UNO cell types displayed a strong reduction in ephrin-B signaling, it was interesting to see that the outwards signaling from cycling radial glia and excitatory neurons was unaffected, suggesting that this signaling pathway may not be globally disrupted but rather in a cell type specific manner.

From a cell adhesion context, Ephrins exhibit considerable crosstalk with cell adhesion receptors such as integrins and cadherins[195,196] to control cell functions

including the polymerization of the actin cytoskeleton and ECM adhesion. *Efnb1* KO mice exhibit morphological changes in their neuroepithelium due to aberrant tissue folding through what are likely disruptions to cell cohesion,[197] and loss of Efnb1 is thought to be required for the maintenance of apical adhesion in apical progenitor cells through localization of integrin β1.[197] Ephrins therefore play a role in maintaining the integrity of cortical structures, and it is possible this extends to changes in organoid size.

Ephrins are better characterized for their contributions to axon guidance, where they primarily act as repulsive elements to axonal growth cones during early neurodevelopment. Disruptions to the EphrinA and Bs results in aberrant cell localization and inappropriate axonal projections, leading to miswiring of circuitry. Indeed, functional deficits are observed in electrophysiology experiments, suggesting that ephrins also have broad impacts on neuronal communication.[197–199]

The last cell adhesion signaling pathway that we examined was the neurexin-neuroligin (NRXN-NLGN) signaling pathway, which are groups of synaptic cell adhesion molecules that help join presynaptic and postsynaptic compartments between neurons. We observed an inverse expression of NRXN1 between the excitatory mature and inhibitory neurons, suggesting that cell adhesion may underlie changes in cell communication and possibly imbalances to E/I circuitry. Substantial reductions in NRXN-NLGN were then observed between the inhibitory neuron population and each cell type, with homotypic interactions presenting the largest observable decrease.

NRXN-NLGN signaling plays a crucial role for neural development and synaptic plasticity; perturbations to ligands from either family result in pre and postsynaptic

signaling deficits that occur in a cell type- and brain region specific-manner,[200] a phenomenon that has been confirmed in *NRXN1* KO human UNOs.[144] KO studies of NRXNs have shown profound deficits in interneuron populations, including alterations to dendritic spine densities, presynaptic release, and action potential-induced presynaptic calcium influx.[201,202] In addition, NRXNs directly interact with $GABA_A$ receptor subunits, independent of NLGNs, to alter GABAergic synaptic transmission[202] which in turn can influence the maturation of glutamatergic excitatory synapses. It is clear that NRXNs and NLGNs regulate interneuron activity to some degree, and that their loss generally leads to disruptions in synaptic transmission. The exact regulatory nature is highly specific to the interneuron subpopulation, however, and so it is difficult to fully appreciate how the UNO interneurons (whose subpopulations remain ambiguous) are affected by this disruption. In addition, this information should be heavily scrutinized as there is a profound diversity in NRXN1 protein isoforms, their synaptic localization, and their combinatorial interactions with various postsynaptic ligands. This diversity has been speculated to create a synaptic cell adhesion "combinatorial code" that may be a master regulator of synaptic diversity in the CNS. In this way it is difficult to infer the exact consequences of altered NRXN-NLGN signaling given broad mRNA expression alone.[203,204]

Taken together, adhesion appears to be broadly disrupted in the 15q13.3del UNOs. Due to their importance in establishing physical connections between cells (the AJs of neighboring radial glia cells, or broadly amongst synapses and neural circuits) changes in nectin or neural CAMs may alter cell-cell proximity and contribute towards cytoarchitectural changes that could include increases in organoid size. It should be

stressed, however, that the above interactions are inferred from the composition of ligand-receptor complexes rather than their mRNA expression in isolation. For this reason, not every constituent of the nectin signaling pathway was significantly reduced in the D120 cell populations, and so validations must be performed functionally rather than just on transcript or protein expression alone.

5.2 Trajectory modules of excitatory and inhibitory neurons display inverse synaptic signaling trends.

Due to the importance of synaptic input (and cell communication broadly) for neuronal maturation, we decided to profile the developmental trajectory of each cell type as well as the aggregate tissue using pseudotime analysis. We found significantly altered distributions in the aggregate cell densities across pseudotime, which resembled a premature emergence of early cell types, followed by a delay of intermediate and late cell types. We then plotted the pseudotemporal distribution of each cell type individually and found all but the mesenchymal-like cells to be significantly altered, with **Table 3** ranking the cell types from most (non-telencephalic hindbrain excitatory neurons) to least (inhibitory neurons).

The lack of change amongst mesenchymal-like cell types is consistent with its presence as a non-ectodermal contaminant (providing mostly structural integrity to the ECM) and suggests that the pseudotemporal changes are specific to neural tissue.

We chose to focus on the trajectory of the excitatory mature neurons, as they are one of the last cell types to emerge and may therefore have the most severe functional consequences in response to pseudotime alterations. Amongst the neural population, we

found enrichment and high specificity to modules pertaining to synaptic activity, (**Supplementary Fig. 8A, Fig. 17**), including terms like, “chemical synaptic transmission”, “trans-synaptic signaling”, and “synapse organization”. After classifying the module broadly as *Synaptic signaling*, we tested the aggregate genes within this module and found a significant increase amongst the excitatory neuron population. Significantly increased genes included three SFARI risk genes: *NRXN1*, a synaptic adhesion model whose overexpression is tied to the emergence of other synaptic genes,[205] *SYNGAP1*, a gene involved in synaptic signaling transmission that is a significant risk factor for global developmental delay, ASD, and epileptic encephalopathy, and *CANCNAG1*, a calcium voltage-gated channel highly enriched in epileptic disorders such as Dravet syndrome, where gain of function mutations results in hyperexcitability,[206,207] and hypersynchronicity[206] in a novel cortico-thalamic assembloid model. Each of these genes represent important modulators of synaptic function, and their increase in the excitatory mature neuron populations may be recapitulating excessive activity seen in epileptic 15q13.3 patients.

In contrast to the increase in synaptic signaling pathways amongst excitatory neurons, the inhibitory neurons showed the opposite trend, with a decrease in the highly specific GABAergic signaling module. Amongst the effected genes included *GAD1*, which encodes the glutamate decarboxylase 1 enzyme responsible for GABA production; *SLC6A1*, which is responsible for its transportation across the synapse, and *ERBB4*, a prominent neuropsychiatric risk gene and RTK protein that binds to the growth factor NRG1 to regulate tangential interneuron migration and circuit formation.[208] Importantly, ERBB4-NRG1 signaling in PV interneurons regulates excitatory pyramidal neuron

signaling by modulating long-term potentiation and synaptic plasticity, [209,210] and it is therefore possible that reductions in inhibitory transmission contribute to or exacerbate the excessive synaptic signaling genes enriched in the excitatory neurons. Alterations in E/I circuitry are consistent with the epileptic phenotypes of 15q13.3 patients and mouse models, where absence seizures are the most common epileptic presentation. Interestingly, reductions in the *ERBB4* target and gamma-aminobutyric acid A ($GABA_A$) receptor α1 subunit *GABARα1* produce cortical absence epileptiform activities in mice,[211] suggesting that interneuron-specific alterations in *ERBB4* can produce epileptic phenotypes consistent with 15q13.3 models. This subunit has previously shown dysfunction in the 15q13.3del background, where cortical pyramidal cells show reduced sensitivity to $GABA_AR$ antagonism in mouse models.[212]

In addition to absence seizures, there is also convergent dysregulation of gamma wave oscillations in 15q13.3 models, which are a type of high-frequency network activity associated with schizophrenia that depends on synchronization of local circuits by PV interneurons. In 15q13.3 patients and mice, there are increased gamma oscillations during resting state but reduced evoked gamma oscillations in response to auditory stimulation, similar to schizophrenic populations[116] and suggestive of disrupted inhibitory transmission. Gamma wave oscillations have been shown to be augmented by *NRG1 in vitro* in an *ERBB4*-dependent manner,[213] further implicated *ERBBR4/NRG1* signaling and interneuron dysfunction in the 15q13.3 microdeletion syndrome.

While most of the literature has focused on the excitatory cortical neurons in 15q13.3del mouse models, a small group of studies have profiled the interneuron progenitors and mature cell types at early and later stages of development, respectively.

Reductions in the firing rate of parvalbumin (PV) fast-spiking interneurons have been identified in the prefrontal cortex of adult mice,[214] and a year later this population was reported to be decreased in both the prelimbic and motor cortex of mature (10-11 week) 15q13.3del mice, using an improved tissue clearing technique that had higher sensitivity for detecting rare interneuron subpopulations.[119] These studies confirmed reductions in both PV interneuron population proportions and maturation. Based on the transcriptome of the interneuron populations in the UNO dataset, it is most likely that they are immature progenitors, with high expression of canonical progenitor markers (*GAD1/2*, *DLX1/2/5/6*) and low subtype-specific markers (*SST, PVALB, CCK/VIP, NPY, HTR3A, TH*). Despite a lack of parvalbumin transcript signal, markers enriched in PV interneurons, including *ERBB4*, were present, suggesting that these progenitors have the potential to become PV interneurons and may later become disrupted in the developing brain. It is also well documented that adhesion proteins are essential modulators of synapse formation, with human cKO models of *L1CAM* showing decreased activity-dependent sodium currents and impairments in action potential generation in a seemingly cell autonomous manner.[215]

Reduction in firing capacities may also be the product of improper axonal guidance or cell migration, both of which were highly enriched GO terms for KEGG and biological pathways, respectively, in the interneuron DEGs. Failure of a neuron to reach its targeted destination in the brain will result in changes to its microenvironment, including synaptic input from neighbouring cells that would normally facilitate its maturation.[10] In this way, there are likely multiple contributors to the predicted changes in synaptic activity predicted by pseudotime analysis.

5.3 Summary

This project provides the first indication of early neurodevelopmental impairments in a human model system of the 15q13.3 microdeletion syndrome. We leveraged an unbiased organoid system to recapitulate key developmental milestones such as neurogenesis and cell type emergence. We identified early disruptions in radial glia at a timepoint that represents early mid-gestation, profiling structural changes in the ventricular-like zones where neurogenesis occurs. We then exploited the model's capacity for cell diversity by maturing it to a timepoint resembling late mid-gestion. Only at this timepoint did we observe broad changes in cell adhesion using three separate analysis pipelines: gene ontology/REVIGO, CellChat, and pseudotime trajectory. We also identified key changes in intercellular communication both in aggregate and between individual cell populations. We then used pseudotime analysis to corroborate signaling changes, by revealing excitatory/inhibitory-specific disruptions in synaptic signaling modules and key regulators of synaptic plasticity. Taken together, this chapter represents an exploratory examination of neurodevelopment in the 15q13.3 microdeletion background. Using this approach, we have made the argument that cell type emergence and maturation is disrupted at multiple timepoints, and that imbalances in excitatory and inhibitory neurons can be modeled using UNOs.

5.4 Caveats to scRNA approaches in unguided models

One of the unique features of unguided neural organoids is the spontaneous nature in which differentiation occurs. Using intrinsic signaling cues, differentiation is achieved with minimal use of small molecules or patterning factors. This enables unbiased

profiling of neural tissue and cell types but generates a considerable degree of heterogeneity in the system. As a result, batch- and line-driven variability can make disease modeling and reproducibility an issue. Even organoids of the same batch can produce variable numbers of rosette structures, which over time will influence the cell type composition between organoids. In this way we observed the emergence of non-telencephalic cell types from the midbrain and hindbrain, which introduce both diversity and heterogeneity into this generally telencephalic model.[216] For this reason, we ensured that multiple organoids and patient lines were used for all experiments using the UNO model.

Indeed, over time we observed an increase in inter-line variability between the three organoids profiled for single cell RNA sequencing. The D40 dataset showed relatively consistent proportions between genotypes, but the D120 population proportions were inconsistent between controls and therefore hindered population proportion analyses at this timepoint. Instead, cell dynamics such as intercellular communication and cellular trajectory were assessed using CellChat and Pseudotime analysis, respectively. These analyses revealed novel disruptions in inferred cellular communication and differentiation, which helped shape the use of guided neural organoids in the fourth chapter. It is imperative that these results are assessed critically, however, as they are predictive in nature and can be influenced by factors such as non-biological zeroes (technical noise and under-sampling) that can bias estimations of gene expression correlations.[217,218] We mitigated the effect of this variable through regression of cells with low transcript, however, it does not account for one-off genes with low expression.

In addition, we are exclusively profiling the transcriptome of single cells, which does not consider essential epigenetic and post translational modifiers that regulate protein expression and function. This assessment therefore provides a window into the gene expression changes in our disease model, but requires robust validations on the protein, cellular, and functional levels.

5.5 Overview of 15q13.3del guided neural organoid phenotyping.

Imbalances in excitatory/inhibitory circuits underlie multiple NDDs and are a clear candidate for the pathobiology underlying the high incidences of epilepsy amongst 15q13.3 microdeletion patients. Despite how essential E/I circuit development and balance is for the developing brain, the underlying biological events are poorly characterized in a human context, and much less so with disease models.[219,220] Only recently have we gained insight into human-specific circuit-regulation for processes such as our increased cortico-cortical connectivity,[175] although much remains uncharacterized.

In chapter 4 we examined E/I circuit development in a more consistent model by taking advantage of the guided neural organoid approach to generate forebrain organoids. These models faithfully produce the mature excitatory and inhibitory neurons that populate the dorsal and ventral forebrain, respectively, and can be combined to generate assembloids for circuit phenotyping. Using forebrain organoids from our entire 15q13.3del cohort, we probed the transcriptome of early (day 50) and late (day 100) dorsal and ventral organoids to better understand what shared and distinct biological pathways underpin the microdeletion syndrome. Amongst the entire cohort we found

persistent dysregulation of the transcriptome in all four datasets, with particular enrichment for zinc finger proteins (ZNFs) localized to chromosome 19p13.2 and 19q13.43. This disruption was most pronounced in the dorsal forebrain dataset, which additionally showed enrichment for chromosomal dysfunction and DNA repair pathways through GSEA. In contrast, the ventral dataset had reductions in neural-specific gene sets including neurogenesis, cell adhesion, and cell migration. Using the neurogenesis pathway as a prompt, we quantified a reduction in the SST interneurons expression that is predominant in at this timepoint. We later examined cell circuitry dynamics in a novel dorsal-ventral assembloid model using multiple lineage and genotype combinations. In this model we identified disruptions in cell migration that is suggestive of a cell autonomous mechanism, as well as morphological changes in the migrating neurons that revealed possible contributions of the microenvironment to cell maturation.

5.6 Broad transcriptional dysregulation in dorsal forebrain organoid may be a product of DNA damage

We first characterized the dorsal forebrain datasets due to the availability of transcriptomic data from 15q13.3del mouse cortical neurons, as well as the numerous studies identifying abnormalities in excitatory cortical neuron development in 15q13.3 and other NDD models.[117,118,221–223] When comparing to a mouse transcriptomic dataset previously generated in our lab, we found a considerable overlap of common DEGs amongst the embryonic day 16 timepoint (30%), as well as postnatal day 21 (24%) and 65 (>50%) when compared to the dorsal DEGs from D50 and D100 (**Table 6**). This gave us confidence that some aspect of the mouse biology was recapitulated in our model system.

Table 7: List of genes shared between dorsal forebrain organoid dataset with pAdj <0.05 and genes identified in the *Uddin and Unda et al., 2018* Df(h15q13)/+ time course study.

	Timepoint		
	E16	**P21**	**Adult**
	Cyr61	Th	Folr1
	Sapcd1	Fam196b	Kcnj13
	Mtmr10	Mtmr10	Cldn2
	Chrna7	Fan1	Zic4
	Fan1	Chrna7	Col8a2
	Klf13	Otud7a	Aldh1a3
	Xist	Klf13	Mtmr10
	Otud7a		Otud7a
			Klf13
			Fan1
			Chrna7
% shared	**30.76923**	**24.13793**	**52.38095**

We noticed, however, that multiple genes were not reported in 15q13.3del mouse transcriptomics; in particular, the ZNF proteins that encompassed nearly half of the top downregulated genes in the D50 and D100 dorsal datasets have not been reported in 15q13.3 literature. Upon further examination of the DEGs, it became apparent that many of the ZNFs were primate-specific (as many as 35%) and therefore could not be identified in a mouse background. Indeed, ZNFs, and in particular those with a Krüppel-box like

domain (KZNFs), have undergone rapid expansion during early primate and hominid evolution through repeated cycles of segmental duplication to create KZNFs with unique genomic targets and molecular functions.[224] Primate-specific KZNFs have been found to participate in gene regulatory networks in the brain and have expression patterns that are distinct to humans.[225,226] It is for these reasons that KZNFs are candidate genetic drivers to the early neocortical expansion that occurred in our hominid ancestors but not in primates. In addition, they represent the largest family of human transcription factors, and so it is possible that reductions or dysregulation to these gene clusters or a subset could affect higher level processing and cognition associated with human brain evolution.[170,227] Since cognitive deficits are heavily associated with NDDs, this warrants further investigation of this poorly characterized family of proteins.

Chromosome 19 is heavily enriched for KZNFs, with the short arm of chr19p13.11–p12 containing as many as 43 unique KZNF genes.[228] *ZNF558* is a gene within this region and is the 5th and 2nd most downregulated gene in the D50 and D100 datasets, respectively. This gene was recently shown to have human-specific expression in forebrain neural progenitor cells and to be a crucial regulator to early neurodevelopmental events such neuronal maturation and projection development in an UNO KD model.[170] We explored expression of the sole protein-coding target to *ZNF558*, *SPATA18*, and found its mRNA expression to be significantly upregulated, in line with reduced function of the ZNF protein. These findings show that 15q13.3del dorsal forebrain organoids produce broad disruption to KZNF proteins, including known regulators of early neurodevelopment. Disruption on the transcriptional level was further bolstered by the GO enrichment for *Transcription regulation* shared by both timepoints.

Specifically, *RNA Polymerase II-mediated transcription* was enriched, suggesting that the downstream targets of the transcription factors are protein coding genes.

It is clear from the available literature that KZNFs are poorly characterized, due in part to their evolutionary divergence, which requires human model systems for proper study. Even well annotated databases such as STRING could not identify a convergent pathway amongst the affected ZNFs, instead classifying the group broadly as transcriptional regulators. It is very possible that many of the ZNFs have unique and possibly development-regulating functions, however, until they are modeled in the appropriate [human] system, their functions remain ambiguous.

5.6.1 Excitatory neurons as a vulnerable cell type to genomic instability.

To determine the causative 15q13.3 gene contributing towards downregulation of chromosome 19 KZNFs, we examined publicly available chromatin immunoprecipitation sequencing (ChIP-Seq) datasets on the ENCODE repository to assess potential binding sites of the 15q13.3 gene and transcription factor, *KLF13*. Using datasets from two unique cell types, we found consistent peak enrichment at a site upstream to the *ZNF558* promoter, suggesting that *KLF13* may have a regulatory role to this and other downstream KZNF genes in the region. This is purely speculative, however, given that the experiments were done in non-neural cell lines and that neurons themselves express highly specific and tightly regulated gene regulatory networks.[229] *KLF13* has also been classified predominantly as a transcriptional repressor, associating with chromatin at the proximal promoter sites to its downstream genetic targets to reduce their expression, and so the reduction of target transcripts in its absence is somewhat unexpected. While it is

certainly possible that *KLF13* has diverse transcriptional regulation, further validation of its binding and regulation of chromosome 19 ZNFs would be required before speculating how it contributes towards changes in transcription.

The contributions of KZNFs to neuron gene regulatory networks has only recently been explored in the field of genetics,[225,226,230] where multiple genes have human-specific expression due to their evolutionary divergence of through segmental duplications. KZNFs have been found to co-evolve with transposable element (TE) regulatory sequences, where they broadly act as repressors to TEs to maintain genomic stability and ensure proper gene expression.[230]

Genomic instability and DNA damage were GSEA terms that were consistent between – and exclusive to – the dorsal forebrain datasets, producing enrichment for gene sets involving *Stress-induced ATR repair*, *Diseases of DNA repair*, *Replication fork processing*, and *Increased resolution of D-Loop structures*. Of note, there was no enrichment in the ventral bulk RNA seq datasets for DNA instability or damage, suggesting that this effect may be cell type- or lineage-specific. We questioned if this phenotype would be consistent amongst the excitatory neural populations within the unguided neural organoids, as they are the most comparable cell type by linage. After examining the significantly upregulated GO terms, we noticed that the excitatory neurons and radial glia populations at the D40 timepoint had exclusive enrichment in DNA damage and repair through GO analysis as well as in semantic similarity reduction, which combines similar terms to produce the most pertinent GO term. DNA damage was also listed amongst their downregulated GO terms, but it was specifically *Cellular response to DNA damage stimulus*. In contrast, the D120 had no enrichment for DNA damage

amongst the inhibitory neurons or other cell types, suggesting that DNA instability is an early event that is not present in inhibitory neurons. The exclusive presence of DNA instability amongst radial glia and predominantly excitatory neurons from UNOs and dorsal organoids suggests the cell type may confer vulnerabilities to cell stress and/or DNA repair.

DNA repair mechanisms are an intricate and essential part for cell survival. DNA damage can occur as a result of chemical modification to the nucleotide base content, singe/double stranded breaks (DSB) to the genome, as well as replication errors or stalling of the replication fork. Disruptions to any of these processes can result in destabilization of the cell genome, with DSBs predicted to be the most harmful.[231,232] Following DNA damage, a cell will initiate one of several repair mechanisms, including base excision repair (BER), nucleotide excision repair (NER), mismatch repair (MMR), homologous recombination (HR) and non-homologous end joining (NHEJ).

Upon examining the genes enriched in the DNA repair terms, we identified those involved with double stranded break repair (*RAD50* and *WRN*) and homologous recombination repair proteins (*XRCC2* and *RAD51*). It is therefore likely that the affected DNA repair is of one or both mechanisms, especially given the additional GSEA term, *Homology directed repair*. Homologous recombination-directed repair often initiates the formation of displacement-loops, or D-loops, which are intermediate structures that enable physical contact between template DNA and a single stranded DNA molecule between sister chromatid of homologous chromosomes. Following synthesis of the new strand, the D-Loops are resolved often through nucleation.[97,233] *D-Loop structure resolution* was another GSEA term amongst the dorsal datasets, which was of interest to

us given the ability of the 15q13.3 gene and DNA nuclease *FAN1* to cleave D-loop structures during HR. Too few or prolonged participation of D-loop structures can prevent HR initiation and completion, respectively, ultimately hindering DNA repair.[234]

DNA damage has previously been reported in a human 15q13.3del model, which used NGN2-overexpression to create excitatory-like glutamatergic neurons from multiple 15q13.3del patient-derived iPSCs. Following bulk RNA sequencing, the group found enrichment in terms related to damaged DNA among the patient lines, and more specifically DSB repair amongst isogenic KO lines of *FAN1* and *MTMR10* exclusively, further suggesting the presence of DNA instability in the 15q13.3 model and the pleiotropic contributions from genes within the CNV. By using a purely excitatory neuron model, this study corroborates our finding that DNA instability is detectable in excitatory neural backgrounds.

While *MTMR10* has not been previously associated with DNA instability or damage, other genes within its family (*MTMR2* and *MTMR13*) have been identified in unbiased DNA damage screens,[235,236] highlighting the possibility of a new role for the gene. In contrast, *FAN1* has been directly associated with DNA repair, promoting replication fork recovery to promote genome stability.[237] *FAN1* variants and KO models show an increased susceptibility to toxic interstrand crosslinking, genomic instability, and repeat instability in expansion diseases such as Huntington's disease and Fragile X syndrome.[99,238,239] Studies have associated *FAN1* with the repair of interstrand crosslink DNA damage. These DNA lesions prevent DNA strand separation and are commonly repaired by HR directed by the Fanconi anemia (FA) pathway, which relies on FAN1 to

identify DNA damage by using its Ubiquitin-binding zinc finger domain and prepare the damage for repair by degrading ssDNA, exposed dsDNA ends, or ss-dsDNA junctions.[98]

DNA replication and stress can directly impact the degree and efficacy of transcription within a cell[240] and was recently identified in a human NPC model of ASD. In this study they inhibited DNA replication to induce cellular stress and performed morphological assays and bulk RNA sequencing of ASD-derived NPCs. They found patient-derived NPCs to have higher rates of cellular division, asymmetric replication fork stalling, and to be more sensitive to replication-induced stress with increased rates of DNA damage.[241] The consequences to this DNA damage had an interesting neural-specific effect, introducing DSBs disproportionately in long neural genes that are enriched in NDD populations. Amongst these genes were multiple protocadherins, which they showed produced a functional change in the adherens junctions and the apical-basal polarity of the NPC rosettes. Amongst the 33 genes reported, our collective forebrain datasets shared over 55%, including: *AUTS2, ERBB4, PCDH15, PCDH11X, CTNNA3, NRG3, RBFOX1, DCC, CTNNA2, MACROD2, ERC2, LSAMP, LPP, CTNND2, CDH18, SDK1, MAGI2, EXOC4,* and *IL1RAPL1*. This striking coverage of genes sensitive to DSB, combined with the transcriptional dysregulation amongst the forebrain downregulated genes, highly suggests that DNA or replication stress may underly some of the changes we see on the mRNA or protein level. Combining potential sensitivity to DSB, followed by impairments to DNA repair pathways and genes that specialize in DSB repair (*FAN1*), it is highly possible that 15q13.3del excitatory neural neurons have genomic instability that affects their maturation and health.

One interesting element of DNA damage enrichment is its cell-type specificity, occurring exclusively in excitatory neurons or their precursors in both UNO and guided forebrain tissue. It is possible the DNA damage is a stress response brought on by reactive oxygen species (ROS) that can be present in metabolically demanding cell types.[242,243] Indeed, excitotoxicity occurs in neurons that participate in seizures, where an increase in signaling (which pseudotime analysis suggested was present in the mature excitatory neuron) can result in DNA damage.[244] This phenomenon has been mostly studied in glutamate-induced excitotoxicity, which is the principal neurotransmitter of excitatory neurons specifically.[245,246] It is therefore possible that hyperactivity of the excitatory neurons is what underlies their vulnerability to DNA damage and is exacerbated by disruptions in long-neural genes.

5.7 Persistent deficits in early neurogenesis may emerge from alterations in radial glial trajectories.

Genes associated with neurogenesis were consistently downregulated in both guided and unguided models, suggesting deficits in neurogenesis across cell types and time in 15q13.3del neuron populations. We confirmed this disruption by quantifying a reduction in $TBR1^+$ newborn neurons at an early timepoint in both tissue type (D40/D50). To examine neurogenesis in the ventral forebrain organoids, we probed for somatostatin (SST), a neuropeptide and marker of the predominant SST interneuron population in the organoid. We found that both the SST mRNA and average protein expression was reduced in the 15q13.3del ventral organoids, in addition to the proportion of SST^+ nuclei, suggesting that changes in SST interneuron activity may be affected in addition to the proportions of this interneuron subpopulation.

Changes in neural population proportions are not a frequent occurrence in 15q13.3 models; in mice there are reductions in PV^+ interneurons in the prelimbic, motor, and somatosensory cortex, astrocytes in the medial prefrontal cortex,[117,119] and no reports of proportion changes in excitatory neurons at present. These changes are all accompanied by synaptic deficits to the local circuits, suggesting that neurogenesis as well as neuronal maturation is altered in this system.

Cell differentiation and maturation can be inferred from changes in pseudotime, which plot the relative emergence of cell types individually or in aggregate. Our analysis predicted significant changes to the pseudotemporal distribution of the aggregate tissue as well as in individual cell types, suggesting that a common mechanism may underlie changes in cellular trajectories. Specifically, the pseudotime density plot highlighted a premature emergence of multiple cell types that was generally sustained for a shorter period of time, which was indicated by the sharp peak (indicating cell type emergence) followed by a decline. This was most evident in the dividing cell types such as the radial glia, which had the second most significant difference in pseudotime distribution. Radial glia provide the progenitor pool for newborn neurons, and their maintenance in cell division is important in early development to ensure neural populations are generated in appropriate proportions. It is possible that changes in neurogenesis or cell type emergence are due to early depletions of NPC populations such as radial glia, which are not sustained throughout time and fail to produce sufficient proportions of neurons. By generating both radial glia and neurons prematurely (shown by an increase in the proportion of these two cell types in **Fig. 8**), the developmental window that is normally protracted is instead shortened, causing disruptions to future newborn neurons, which in

our models would be TBR1$^+$ neurons within the rosette structures. Altered ratios of neurons: NPCs are typically accompanied by changes in proliferation,[247] which were not observed at the D40 timepoint but was not assessed at D120. For these reasons, there is a disconnect between our findings of unchanged proliferation, *vs*. a decreased proportion of newborn neurons.

It is important to note that TBR1 mRNA transcript was not significantly decreased in the scRNA seq of the UNO populations, although the detection of TBR1 mRNA was minimal compared to other neuronal markers (**Supplementary Fig. 2-3 dot plots**), and it is possible the chosen read depth of 50 000 was not sensitive enough to pick up differences in transcript expression. Likewise, TBR1 mRNA were not significantly altered in the bulk RNA seq dataset. This dataset is derived from aggregate tissue, however, which is prone to heterogeneity and dilution of cell-type specific markers. Instead, the protein quantification of TBR1 was restricted the SOX2$^+$ rosette structures to try and mitigate the inherent variability of the system. This validation was only performed on a single family, however, and would need to be repeated across the cohort in order to have implications towards the CNV at large.

5.8 Ventral forebrain tissue possesses unique disruptions in migration pathways and temporal shifts in axonal protein expression.

Due to the minimal literature on the contribution of inhibitory neurons in NDDs, and in particular models of the 15q13.3 microdeletion syndrome, we questioned how early 3D MGE-derived interneuron development would be impacted by the 15q13.3 microdeletion across time. Preliminary data from the D120 UNO scRNA seq suggested cell type

emergence and synaptic activity may be affected, but the data was derived from a heterogeneous system that lacks interneuron-specific cell niches (such as the ganglionic eminence) that drive fate interneuron fate specification.

Previous research has implicated PV interneurons to be the primary interneuron subtype effected by the 15q13.3 microdeletion, with reports of reduced cell density and neural activity within the cortex. SST interneurons, which precede the PV interneurons developmentally, are the principal neural cell type in ventral organoids from developmental days 25-200, after which PV interneurons begin to emerge. This experiment therefore offered new insights into how the 15q13.3 microdeletion may shape an early and previously unexplored interneuron subtype.

15q.13.3del ventral organoids, like their dorsal counterpart, showed a significant enrichment for transcription factors amongst the downregulated genes at both timepoints, although to a lesser extend by both fold enrichment and number of transcription factors involved. Instead, the D50 and D100 datasets had more pronounced and convergent enrichment for reductions in cell adhesion, both by GO term analysis and GSEA approaches. Adhesion proteins in the D50 dataset included a variety of cadherins, protocadherins, desmosome proteins, collages, and ECM constituents to suggest broad changes in cell adhesion at this timepoint. Conversely, the D100 dataset had a lower enrichment for adhesion proteins and only featured integrins and some ECM constituents in the enriched gene list. Cell migration was exclusively enriched at the early D50 timepoint, including neural CAMs genes like *CHL1*, transcription factor and GABAergic neuron regulator and migration regulator, *SOX14*, and multiple signaling proteins like Sonic hedgehog (*Shh*) and Ras Homolog Family Member H (*RHOH*).[248–250] The size and

diversity in the gene set list made it difficult to predict how specifically migration would be impaired (radial, tangential, neural placement, etc.) but was of interest as interneuron migration had never been characterized in the 15q13.3 microdeletion syndrome.

Interneuron migration is crucial for circuit development in the developing CNS and given the enrichment for migratory genes in the D50 downregulated gene set, it is possible the D100 will reflect the consequences to this deficit. We noticed a unique enrichment for G-Coupled protein receptor (G-PCR) signaling amongst the D100 downregulated genes, which represent a large family of proteins that orchestrate synaptic signaling and plasticity between neurons through detection of external stimulus and initiation of downstream signaling. Since interneuron-specific signaling receptors such as a GABA(B) fall in the G-PCR family, we screened for GABA receptors and found downregulation of multiple G-PCR genes, including *GABARG1, GABRE,* and *GABBR2*. This disruption in GABA receptor expression suggests that broader synaptic dysfunction may be occurring in the ventral forebrain organoids. which is consistent with the pseudotemporal enrichment of the inhibitory neurons for GABA transmission.

To understand developmental-dependent transcriptional changes, we assessed for genes that were overexpressed in early ventral development, and then later become downregulated at the D100 timepoint. Of the 33 genes that change with the developmental trajectory, 12% were enriched for axon development and ensheathment (myelination), including the myelin basic protein (*MBP*). This further supports the notion that D100 ventral organoids have exacerbated signaling deficits compared to the immature D50 timepoint, as myelin is essential for the propagation and conduction of action potentials. Amongst interneurons, however, it is typically PV interneurons that

receive myelination rather than SST interneurons.[251,252] This change could be due to alterations in neural fate amongst the 15q13.3del organoids, which is supported by altered trajectory of the calbindin-1 gene (*CALB1*), which marks a subtype of interneurons. The presence of oligodendrocytes or OPCs in a ventral forebrain model is perplexing, however, as they are thought to emerge primarily from cortical VZs rather than the ganglionic eminence, although this is an evolving perspective.[253]

Early 15q13.3 CNV fetal interneuron development has only been characterized in two studies: one using the 15q13.3del mouse model and another in 15q13.3dup mixed inhibitory-excitatory pseudo 3D cultures. In the mouse study, interneuron progenitor proliferation was assessed in the medial and caudal ganglionic eminences in 15q13.3 and *Klf13*$^{+/-}$ KO mice at embryonic days 13.5-17.5. Using these embryonic timepoints they found reductions in the proliferation of medial and caudal ganglionic eminence progenitors at discrete and non-overlapping timepoints, which was recapitulated in the *Klf13* heterozygous KO mouse. *Klf13* was chosen for its singular expression in early mouse interneuron progenitors amongst the other 15q13.3 genes, suggesting a putative role of *KLF13* as a driver gene to interneuron-specific abnormalities in neural development. *KLF13* regulation extends to postmitotic neurons, as the study also showed reductions in MGE SST^{+} and PV^{+} interneurons in postnatal day 15 pups – the latter of which has been reported in 15q13.3del mice as well.

Our transcriptomics analysis revealed early implications for cell adhesion and migration disruption that are highly enriched in early 15q13.3 ventral organoids. Assessment of timepoint-specific changes in DEG trajectory revealed alterations in axon

myelinating proteins, suggesting synaptic transmission may be impacted in this model, which is further supported by reductions in broad GABA receptor genes.

5.9 Dorsal and ventral forebrain organoids converge on transcriptional dysregulation at Chr19 and Chr5

After identifying tissue-specific changes amongst the dorsal and ventral forebrain organoids for DNA damage and migration, respectively, we questioned what pathways the two lineages may converge on. After examining the common DEGs shared by each lineage and timepoint, we found that the enrichment for ZNFs at chromosome 19 was recapitulated, which was unsurprising as we also saw this in each dataset independently. We were surprised, however, to find additional enrichment of a cluster of 11 genes on chromosome 5q31.3, a site known for its production of neural-specific cell adhesion protocadherins.[254] This genetic locus has previously been identified through a DNA methylation sequencing as an enriched site in 15q13.3 iPSC and NGN2 glutamatergic-like excitatory neuron models, suggesting epigenetic modifications may affect the accessibility of this region. In this report they generated CRISPR-Cas9 edited isogenic KOs of 15q13.3 genes *KLF13*, *FAN1*, *MTMR10*, and *OTUD7A* and repeated the bulk RNA sequencing to better understand the contributions of individual genes to this phenotype. Surprisingly, the FAN1 KO exclusively was capable of downregulating over 70 protocadherin genes, suggesting it may play a role in regulating cell adhesion to some capacity.

A DEG shared between this dataset and our own was *PCDHGA10*, a protocadherin associated with ASD and neuropsychiatric disorders such as schizophrenia.[255] Recently, this gene was found to be upregulated in an NDD organoid model of PANX1, an ASD

risk gene and paracrine signaling factor that inhibits neuron differentiation and migration. *PANX1* KO organoids produced a decrease in size that is driven by disruptions in the neuroepithelial expansion due to dysregulation of cell-cell and cell-matrix and adhesion proteins. *PCDHGA10* was an upregulated gene in this set, suggesting that it may have a role in shaping organoid size and structure.[256] Interestingly, PANX1 is upregulated in the ventral organoids at an early timepoint but is unchanged in the dorsal tissue.

During early development of the forebrain organoids, we observed an increase in organoid size that was exclusive to the ventral organoids form two 15q13.3del families. Downregulated genes in this tissue type consistently showed enrichment for cell adhesion pathways, including not only the protocadherin genes localized to chromosome 5, but also cadherins like *CDH1*, collagen proteins, and constituents of the ECM. It is therefore possible that ventral organoids are particularly sensitive to changes in adhesion, which may underlie structural differences that contribute towards their growth. Of note, *FAN1* expression in ventral organoids is amongst the highest of the 15q13.3 genes (**Supplementary Fig. 16-17**), showing comparable expression to *KLF13* by the D100 timepoint and suggesting it may play a bigger role in ventral organoid development. Given the reduction protocadherins in *FAN1* KO NGN2 neurons,[121] it is possible that loss of FAN1 affects cell adhesion disproportionately in tissues that confer higher FAN1 expression normally. However, as this change is ventral tissue size is only consistent amongst two families, it cannot be broadly applied to the 15q13.3 phenotype.

5.10 Caveats to organoid modeling.

The objective of chapter four was to determine lineage- and timepoint-specific alterations in the 15q13.3 microdeletion background. We reasoned that using the guided

forebrain model would mitigate the heterogeneity found in the UNO model and would better identify changes specific to excitatory and inhibitor neurons.

The experimental design poses a few caveats, however, relating to the differentiation timeline of ventral organoids and our method of assessment (bulk transcriptomics). First, the ventral organoid system produces interneuron populations slowly, with PV interneurons emerging only after 200+ days of growth. Since we profiled the transcriptome at days 50 and 100, we miss the potential contribution of this cell type to the microdeletion syndrome. The use of bulk transcriptomics also has intrinsic limitations to parsing out diverse cell types, in that it examines mRNA in aggregate. For these two reasons, an examination of a later timepoint would benefit the study, and when supplemented with single cell profiling would better identify which interneuron subtypes are most vulnerable to the 15q13.3 microdeletion syndrome.

The last caveat to organoid modeling more broadly is the lack of standardization in the field to phenotyping and what constitutes a replicate.[257] Disagreements arise about the use of single organoids as a biological replicate, *vs.* those derived from the same well vs. those from different families, making standardization and experimental design difficult. We ensure that our transcriptomics experiments included a minimum of three families and a maximum of our entire cohort (six families), and likewise with phenotypical validations of the UNOs and assembloid experiments (discussed below). Only a single family, however, was used for validation of cell type emergence in the dorsal and ventral organoids (*TBR1* and *SST*, respectively), which does not align with the “gold standard” approach to patient phenotyping. We will supplement this data in the near future with 3-4 more families, however, at present, it is lacking.

5.11 Argument for cell autonomous migration deficits of 15q13.3del interneurons

Our previous transcriptomics experiments highlighted persistent but unique disruptions in the dorsal and ventral forebrain tissue at two timepoints in development. Our findings are the first to suggest broad transcriptional dysregulation in a neural model of the 15q13.3 microdeletion, and ventral forebrain-specific alterations in interneuron migration. Due to examination of these tissues in isolation, however, there is no way of knowing how these two tissue types influence one another and if synergistic mechanisms alter or exacerbate the phenotypes previously identified.

In the fetal brain, these regions do not develop in isolation, but rather integrate into circuit networks composed of dorsal forebrain excitatory neurons combined with inhibitory interneurons which migrate from the ventral subpallium during mid-gestation. This migratory process is essential for neuron maturation and the establishment of E/I balance, which are frequently disrupted in NDDs. The incorporation of dorsal and ventral forebrain organoids into a single “assembloid” system has been vetted as a more complex model of neurodevelopment, capable of saltatory interneuron migration which in turn influences the transcriptome of migrating cells and their synaptic input onto neighbouring excitatory neurons.[131]

We exploited optical clearing methods to capture migration trajectories within the entire assembloid system. This technique was chosen for its ability to examine fluorophore dynamics in whole-tissue without the need for sectioning, thereby retaining the three-dimensional structure of the assembloid system. The advantages to optical

clearing are especially relevant for tissue that produce neurons, as their processes project into multiple, unpredictable directions that are difficult to capture in a thin section.[258]

We characterized the migration of ventral forebrain interneurons in the 15q13.3del background using a novel combination approach previously unreported in the literature. By mixing different genotypes and forebrain lineages together, we were able to infer the contribution of the dorsal microenvironment and compare to cell autonomous or genotype-dependent impacts on interneuron migration. Using this system, we quantified reductions in interneuron migration in each assembloid combinations that contained a 15q13.3del ventral organoid. In contrast to this, we found morphological deficits amongst all but the WT-WT combination pairing, suggesting that the circuit development within a dorsal-ventral assembloid is a complex process relying on what is likely a combination of cell intrinsic and extrinsic mechanisms.

By combining the assembloids during a period of early neurogenesis (day 25), we were able to capture a period where expression of early newborn neurons (TBR1) and neuronal cell adhesion factors (NCAM1) begins.[139] Despite reports of cell migration as early as 7 days (According to manufacturer, STEMCELL Technologies™), we waited a month to ensure that an adequate proportion of cells could be quantified and to allow for a matched comparison to the ventral forebrain sequencing. We were further incentivized to examine migration at this timepoint after comparing the common downregulated genes shared between the D50 and D100 ventral interneurons and the D120 inhibitory neuron subpopulation from the bulk RNA Seq and scRNA Seq dataset, respectively. By examining the enriched biological pathway GO terms, neuron migration was a top pathway by fold enrichment, containing several important migration genes such as: *DAB1*

(a Reelin constituent that also marks a subclass of SST interneurons[259]), *SLIT2*, *EFNB2* (an Ephrin), and NRP1, a neuropillin involved in semaphoring signaling.

The migration enrichment is speculative and not derived from the assembloid system itself, which is known to accelerate the maturation of the tissue and enable more complex microcircuitry. We generated the four combinations (WT-D+ WT-V, WT-D+HET-V, HET-D+WT-V, and HET-D+HET-V) and assessed for migration after on month of innervation. We were surprised to find two combinations produce a reduction in migration: the HET-HET, which is representative of a typical 15q13.3del background, as well as the WT-HET, wherein the microdeletion is restrained to cell types within the ventral organoid only. No migration change was found in the HET-WT combination, suggesting that heterozygosity of 15q13.3 genes in the dorsal forebrain alone does not affect migration proportions. It is therefore more likely that the events governing 15q13.3 interneuron migration deficits are due to cell intrinsic properties, such as changes to cell identity, reductions in cell adhesion, or alterations in signaling amongst, which I will examine below.

An important experimental consideration for this assay was to normalize the number of migrating neurons to the total population of interneurons; in this way, confounders such as a viral labeling efficiency and possible differences in cell proportions are mitigated. The use of the human Synapsin promoter allowed us to examine all neural cell types in the assembloid but introduces ambiguity as to the identity of the migrating neurons. Our bulk transcriptomics experiment suggests that changes in maturation and differentiation may be occurring in the ventral organoids, with reductions in genes such as *SST* and *CALB1* (interneuron subtype markers) in addition to GABA receptors

GABARAP, and interneuron-specific transcription factors *NKX2*-1 and *NKX2*-2. It is therefore possible that the 15q13.3del organoids have an altered maturation that affects the identity of the cell type and its capacity to migrate. Changes in neural population composition is corroborated by reductions in interneural subtype markers *DAB1* (a marker for SST interneuron subpopulations), *SST*, and *CALB.* By contrast, one of the top upregulated gene in both ventral datasets was *TBR1*, which is highly enriched in the developing dorsal telencephalon.[260] *TBR1* overexpression has important ties to migration, and in mice produces disruptions in both excitatory and inhibitory neuron migration in the developing cortex.[261] It is therefore essential for transcription factors that govern neuron differentiation and identity to be tightly regulated during development.[11] The affected pathways in the ventral bulk Seq dataset may provide further insight towards how interneuron migration may be impaired. Amongst the top candidate pathways are cell adhesion (comprised of netrins, cadherins, and protocadherins), and Slit/Robo signaling.

Adhesion proteins play several important roles for cellular migration. During interneuron migration, cytoskeletal dynamics change drastically within interneurons as they extend leading processes which later become branched as the cell interacts with environmental cues. Once a single branch is stabilized, the nucleus translocates towards the branch in a saltatory motion, termed nucleokinesis. These processes are essential for interneuron migration and require the function of adhesion molecules such as cadherins to sustain adhesive strength and proper dynamics.[262,263]

Cadherins were a prominent cell adhesion molecule amongst the downregulate genes – so much so that, “Positive regulation of cell-cell adhesion mediated by cadherin,

GO:2000049" was also a significantly enriched GO term. Cadherin protein complexes have been well characterized in excitatory neurons for their role in facilitating radial migration, but whether this is true for interneurons is now the subject of more recent research. In mouse studies, *N-cadherin* (*Cdh2*) conditional and homozygous KO mice have shown that in the absence of this adhesion protein there are deficits in tangential migration of $CALB2^+$ and SST^+ cortical interneurons as well disruptions to their directionality and motility, respectively.[264,265] Additional interneuron migration deficits have been reported in mouse KO models of Pcdh19 (a protocadherin), and *Celsr3*, an interneuron-specific cadherin protein that was downregulated in the D100 gene set.[264,266,267] Of note, *Celsr3* KO mice have reductions in the neuregulin-1 protein (NRG1), an important mediator to radial and tangential interneuron migration, and shared downregulated gene in D50 and D100 ventral organoids. Cadherins have also been assessed in human interneurons, where they are disrupted in iPSC models of Down syndrome that also have deficits in GABAergic interneuron migration, both *in vitro* and in mouse xenotransplantation assays.[142] These studies demonstrate the importance of cadherins in regulating proper interneuron migration in both mouse and human models and allude to their involvement in migratory signaling pathways like ERBB4-NRG1.[268,269]

Another neural specific signaling pathway is Slit/Robo signaling, which has a well-defined role in axon pathfinding and angiogenesis. The signaling components are evolutionarily conserved proteins including the Slit ligands (*Slit1, 2, 3*), which bind to the Roundabout (*Robo*) receptors (*Robo1, Robo2, Robo3/Rig-1*, and *Robo4*) initially found on the growth cones of commissural axons in the developing CNS. Understanding of the

signaling pattern has since broadened to include the regulation of various tissue and cell types, including inhibitory interneurons. During embryonic development, broad expression of Slit proteins is observed in the intermediate and marginal zone of the cortex, which serves as a migratory path for interneurons of the MGE. *Slit1* is strongly expressed throughout the ventricular zones of the ganglionic eminence as well as at the ventral midline and other regions in the mouse forebrain,[270–272] suggesting dynamic signaling throughout multiple brain regions. Interneuron expression of Robo receptors (specifically *Robo1* and *Robo1*) has been observed *in vitro* and *in vivo* during periods of corticogenesis,[14,273] however, the exact contributions of Slit/Robo signaling to their migration remains unclear. For instance, early explant studies have found LGE-derived interneurons to be repulsed by Slit secreted from the ventricular zone, with inhibition of tangential migration following local addition of Slit to the corticostriatal boundaries.[274] Despite this, *Slit1*$^{-/-}$ / *Slit2*$^{-/-}$ double KO mice show normal tangential migration of LGE interneurons,[275] suggesting that the migratory process is both dynamic and reliant on a variety of ligand and receptor interactions.[276]

Upon examination of the affected Slit/Robo genes in the D50 enriched gene set, we identified both receptors (*ROBO3*) and ligands (*SLIT1* and *SLIT2*), suggesting that cell types within the ventral organoids may be acting as both senders and receivers to the signaling complex. Due to the aggregate nature of the sample, however, the proportions and signaling patterns between cell types can not be delineated. It is possible that the 15q13.3del ventral organoids produce a smaller subpopulation of interneurons with typical expression of migration signaling factors such as Slit-Robo, which would explain

why the proportion of migrating neurons is reduced but that the migration distance remains unaffected.

It is also possible that there is a more pervasive deficit in migration than what was reported, but that the assay or timepoint chosen lacks the sensitivity to identify these abnormalities. It has been reported in similar forebrain assembloid models that the density of migrating cells can increase until day 46 post-assembly, however this is using a protocol that assembles at a much earlier timepoint than the one chosen in our model system.

Due to the 3D nature of the assembloid system, we are also able to characterize the spatial location of the migrating cell relative to the outer boarder of the ventral organoid. We did not see changes in the putative migration distance of each migrating cell, although since this assay is performed at a fixed timepoint, it is impossible to track the actual path taken by each neuron. Use of live-imaging technologies would allow insight into not only the migratory path but also kinetic properties such as saltatory lengths and speed characteristic of interneuron migration.[131] It is likely that a combination of the deficits listed above contribute towards interneuron migration, and only through functional validations (including pharmacological or genetic rescue) can the true biopathology be understood.

5.12 Morphology of migrated interneurons is influenced by genotype and lineage environment

Following their arrival to appropriate regions in the cortex, interneurons undergo changes in cell fate, morphology, and physiology that are unique to the cortical region

that they reside in.[11,259,277,278] These changes are heavily impacted by the environmental cues of the cortex and have only recently been investigated in heterotropic transplantation studies where grafted cells display characteristic similar to the endogenous interneuron populations.[279,280] Changes in interneuron identity post-migration is recapitulated in forebrain assembloids, where migrated cells exhibit increases in dendritic arborization and activity-related genes such as *FOS* and the AMPA-receptor trafficking regulator *GRIP2*[131] relative to stationary neurons. In forebrain assembloids from three 15q13.3del families, we identified a decrease in the neurite length of migrating interneurons amongst all but the WT-WT combinations. This finding is consistent with our previous characterization of reduced axon length in 15q13.3del 2D NGN2 glutamatergic-like excitatory neurons in three separate families. This experiment, however, provides a more nuanced look at how lineage and genotype contribute towards morphology.

The decrease in length was comparable amongst each of the three conditions (WT-HET, HET-WT, and HET-HET), suggesting that changes in 15q13.3 gene expression in either or both the dorsal and ventral forebrain have an influence on neuronal morphology, but that they do not act synergistically in a HET-HET background. It is possible that shared or redundant mechanisms between tissue types underlie these changes, which can be inferred by common DEGs or pathway enrichment between the dorsal and ventral bulk RNA sequencing. The sole biological pathways shared amongst the common DEGs were cell adhesion pathways, of which the predominant genes were from the family of protocadherins. Protocadherins are adhesion genes that help initiate neurite outgrowth and synapse formation through interactions with cytoskeletal regulators and other adhesion proteins.

Pcdha KO mice have drastic reductions in the neurite complexity and length of PV interneurons in vitro, as well as deficits in inhibitory synapse formation, suggesting the importance of these adhesion genes for inhibitory neuron maturation. Importantly, this phenotype is recapitulated in neuropsychiatric models that share disruptions in protocadherin expression, suggesting its role is shared across species.[281]

Loss of protocadherins could also influence the synaptic inputs of the excitatory neurons onto the migrated interneurons. The combination of multiple protocadherins can generate what is thought to be a recognition unit, a structure on the cell surface that facilitates auto-recognition and prevents neurons from synapsing onto themselves (termed *isoneuronal self-avoidance*). This is likely achieved by emission of a repulsive signal and is theoretically meant to facilitate the adhesion between synapses from different cells.[282,283] Loss of protocadherins could potentially disrupt these recognition units in the dorsal forebrain neurons, resulting in an increase in isoneural connections and reductions to heteroneuronal adhesions between different cell types. Due to the importance of synaptic input on interneuron development, and specifically in shaping neurite morphology,[284,285] it is possible that this loss of connectivity drives an immature phenotype that is observed through a reduction in neurite length.

Due to the extensive diversity in interneuron populations, it is equally possible that what is being observed is a change in interneuron identity rather than maturity. This system utilizes a pan-neuronal human synapsin marker that does not distinguish subclasses of interneurons, and so it is impossible to determine the identity of the migrating neurons. Axon morphology varies widely between interneuron populations and is not always a predictor of maturation.[286–288] In addition, heterochronic transplantation

studies of MGE-derived progenitors at different timepoints have shown that the cortical microenvironment can largely influence interneuron identity and laminar fate.[279] This experiment would therefore benefit from broader characterization of the interneuron subclasses that participate in migration. Due to the high density of labeled neurons within the ventral organoids, we were unable to characterize the morphology of the stagnant cells, however it would be interesting to see if the morphology deficits are shared between stagnant and migrating cells or are exclusive to a single population.

5.13 Summary and applications to NDD phenotyping.

The objective of this thesis was to examine early neurodevelopmental milestones in an NDD model to identify which timepoints and cell types are the primary contributors towards pathogenesis. We leveraged organoids for their ability to mimic early fetal development and used three models to examine cell type emergence and circuitry, with unguided neural organoids, guided forebrain organoids and assembloids, respectively. Using these three models, which have historically been examined individually, we identified convergent and cell type-specific disruptions that would not be fully appreciated or apparent had we employed a single organoid model.

Using the largest 15q13.3 microdeletion cohort to date, we found global alterations in transcription that included enrichment of a cluster of evolutionary divergent zinc finger proteins on chromosome 19p13.2 and 19q13.43. While shared, the dorsal organoids showed higher sensitivity to transcriptional changes with roughly 50% more transcription factors affected in the top ranked gene sets. The dorsal dataset exclusively showed an upregulation in genes relating to DNA repair, a finding that was consistent in the radial

glia and excitatory neuron population in the D40 UNO dataset as well. The possibility of excitatory-neuron specific vulnerability to DNA stress and damage has not been explored in this genetic model and is complemented by what appears to be interneuron-specific migration deficits amongst the ventral organoids and inhibitory subpopulation in the UNOs. In this way we have identified convergent and cell type-specific alterations in the dorsal and ventral system that were not apparent in the UNOs alone.

This study provides the first evidence that interneuron migration may be impaired in individuals with the 15q13.3 microdeletion. Importantly, by using a combination of genotypes and lineages we were able characterize this migration in neurotypical vs disease condition dorsal forebrain environments to assess the influence of that factor on migration. This migration impairment is in line with previous mouse studies reporting reductions in interneuron populations, but importantly explores what may be the underlying cause to that phenomenon.

Our preliminary assembloid experiments suggest there are defects in the migration of inhibitory neurons in 15q13.3del forebrain assembloids. Reductions in neurite length, which in this system represent putative leading branches necessary for migration, also suggest an immature morphology in these migrated cells. Reductions in inhibitory neurons support the epileptic phenotypes amongst 15q13.3 patients, as interneurons play a critical role in regulating synaptic activity in cortical networks. Dorsal-ventral forebrain assembloids provide more than a model for cellular migration, however; the circuits formed between excitatory and inhibitory neurons are responsive to signaling cues and are functionally active.[140,173,289] This integration can be partially assessed through migration studies but must be complimented by functional neural population studies to

determine the neural subtypes that participate in the circuits as well as the cellular crosstalk between these cells.

The advantages to using multiple organoid model systems is clear when modeling the 15q13.3 microdeletion syndrome, as it is an incredibly complex CNV with mechanisms of dysfunction that are unique to different cell types. This rigor in modeling should be applied in future NDD studies, as it enables the user to parse out cell-type and convergent mechanisms of disease in order to fully understand the pathobiology.

6 Future directions

This project is the first to examine human fetal neurodevelopment in the 15q13.3 microdeletion background. Using a variety of organoid modeling and transcriptomic approaches, we were able to identify global changes in cell adhesion pathways as well as cell-type specific disruptions in neuron migration and DNA repair. Inference-based analyses predicted global changes to both cellular communication and their developmental trajectories, suggesting a complex phenotype that warrants further investigation. To fully understand and validate the putative pathobiology underlying these abnormalities, I have proposed follow-up studies that would compliment the current experimental findings.

6.1 Exploring the mechanistic underpinnings to organoid growth and abnormal radial glia populations

Increases in organoid sizes are a shared phenotype between the UNO and ventral organoids, suggesting a common underlying pathology in the two model systems. In this thesis I have proposed deficits in cell adhesion as an underlying cause to this change in

size. However, to fully rule out contributions from cell proliferation, further assessments should be made. Specifically, earlier timepoints should be examined for proliferation to see if some aspect of growth is influenced by early proliferation events that we did not capture at the D40 timepoint.

To confirm the possibility of neurogenesis being altered in the D40 UNOs, a 1-hour BrdU pulse could be performed to label cycling progenitors, and the UNOs could be assessed days later for the proportion of progenitors that gave rise to neurons. It is also critical to probe for the neural populations more broadly, as excitatory neurons ("ExN") had the highest increase amongst the 15q13.3del UNOs by proportion and are known to be disrupted functionally in late stage 15q13.3 deletion models.[56,78,79,115,290] Due to the sensitivity of scRNA Seq as an assay, however, it is important that multiple validations, both population-wide and functional, are performed to support these findings. Some examples specifically pertaining to organoid size include cytoskeletal assessments for total cell volume (through use of dye fillers), distribution in 3D space (broadly and with cell-type tagging to determine causative populations), and ECM dynamics that may influence cellular compositions and spacing.

6.1.1 Validating inference-based analyses on cell communication and trajectory

CellChat analysis predicted global deficits in cell adhesion signaling, which was supported by an enrichment of adhesion-related GO term enrichment in individual cell types. The functional consequences to a loss of adherence may include cytoskeletal and ECM changes, which can be assessed in culture with cell attachment (2D) and neurosphere (3D) assays. Cell attachment assays look at structures and cell adhesion very

broadly, however, and protein-specific validations would require targeted immunofluorescent approaches. As mentioned, the CellChat platform is entirely predictive and uses a generalized training dataset that is not neuronal in nature. Due to the unique composition of neural ligands and receptors,[160] it would benefit the analysis to incorporate a neuron-specific platform such as NeuronChat, which was recently made publicly available by the Nie lab.[291] Cellular crosstalk must ultimately be validated in our models, and the most functionally relevant way to do so would be through electrophysiology profiling of the neural circuitry (see **Chapter 7.5**)

Pseudotime analyses present similar challenges as it is also an inferenced based approach. To validate changes to cell type emergence, lineage tracing techniques such as the iTracer platform (pioneered by Dr. Barabra Treutlein's team) can be employed. This technique uses an inducible CRISPR-Cas9 system to introduce a traceable "scar" into a cell population of interest, which is retained in daughter cells throughout differentiation. Simultaneously, the system introduces unique barcodes that are only incorporated into the starting cell population, enabling identification of cells throughout development.[292]

6.2 Assessing DNA instability and trinucleotide repeats using long read sequencing

Due to consistent changes in genomic content at chromosome 19 and the enrichment of biological pathways like replication fork processing, we must consider the possibility of DNA instability occurring within the 15q13.3 organoids as the result of disruption of the DNA nuclease, *FAN1*.

DNA damage has previously been examined in a *FAN1* KO model using a cytokinesis block nuclei assay, which stains for the presence of micronuclei (chromosomal fragments) that occur during DNA damage. In this study, *FAN1* KO cancer cells were challenged with a replication fork inhibitor and responded with an increase in total DNA damage.[239] A similar experiment could be used in dissociated neurons derived from our organoid models to see if this phenomenon persists in neurons at baseline or with a replication inhibitor.

A second method to identifying DNA instability is through long read sequencing, which was named 2022 Nature Method of the year[293] for its incorporation of longer DNA fragments to better resolve repetitive and complex regions of the genome. This technique can identify CNVs and other structural variants, including repeat expansions (AKA "microsatellites"). Trinucleotide repeats are known to be enriched in populations possessing FAN1 variants and are heavily associated with neurodevelopmental disorders such as a Fragile X, which is the most common inherited cause of intellectual disability.[294] The characterization of trinucleotide repeats in 15q13.3 microdeletion patients would provide a novel perspective in how CNVs and genomic instability shape neurodevelopmental disorders, and importantly could be rescued with antisense oligonucleotides as a form of treatment.[295] It would also provide important resolution to the 15q13.3 CNV for each patient, which could help with interpreting variability between patient lines.

6.3 Determination of KLF13 binding sites and transcriptional interactors in neural tissue using chromatin immunoprecipitation sequencing (ChIP-Seq)

Another possible driver to the transcriptional dysregulation seen in the 15q13.3del organoid models is KLF13, a Krüppel-like domain transcription factor whose neural targets (if any) remain uncharacterized. Chromatin immunoprecipitation sequencing allows users to assess putative binding sites of a transcription factor of interest, either through pulldown of the endogenous transcription factor protein or through overexpression of a tagged protein of interest. Presently there are two publicly available KLF13 ChIP Seq experiments (shown in **Fig. 25**), which were performed on HepG2 (a hepatoblastoma cell line used to study the liver) and K562 (cancer cell line) cells on 3xFLAG- and GFP-tagged KLF13 samples, respectively. This dataset was used to argue putative binding of KLF13 to the top downregulated gene, *ZNF558*; however, it well understood that neurons possess highly unique transcriptional signatures and require robust characterization of their transcriptome in order to definitively say a transcription factor regulates a gene. Understanding the neural-specific targets to KLF13 would shed light on its potential contributions to the 15q13.3del phenotype and provide possible targets for rescue (genetically or pharmacologically).

6.4 Spatial RNA sequencing as a means to add spatial resolution to migration dynamics

Our assembloid data provides promising insight into early inhibitory migration deficits in the 15q13.3 microdeletion background. A major limitation to the assay is the ambiguity of the neural identity of the migrating cells, as a pan-neuronal human synapsin

promoter was used for cell tagging. The identity of the stagnant vs migrating neuronal subpopulations therefore remains unknown, as well as the composition of the microenvironment surrounding the migrating cells. To fully profile the dorsal cellular microenvironment and the identity of cell types participating in dorsal-ventral circuitry, we are preparing for a spatial sequencing experiment using 15q13.3del dorsal-ventral forebrain assembloids and the 10x Visium pipeline.

Using the four combinations previously phenotyped, we will assess which cell types are present in the assembloid models, and their orientation in 3D space. To confidently identify the cell types, this experiment will run in parallel with single cell RNA sequencing of individual dorsal and ventral organoids derived from three 15q13.3del families. Using cellular annotations from the scRNA Seq dataset, we will be able to deconvolute the spatial sequencing spots (presently set to 55 µm diameter) and gain meaningful information of the distribution and crosstalk between excitatory and inhibitory neurons. Importantly, we can visually assess the gradients and distribution of morphogens and axon guidance molecules (such as Slit/Robo) which have been implicated from the bulk RNA Seq datasets. Inference based analyses such as CellChat can also be used with higher fidelity between neurons inside of adjacent spots to better understand the functional consequences to these abnormalities. With this information, we hope to better understand how neural circuitry is impaired in the 15q13.3 microdeletion syndrome to better provide insight on therapeutics.

6.5 Functional validations to delays in neuronal maturation using calcium imaging, patch electrophysiology, and rabies tracing

While inference-based platforms can provide an idea of how cell communication is impaired in our model, functional assays are the gold standard approach to understanding neural activity. We intend to perform calcium imaging on AAV9-CAG-somaGCamp6f2 infected forebrain assembloids to generate information on the basal activity of the assembloid system.[151] Population-wide calcium transients will be recorded via live-imaging confocal microscopy two weeks post-infection as previously reported in the assembloid model.[296] This high-throughput assay will inform us on how global changes in neural activity may be impacted in this model.

We predict that the disruptions in inhibitory neuron migration will impact the circuitry that is formed between inhibitory and excitatory neurons and will result in an E/I imbalance typical of an epileptic model. We will compliment these studies with cell type-specific phenotyping using slice electrophysiology with cell type specific fluorescent labelling or addition of cell type specific blockers such as glutamate (excitatory) or GABA (inhibitory) receptor antagonists. Understanding how circuitry is impaired in patient-derived 3D tissue provides a rare glimpse into early developmental impairments that would otherwise be missed in the clinic. With these tools we hope to provide meaningful assessments of the 15q13.3 microdeletion during early circuit development that may eventually give rise to therapeutic intervention.

A third and equally fundamental assay we could perform to assess synaptic connectivity is rabies tracing, which has only been used in select organoid

papers,[135,136,206,297,298] and only three times before in assembloid systems from Dr. Sergio Pașca's lab.[135,136,206] Rabies tracing leverages the natural ability of the rabies virus to travel retrogradely between neural synapses allow investigation of synaptic connectivity starting from a cell population of interest. The virus is modified to produce a fluorescent label that can be used to trace neural circuits, assess neuronal morphology, and measure synaptic activity. Importantly, the application of a cell type specific promoter such as *Dlx2* (for interneurons) and *CaMKII* (for excitatory neurons) can be employed to determine the upstream circuitry from either cell type. Using rabies tracing in our assembloid models could inform the functional connectivity within the system and elucidate which of the migrating cells are actively participating in synaptic connections. Two caveats presently exist to rabies tracing, including its variable infectivity to a starting population, and the toxicity of the rabies virus itself, which can be mitigated with the use of self-inactivating rabies.[299,300]

6.6 Identification of driver genes to 15q13.3 human-specific interneuron dysregulation using isogenic gene knock-out cell lines

CNVs have long been used as a model to identify "driver genes" that disproportionately influence the disease phenotype. Sequencing algorithms such as *Orgo-Seq* from George Church's lab have been generated to identify putative cell-type specific driver genes from organoid bulk and single cell transcriptomics from ASD cohorts. The system is entirely predictive, however, and requires validations mechanistically and on the protein level. To truly validate the role of individual genes in the CNV, isogenic gene KO PSC lines must be and generated into organoids to recapitulate a pathological phenotype, to some degree. A convincing example of this approach was performed in cortical organoid models of the

22q11.2 CNV; using bulk RNA sequencing data from patient-derived organoids, GSEA results suggested the influence of the *DGCR8* gene, which was CRISPR-edited into KO hiPSC line that later recapitulated defects in neuronal activity and calcium signaling previously observed in the CNV background.[139]

A previous study generated isogenic KO iPSC lines of 15q13.3 genes and found convergence of select genes on the same pathways; it is therefore likely that pleotropic mechanisms regulate the 15q13.3 microdeletion syndrome and that they act synergistically to produce the clinical presentations observed in the affected patients. The identification of one or multiple driver genes opens the door to gene therapy as well as biological mechanisms underlying pathology. This project has demonstrated that early fetal neurodevelopment is impacted by the 15q13.3 microdeletion syndrome and providing single genes rather than CNVs to screen for is a much more feasible biomarker than an entire CNV.

7 References

1. Sadler, T. W. Embryology of neural tube development. *Am J Med Genet C Semin Med Genet* **135C**, 2–8 (2005).

2. Le Dréau, G. & Martí, E. Dorsal-ventral patterning of the neural tube: A tale of three signals. *Dev Neurobiol* **72**, 1471–1481 (2012).

3. Dessaud, E., McMahon, A. P. & Briscoe, J. Pattern formation in the vertebrate neural tube: a sonic hedgehog morphogen-regulated transcriptional network. *Development* **135**, 2489–2503 (2008).

4. Götz, M., Hartfuss, E. & Malatesta, P. Radial glial cells as neuronal precursors: a new perspective on the correlation of morphology and lineage restriction in the developing cerebral cortex of mice. *Brain Res Bull* **57**, 777–788 (2002).

5. Stiles, J. & Jernigan, T. L. The Basics of Brain Development. *Neuropsychol Rev* **20**, 327–348 (2010).

6. Pollen, A. A. *et al.* Molecular Identity of Human Outer Radial Glia during Cortical Development. *Cell* **163**, 55–67 (2015).

7. Delgado, R. N. *et al.* Individual human cortical progenitors can produce excitatory and inhibitory neurons. *Nature* **601**, 397–403 (2022).

8. Chen, C.-H. *et al.* Hierarchical Genetic Organization of Human Cortical Surface Area. *Science (1979)* **335**, 1634–1636 (2012).

9. Rakic, P. Radial versus tangential migration of neuronal clones in the developing cerebral cortex. *Proceedings of the National Academy of Sciences* **92**, 11323–11327 (1995).

10. López-Mengual, A. *et al.* Involvement of Mechanical Cues in the Migration of Cajal-Retzius Cells in the Marginal Zone During Neocortical Development. *Front Cell Dev Biol* **10**, (2022).

11. Guo, J. & Anton, E. S. Decision making during interneuron migration in the developing cerebral cortex. *Trends in Cell Biology* vol. 24 Preprint at https://doi.org/10.1016/j.tcb.2013.12.001 (2014).

12. *Extension of Long Leading Processes and Neuronal Migration in the Mammalian Brain Directed by the Chemoattractant Netrin-1*.

13. Andrews, W. D., Barber, M. & Parnavelas, J. G. Slit-Robo interactions during cortical development. in *Journal of Anatomy* vol. 211 188–198 (2007).

14. Andrews, W. *et al.* The role of Slit-Robo signaling in the generation, migration and morphological differentiation of cortical interneurons. *Dev Biol* **313**, 648–658 (2008).

15. Hernández-Miranda, L. R. *et al.* Robo1 regulates semaphorin signaling to guide the migration of cortical interneurons through the ventral forebrain. *Journal of Neuroscience* **31**, 6174–6187 (2011).

16. Fernandez, A., Pasquet-Levy, M., Laure, G., Thümmler, S. & Askenazy, F. Neurodevelopmental Disorders, Psychiatric Comorbidities and Associated Pathologies in Patients with Childhood-Onset Schizophrenia and Premorbid Autistic Symptoms. *Canadian Journal of Psychiatry* **66**, (2021).

17. Chow, J. *et al.* Dissecting the genetic basis of comorbid epilepsy phenotypes in neurodevelopmental disorders. *Genome Med* **11**, (2019).

18. Solberg, B. S. *et al.* Patterns of Psychiatric Comorbidity and Genetic Correlations Provide New Insights Into Differences Between Attention-Deficit/Hyperactivity Disorder and Autism Spectrum Disorder. *Biol Psychiatry* **86**, (2019).

19. Polanczyk, G. V., Willcutt, E. G., Salum, G. A., Kieling, C. & Rohde, L. A. ADHD prevalence estimates across three decades: An updated systematic review and meta-regression analysis. *Int J Epidemiol* **43**, (2014).

20. He, H. *et al.* Trends in the incidence and DALYs of schizophrenia at the global, regional and national levels: Results from the Global Burden of Disease Study 2017. *Epidemiol Psychiatr Sci* (2019) doi:10.1017/S2045796019000891.

21. Fiest, K. M. *et al.* Prevalence and incidence of epilepsy. *Neurology* vol. 88 Preprint at https://doi.org/10.1212/WNL.0000000000003509 (2017).

22. Solmi, M. *et al.* Incidence, prevalence, and global burden of autism spectrum disorder from 1990 to 2019 across 204 countries. *Mol Psychiatry* **27**, (2022).

23. Agnew-Blais, J. C. *et al.* Evaluation of the persistence, remission, and emergence of Attention-deficit/hyperactivity disorder in young adulthood. *JAMA Psychiatry* **73**, (2016).

24. Hansen, S. N., Schendel, D. E. & Parner, E. T. Explaining the increase in the prevalence of autism spectrum disorders: The proportion attributable to changes in reporting practices. *JAMA Pediatr* **169**, (2015).

25. Francés, L. *et al.* Current state of knowledge on the prevalence of neurodevelopmental disorders in childhood according to the DSM-5: a systematic review in accordance with the PRISMA criteria. *Child and Adolescent Psychiatry and Mental Health* vol. 16 Preprint at https://doi.org/10.1186/s13034-022-00462-1 (2022).

26. Pan, P.-Y., Bölte, S., Kaur, P., Jamil, S. & Jonsson, U. Neurological disorders in autism: A systematic review and meta-analysis. *Autism* **25**, 812–830 (2021).

27. Ghirardi, L. *et al.* The familial co-aggregation of ASD and ADHD: A register-based cohort study. *Mol Psychiatry* **23**, 257–262 (2018).

28. Buie, T. *et al.* Evaluation, Diagnosis, and Treatment of Gastrointestinal Disorders in Individuals With ASDs: A Consensus Report. *Pediatrics* **125**, S1–S18 (2010).

29. Mazefsky, C. A., Conner, C. M. & Oswald, D. P. Association between depression and anxiety in high-functioning children with autism spectrum disorders and maternal mood symptoms. *Autism Research* **3**, 120–127 (2010).

30. *Epilepsy*. (2023).

31. Fisher, R. S. *et al.* Operational classification of seizure types by the International League Against Epilepsy: Position Paper of the ILAE Commission for Classification and Terminology. *Epilepsia* **58**, 522–530 (2017).

32. Stafstrom, C. E. & Carmant, L. Seizures and epilepsy: An overview for neuroscientists. *Cold Spring Harb Perspect Biol* **7**, 1–19 (2015).

33. Moyses-Oliveira, M., Yadav, R., Erdin, S. & Talkowski, M. E. New gene discoveries highlight functional convergence in autism and related neurodevelopmental disorders. *Current Opinion in Genetics and Development* vol. 65 195–206 Preprint at https://doi.org/10.1016/j.gde.2020.07.001 (2020).

34. Satterstrom, F. K. *et al.* Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell* **180**, 568-584.e23 (2020).

35. Lieberwirth, J. K. *et al.* AutoCaSc: Prioritizing candidate genes for neurodevelopmental disorders. *Hum Mutat* **43**, 1795–1807 (2022).

36. Folstein, S. & Rutter, M. INFANTILE AUTISM: A GENETIC STUDY OF 21 TWIN PAIRS. *Journal of Child Psychology and Psychiatry* **18**, 297–321 (1977).

37. Bailey, A. *et al.* Autism as a strongly genetic disorder: evidence from a British twin study. *Psychol Med* **25**, 63–77 (1995).

38. Tick, B., Bolton, P., Happé, F., Rutter, M. & Rijsdijk, F. Heritability of autism spectrum disorders: a meta-analysis of twin studies. *Journal of Child Psychology and Psychiatry* **57**, 585–595 (2016).

39. Lemvigh, C. K., Glenthøj, B. Y. & Fagerlund, B. A nation-wide twin study of social cognition in schizophrenia spectrum disorders. *Schizophrenia* **8**, 12 (2022).

40. Larsson, H. *et al.* Genetic and environmental influences on adult attention deficit hyperactivity disorder symptoms: a large Swedish population-based study of twins. *Psychol Med* **43**, 197–207 (2013).

41. Sandin, S. *et al.* The Heritability of Autism Spectrum Disorder. *JAMA* **318**, 1182 (2017).

42. Sada-Fuente, E. *et al.* Common genetic variants contribute to heritability of age at onset of schizophrenia. *Transl Psychiatry* **13**, (2023).

43. Sullivan, P. F., Kendler, K. S. & Neale, M. C. *Schizophrenia as a Complex Trait Evidence From a Meta-analysis of Twin Studies*. *Arch Gen Psychiatry* vol. 60 http://www.vcu.edu/mx/examples.html (2003).

44. Lichtenstein, P. *et al.* Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *The Lancet* **373**, 234–239 (2009).

45. Speed, D. *et al.* Describing the genetic architecture of epilepsy through heritability analysis. *Brain* **137**, 2680–2689 (2014).

46. Anney, R. *et al.* Individual common variants exert weak effects on the risk for autism spectrum disorders. *Hum Mol Genet* **21**, 4781–4792 (2012).

47. Ji, X., Kember, R. L., Brown, C. D. & Bućan, M. Increased burden of deleterious variants in essential genes in autism spectrum disorder. *Proc Natl Acad Sci U S A* **113**, 15054–15059 (2016).

48. Uddin, M. *et al.* Brain-expressed exons under purifying selection are enriched for de novo mutations in autism spectrum disorder. *Nat Genet* **46**, 742–747 (2014).

49. Rauch, A. *et al.* Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *The Lancet* **380**, 1674–1682 (2012).

50. Fromer, M. *et al.* Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nat Neurosci* **19**, 1442–1453 (2016).

51. Tennessen, J. A. *et al.* Evolution and Functional Impact of Rare Coding Variation from Deep Sequencing of Human Exomes. *Science (1979)* **337**, 64–69 (2012).

52. Sanders, S. J. *et al.* De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* **485**, 237–241 (2012).

53. Feliciano, P. *et al.* eP121: Integrating de novo and inherited variants in over 42,607 autism cases identifies variants in new moderate risk genes. *Genetics in Medicine* **24**, (2022).

54. Richter, M. *et al.* Altered TAOK2 activity causes autism-related neurodevelopmental and cognitive abnormalities through RhoA signaling. *Mol Psychiatry* 1 (2018) doi:10.1038/s41380-018-0025-5.

55. Deneault, E. *et al.* Complete Disruption of Autism-Susceptibility Genes by Gene Editing Predominantly Reduces Functional Connectivity of Isogenic Human Neurons. *Stem Cell Reports* **11**, 1211–1225 (2018).

56. Uddin, M. *et al.* OTUD7A Regulates Neurodevelopmental Phenotypes in the 15q13.3 Microdeletion Syndrome. *The American Journal of Human Genetics* **102**, 278–295 (2018).

57. Pinto, D. *et al.* Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* **466**, 368–372 (2010).

58. MacDonald, J. R., Ziman, R., Yuen, R. K. C., Feuk, L. & Scherer, S. W. The Database of Genomic Variants: A curated collection of structural variation in the human genome. *Nucleic Acids Res* **42**, (2014).

59. Zarrei, M., MacDonald, J. R., Merico, D. & Scherer, S. W. A copy number variation map of the human genome. *Nature Reviews Genetics* vol. 16 172–183 Preprint at https://doi.org/10.1038/nrg3871 (2015).

60. Redon, R. *et al.* Global variation in copy number in the human genome. *Nature* **444**, 444–454 (2006).

61. Auton, A. *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).

62. Dennis, M. Y. & Eichler, E. E. Human adaptation and evolution by segmental duplication. *Curr Opin Genet Dev* **41**, 44–52 (2016).

63. Dennis, M. Y. *et al.* Evolution of Human-Specific Neural SRGAP2 Genes by Incomplete Segmental Duplication. *Cell* **149**, 912–922 (2012).

64. Marshall, C. R. *et al.* Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. *Nat Genet* **49**, 27–35 (2017).

65. Takumi, T. & Tamada, K. CNV biology in neurodevelopmental disorders. *Curr Opin Neurobiol* **48**, 183–192 (2018).

66. Wilfert, A. B., Sulovari, A., Turner, T. N., Coe, B. P. & Eichler, E. E. Recurrent de novo mutations in neurodevelopmental disorders: properties and clinical implications. *Genome Med* **9**, 101 (2017).

67. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).

68. Ben-Shachar, S. *et al.* Microdeletion 15q13.3: A locus with incomplete penetrance for autism, mental retardation, and psychiatric disorders. *J Med Genet* **46**, 382–388 (2009).

69. Cooper, G. M. *et al.* A copy number variation morbidity map of developmental delay. *Nat Genet* **43**, 838–846 (2011).

70. Dibbens, L. M. *et al.* Familial and sporadic 15q13.3 microdeletions in idiopathic generalized epilepsy: precedent for disorders with complex inheritance. *Hum Mol Genet* **18**, 3626–3631 (2009).

71. Helbig, I. *et al.* 15q13.3 microdeletions increase risk of idiopathic generalized epilepsy. *Nat Genet* **41**, 160–162 (2009).

72. Gillentine, M. A. *et al.* The Cognitive and Behavioral Phenotypes of Individuals with CHRNA7 Duplications. *J Autism Dev Disord* **47**, 549–562 (2017).

73. Antonacci, F. *et al.* Palindromic GOLGA8 core duplicons promote chromosome 15q13.3 microdeletion and evolutionary instability. *Nat Genet* **46**, 1293–302 (2014).

74. Shinawi, M. *et al.* A small recurrent deletion within 15q13.3 is associated with a range of neurodevelopmental phenotypes. *Nat Genet* **41**, 1269–1271 (2009).

75. Gillentine, M. A. & Schaaf, C. P. The human clinical phenotypes of altered CHRNA7 copy number. *Biochemical Pharmacology* vol. 97 352–362 Preprint at https://doi.org/10.1016/j.bcp.2015.06.012 (2015).

76. Gillentine, M. A. Exploring 15q13.3 copy number variants in iPSCs. in *Current Topics in iPSCs Technology* 333–360 (Elsevier, 2022). doi:10.1016/B978-0-323-99892-5.00017-7.

77. Tropeano, M., Andrieux, J., Vassos, E. & Collier, D. A. Clinical utility gene card for: 15q13.3 microdeletion syndrome. *European Journal of Human Genetics* **22**, 1338 (2014).

78. Sharp, A. J. *et al.* A recurrent 15q13.3 microdeletion syndrome associated with mental retardation and seizures. *Nat Genet* **40**, 322–328 (2008).

79. Lowther, C. *et al.* Delineating the 15q13.3 microdeletion phenotype: A case series and comprehensive review of the literature. *Genetics in Medicine* **17**, 149–157 (2015).

80. Ziats, M. N. *et al.* The complex behavioral phenotype of 15q13.3 microdeletion syndrome. *Genetics in Medicine* **18**, 1111–1118 (2016).

81. Shen, J. & Yakel, J. L. Nicotinic acetylcholine receptor-mediated calcium signaling in the nervous system. *Acta Pharmacol Sin* **30**, 673–680 (2009).

82. Nikiforuk, A., Kos, T., Hołuj, M., Potasiewicz, A. & Popik, P. Positive allosteric modulators of alpha 7 nicotinic acetylcholine receptors reverse ketamine-induced schizophrenia-like deficits in rats. *Neuropharmacology* **101**, 389–400 (2016).

83. Gass, N. *et al.* An acetylcholine alpha7 positive allosteric modulator rescues a schizophrenia-associated brain endophenotype in the 15q13.3 microdeletion, encompassing CHRNA7. *European Neuropsychopharmacology* **26**, 1150–1160 (2016).

84. Ben-David, Y. *et al.* RIC-3 expression and splicing regulate nAChR functional expression. *Mol Brain* **9**, 47 (2016).

85. Unda, B. K. *et al.* Impaired OTUD7A-dependent Ankyrin regulation mediates neuronal dysfunction in mouse and human models of the 15q13.3 microdeletion syndrome. *Mol Psychiatry* **28**, 1747–1769 (2023).

86. Lavallée, G. *et al.* The Kruppel-like transcription factor KLF13 is a novel regulator of heart development. *EMBO Journal* **25**, 5201–5213 (2006).

87. Scohy, S. *et al.* Identification of KLF13 and KLF14 (SP6), novel members of the SP/XKLF transcription factor family. *Genomics* **70**, 93–101 (2000).

88. *KLF13 loss-of-function variation contributes to familial congenital heart defects*. https://gnomad.broadinstitute.org.

89. Guo, Y. H. *et al.* KLF13 Loss-of-Function Mutations Underlying Familial Dilated Cardiomyopathy. *J Am Heart Assoc* **11**, (2022).

90. Lavallée, G. *et al.* The Kruppel-like transcription factor KLF13 is a novel regulator of heart development. *EMBO Journal* **25**, 5201–5213 (2006).

91. Lavallée, G. *et al.* The Kruppel-like transcription factor KLF13 is a novel regulator of heart development. *EMBO Journal* **25**, 5201–5213 (2006).

92. Malwade, S. *et al.* Identification of Vulnerable Interneuron Subtypes in 15q13.3 Microdeletion Syndrome Using Single-Cell Transcriptomics. *Biol Psychiatry* **91**, 727–739 (2022).

93. Ávila-Mendoza, J., Subramani, A. & Denver, R. J. Krüppel-Like Factors 9 and 13 Block Axon Growth by Transcriptional Repression of Key Components of the cAMP Signaling Pathway. *Front Mol Neurosci* **13**, (2020).

94. Ávila-Mendoza, J., Subramani, A., Sifuentes, C. J. & Denver, R. J. Molecular Mechanisms for Krüppel-Like Factor 13 Actions in Hippocampal Neurons. *Mol Neurobiol* **57**, 3785–3802 (2020).

95. Ávila-Mendoza, J. *et al.* KLF13 Regulates the Activity of the GH-Induced JAK/STAT Signaling by Targeting Genes Involved in the Pathway. *Int J Mol Sci* **24**, 11187 (2023).

96. Moore, D. L., Apara, A. & Goldberg, J. L. Krüppel-like transcription factors in the nervous system: Novel players in neurite outgrowth and axon regeneration. *Molecular and Cellular Neuroscience* **47**, 233–243 (2011).

97. Kratz, K. *et al.* Deficiency of FANCD2-Associated Nuclease KIAA1018/FAN1 Sensitizes Cells to Interstrand Crosslinking Agents. *Cell* **142**, 77–88 (2010).

98. Smogorzewska, A. *et al.* A Genetic Screen Identifies FAN1, a Fanconi Anemia-Associated Nuclease Necessary for DNA Interstrand Crosslink Repair. *Mol Cell* **39**, 36–47 (2010).

99. MacKay, C. *et al.* Identification of KIAA1018/FAN1, a DNA Repair Nuclease Recruited to DNA Damage by Monoubiquitinated FANCD2. *Cell* **142**, 65–76 (2010).

100. Deshmukh, A. L. *et al.* FAN1 exo- not endo-nuclease pausing on disease-associated slipped-DNA repeats: A mechanism of repeat instability. *Cell Rep* **37**, (2021).

101. Chaudhury, I., Stroik, D. R. & Sobeck, A. FANCD2-Controlled Chromatin Access of the Fanconi-Associated Nuclease FAN1 Is Crucial for the Recovery of Stalled Replication Forks. *Mol Cell Biol* **34**, 3939–3954 (2014).

102. Ionita-Laza, I. *et al.* Scan statistic-based analysis of exome sequencing data identifies FAN1 at 15q13.3 as a susceptibility gene for schizophrenia and autism. *Proc Natl Acad Sci U S A* **111**, 343–348 (2014).

103. Zhao, X. N. & Usdin, K. FAN1 protects against repeat expansions in a Fragile X mouse model. *DNA Repair (Amst)* **69**, 1–5 (2018).

104. McAllister, B. *et al.* Exome sequencing of individuals with Huntington's disease implicates FAN1 nuclease activity in slowing CAG expansion and disease onset. *Nat Neurosci* **25**, 446–457 (2022).

105. Xing, L. *et al.* Expression of human-specific *ARHGAP11B* in mice leads to neocortex expansion and increased memory flexibility. *EMBO J* **40**, (2021).

106. Fischer, J. *et al.* Human-specific ARHGAP11B ensures human-like basal progenitor levels in hominid cerebral organoids . *EMBO Rep* **23**, (2022).

107. Namba, T. *et al.* Human-Specific ARHGAP11B Acts in Mitochondria to Expand Neocortical Progenitors by Glutaminolysis. *Neuron* **105**, 867-881.e9 (2020).

108. Fischer, J. *et al.* Human-specific <*scp*>*ARHGAP11B*</*scp*> ensures human-like basal progenitor levels in hominid cerebral organoids. *EMBO Rep* **23**, (2022).

109. Rakic, P. Evolution of the neocortex: A perspective from developmental biology. *Nature Reviews Neuroscience* vol. 10 724–735 Preprint at https://doi.org/10.1038/nrn2719 (2009).

110. Hill, R. S. & Walsh, C. A. Molecular insights into human brain evolution. *Nature* vol. 437 64–67 Preprint at https://doi.org/10.1038/nature04103 (2005).

111. Nonaka-Kinoshita, M. *et al.* Regulation of cerebral cortex size and folding by expansion of basal progenitors. *EMBO J* **32**, 1817–1828 (2013).

112. Rojas, D. C. *et al.* Auditory evoked magnetic fields in adults with fragile X syndrome. *Neuroreport* **12**, 2573–2576 (2001).

113. Rotschafer, S. & Razak, K. Altered auditory processing in a mouse model of fragile X syndrome. *Brain Res* **1506**, 12–24 (2013).

114. Yin, J. *et al.* Otud7a Knockout Mice Recapitulate Many Neurological Features of 15q13.3 Microdeletion Syndrome. *The American Journal of Human Genetics* **102**, 296–308 (2018).

115. Kogan, J. H. *et al.* Mouse Model of Chromosome 15q13.3 Microdeletion Syndrome Demonstrates Features Related to Autism Spectrum Disorder. *Journal of Neuroscience* **35**, 16282–16294 (2015).

116. Fejgin, K. *et al.* A mouse model that recapitulates cardinal features of the 15q13.3 microdeletion syndrome including schizophrenia- and epilepsy-related alterations. *Biol Psychiatry* (2014) doi:10.1016/j.biopsych.2013.08.014.

117. Al-Absi, A. R., Qvist, P., Glerup, S., Sanchez, C. & Nyengaard, J. R. Df(h15q13)/+ Mouse Model Reveals Loss of Astrocytes and Synaptic-Related Changes of the Excitatory and Inhibitory Circuits in the Medial Prefrontal Cortex. *Cerebral Cortex* **31**, 1609–1621 (2021).

118. Unda, B. K. *et al.* Impaired OTUD7A-dependent Ankyrin regulation mediates neuronal dysfunction in mouse and human models of the 15q13.3 microdeletion syndrome. *Mol Psychiatry* **28**, 1747–1769 (2023).

119. Bastrup, J. & Larsen, P. H. Optimized CLARITY technique detects reduced parvalbumin density in a genetic model of schizophrenia. *J Neurosci Methods* **283**, 23–32 (2017).

120. Takahashi, K. & Yamanaka, S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* **126**, 663–676 (2006).

121. Zhang, S. *et al.* Network Effects of the 15q13.3 Microdeletion on the Transcriptome and Epigenome in Human-Induced Neurons. *Biol Psychiatry* **89**, 497–509 (2021).

122. Gillentine, M. A. *et al.* Functional Consequences of CHRNA7 Copy-Number Alterations in Induced Pluripotent Stem Cells and Neural Progenitor Cells. *Am J Hum Genet* **101**, 874–887 (2017).

123. Kang, H. J. *et al.* Spatio-temporal transcriptome of the human brain. *Nature* **478**, 483–489 (2011).

124. Stein, J. L. *et al.* A Quantitative Framework to Evaluate Modeling of Cortical Development by Neural Stem Cells. *Neuron* **83**, 69–86 (2014).

125. Jaffe, A. E. *et al.* Developmental regulation of human cortex transcription and its clinical relevance at single base resolution. *Nat Neurosci* **18**, 154–161 (2015).

126. Mariani, J., Coppola, G., Pelphrey, K. A., Howe, J. R. & Vaccarino Correspondence, F. M. FOXG1-Dependent Dysregulation of GABA/ Glutamate Neuron Differentiation in Autism Spectrum Disorders. *Cell* **162**, 375–390 (2015).

127. Cheroni, C. *et al.* Benchmarking brain organoid recapitulation of fetal corticogenesis. *Transl Psychiatry* **12**, 520 (2022).

128. Velasco, S. *et al.* Individual brain organoids reproducibly form cell diversity of the human cerebral cortex. *Nature* **570**, 523–527 (2019).

129. Qian, X. *et al.* Brain-Region-Specific Organoids Using Mini-bioreactors for Modeling ZIKV Exposure. *Cell* **165**, 1238–1254 (2016).

130. Lancaster, M. A. & Knoblich, J. A. Generation of cerebral organoids from human pluripotent stem cells. *Nat Protoc* **9**, 2329–2340 (2014).

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

132. Fair, S. R. *et al.* Electrophysiological Maturation of Cerebral Organoids Correlates with Dynamic Morphological and Cellular Development. *Stem Cell Reports* **15**, 855–868 (2020).

133. Kanton, S., Treutlein, B. & Camp, J. G. *Single-cell genomic analysis of human cerebral organoids. Methods in cell biology.* vol. 159 (Academic Press, 2020).

134. Camp, J. G. *et al.* Human cerebral organoids recapitulate gene expression programs of fetal neocortex development. *Proc Natl Acad Sci U S A* **112**, (2015).

135. Miura, Y. *et al.* Generation of human striatal organoids and cortico-striatal assembloids from human pluripotent stem cells. *Nat Biotechnol* **38**, (2020).

136. Andersen, J. *et al.* Generation of Functional Human 3D Cortico-Motor Assembloids. *Cell* **183**, 1913-1929.e26 (2020).

137. Park, J. *et al.* A logical network-based drug-screening platform for Alzheimerâ€™s disease representing pathological features of human brain organoids. *Nat Commun* 1–13 doi:10.1038/s41467-020-20440-5.

138. Hernández, D. *et al.* Culture Variabilities of Human iPSC-Derived Cerebral Organoids Are a Major Issue for the Modelling of Phenotypes Observed in Alzheimer's Disease. *Stem Cell Rev Rep* **18**, 718–731 (2022).

139. Khan, T. A. *et al. Neuronal defects in a human cellular model of 22q11.2 deletion syndrome*. doi:10.1038/s41591-020-1043-9.

140. Birey, F. *et al.* Dissecting the molecular basis of human interneuron migration in forebrain assembloids from Timothy syndrome. *Cell Stem Cell* **29**, (2022).

141. Negraes, P. D. *et al.* Altered network and rescue of human neurons derived from individuals with early-onset genetic epilepsy. *Mol Psychiatry* doi:10.1038/s41380-021-01104-2.

142. Huo, H.-Q. *et al.* Modeling Down Syndrome with Patient iPSCs Reveals Cellular and Migration Deficits of GABAergic Neurons. *Stem Cell Reports* **10**, 1251–1266 (2018).

143. Subramanian, A. *et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles*. www.pnas.orgcgidoi10.1073pnas.0506580102 (2005).

144. Sebastian, R. *et al.* Schizophrenia-associated NRXN1 deletions induce developmental-timing- and cell-type-specific vulnerabilities in human brain organoids. *Nat Commun* **14**, (2023).

145. Van den Berge, K. *et al.* Trajectory-based differential expression analysis for single-cell sequencing data. *Nat Commun* **11**, 1201 (2020).

146. Trapnell, C. *et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* **32**, 381–386 (2014).

147. Trapnell, C. *et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* **32**, 381–386 (2014).

148. Ho, S.-M. *et al.* Rapid Ngn2-induction of excitatory neurons from hiPSC-derived neural progenitor cells. *Methods* **101**, 113–124 (2016).

149. He, Z. *et al.* Lineage recording reveals dynamics of cerebral organoid regionalization. (2020).

150. Lancaster, M. A. *et al.* Cerebral organoids model human brain development and microcephaly. *Nature* **501**, 373–379 (2013).

151. Velasco, S., Paulsen, B. & Arlotta, P. Highly reproducible human brain organoids recapitulate cerebral cortex cellular diversity. (2019) doi:10.21203/rs.2.9542/v1.

152. Nojima, S. *et al.* CUBIC pathology: Three-dimensional imaging for pathological diagnosis. *Sci Rep* **7**, (2017).

153. Zhao, S. *et al.* Comparison of stranded and non-stranded RNA-seq transcriptome profiling and investigation of gene overlap. *BMC Genomics* **16**, (2015).

154. MacManes, M. D. On the optimal trimming of high-throughput mRNA sequence data. *Front Genet* **5**, (2014).

155. Xijin Ge, S., Jung, D. & Yao, R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. doi:10.5281/zenodo.1451847.

156. Sherwood, J., Quadrato^*^, G., Sherwood^*^, J. L. & Arlotta, P. Long-term culture and electrophysiological characterization of human brain organoids. *Protoc Exch* (2017) doi:10.1038/protex.2017.049.

157. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol* **33**, (2015).

158. Chen, G., Ning, B. & Shi, T. Single-cell RNA-seq technologies and related computational data analysis. *Frontiers in Genetics* vol. 10 Preprint at https://doi.org/10.3389/fgene.2019.00317 (2019).

159. Kanton, S. *et al.* Organoid single-cell genomic atlas uncovers human-specific features of brain development. *Nature* **574**, (2019).

160. Jin, S. *et al.* Inference and analysis of cell-cell communication using CellChat. *Nat Commun* **12**, (2021).

161. Cao, J. *et al.* The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502 (2019).

162. Lancaster, M. A. *et al.* Guided self-organization and cortical plate formation in human brain organoids. *Nat Biotechnol* **35**, 659–666 (2017).

163. Kelava, I., Chiaradia, I., Pellegrini, L., Kalinka, A. T. & Lancaster, M. A. Androgens increase excitatory neurogenic potential in human brain organoids. *Nature* **602**, (2022).

164. Hagey, D. W. & Muhr, J. Sox2 Acts in a Dose-Dependent Fashion to Regulate Proliferation of Cortical Progenitors. *Cell Rep* **9**, 1908–1920 (2014).

165. Hevner, R. F. *et al.* Tbr1 regulates differentiation of the preplate and layer 6. *Neuron* **29**, (2001).

166. Zhou, Y., Song, H. & Ming, G. li. Genetics of human brain development. *Nature Reviews Genetics* Preprint at https://doi.org/10.1038/s41576-023-00626-5 (2023).

167. Supek, F., Bošnjak, M., Škunca, N. & Šmuc, T. Revigo summarizes and visualizes long lists of gene ontology terms. *PLoS One* **6**, (2011).

168. Tatti, R., Haley, M. S., Swanson, O. K., Tselha, T. & Maffei, A. Neurophysiology and Regulation of the Balance Between Excitation and Inhibition in Neocortical Circuits. *Biol Psychiatry* **81**, 821–831 (2017).

169. Narsinh, K. H. *et al.* Endovascular Biopsy of Vertebrobasilar Aneurysm in Patient With Polyarteritis Nodosa. *Front Neurol* **12**, (2021).

170. Johansson, P. A. *et al.* A cis-acting structural variation at the ZNF558 locus controls a gene regulatory network in human brain development. *Cell Stem Cell* **29**, 52-69.e8 (2022).

171. Ohshima, T. Neuronal migration and protein kinases. *Frontiers in Neuroscience* vol. 9 Preprint at https://doi.org/10.3389/fnins.2014.00458 (2015).

172. Legler, D. F. & Thelen, M. New insights in chemokine signaling. *F1000Research* vol. 7 Preprint at https://doi.org/10.12688/f1000research.13130.1 (2018).

173. Sloan, S. A., Andersen, J., Paşca, A. M., Birey, F. & Paşca, S. P. Generation and assembly of human brain region–specific three-dimensional cultures. *Nat Protoc* **13**, (2018).

174. Zhang, Y., Lowe, S., Ding, A. Z. & Li, X. Notch-dependent binary fate choice regulates the Netrin pathway to control axon guidance of Drosophila visual projection neurons. *Cell Rep* **42**, 112143 (2023).

175. Schmidt, E. R. E. *et al.* A human-specific modifier of cortical connectivity and circuit function. *Nature* **599**, 640–644 (2021).

176. Gao, Z. & Godbout, R. Reelin-Disabled-1 signaling in neuronal migration: splicing takes the stage. *Cellular and Molecular Life Sciences* **70**, 2319–2329 (2013).

177. Trakoshis, S. *et al.* Intrinsic excitation-inhibition imbalance affects medial prefrontal cortex differently in autistic men versus women. *Elife* **9**, (2020).

178. Gao, R. & Penzes, P. Common Mechanisms of Excitatory and Inhibitory Imbalance in Schizophrenia and Autism Spectrum Disorders. *Curr Mol Med* **15**, (2015).

179. Fritschy, J. M. Epilepsy, E/I balance and GABAA receptor plasticity. *Front Mol Neurosci* **1**, (2008).

180. Valiente, M. & Martini, F. J. Migration of cortical interneurons relies on branched leading process dynamics. *Cell Adhesion and Migration* vol. 3 Preprint at https://doi.org/10.4161/cam.3.3.8832 (2009).

181. Peyre, E., Silva, C. G. & Nguyen, L. Crosstalk between intracellular and extracellular signals regulating interneuron production, migration and integration into the cortex. *Front Cell Neurosci* **9**, (2015).

182. Bagley, J. A., Reumann, D., Bian, S. & Knoblich, J. A. Fused dorsal-ventral cerebral organoids model human cortical interneuron migration. *bioRxiv* (2017).

183. Smajlagić, D. *et al.* Population prevalence and inheritance pattern of recurrent CNVs associated with neurodevelopmental disorders in 12,252 newborns and their parents. *European Journal of Human Genetics* **29**, 205–215 (2021).

184. Thygesen, J. H. *et al.* Neurodevelopmental risk copy number variants in adults with intellectual disabilities and comorbid psychiatric disorders. *British Journal of Psychiatry* **212**, 287–294 (2018).

185. Tomorsky, J., Parker, P. R. L., Doe, C. Q. & Niell, C. M. Precise levels of nectin-3 are required for proper synapse formation in postnatal visual cortex. *Neural Dev* **15**, (2020).

186. Prandovszky, E. *et al.* Nectin-1 (HveC) is expressed at high levels in neural subtypes that regulate radial migration of cortical and cerebellar neurons of the developing human and murine brain. *J Neurovirol* **14**, 164–172 (2008).

187. Dash, S. *et al.* Heterophilic recognition between E-cadherin and N-cadherin relies on same canonical binding interface as required for E-cadherin homodimerization. *Arch Biochem Biophys* **727**, 109329 (2022).

188. Takahashi, K. *et al.* Nectin/PRR: An Immunoglobulin-like Cell Adhesion Molecule Recruited to Cadherin-based Adherens Junctions through Interaction with Afadin, a PDZ Domain–containing Protein. *J Cell Biol* **145**, 539–549 (1999).

189. Ikeda, W. *et al.* Afadin. *J Cell Biol* **146**, 1117–1132 (1999).

190. Harris, T. J. C. & Tepass, U. Adherens junctions: from molecules to morphogenesis. *Nat Rev Mol Cell Biol* **11**, 502–514 (2010).

191. Inoue, T. *et al.* Nectin-1 spots as a novel adhesion apparatus that tethers mitral cell lateral dendrites in a dendritic meshwork structure of the developing mouse olfactory bulb. *Journal of Comparative Neurology* **523**, 1824–1839 (2015).

192. Miyata, M. *et al.* Localization of nectin-2δ at perivascular astrocytic endfoot processes and degeneration of astrocytes and neurons in nectin-2 knockout mouse brain. *Brain Res* **1649**, 90–101 (2016).

193. Ribic, A., Crair, M. C. & Biederer, T. Synapse-Selective Control of Cortical Maturation and Plasticity by Parvalbumin-Autonomous Action of SynCAM 1. *Cell Rep* **26**, 381-393.e6 (2019).

194. Weller, S. & Gärtner, J. Genetic and clinical aspects of X-linked hydrocephalus (L1 disease): Mutations in the *L1CAM* gene. *Hum Mutat* **18**, 1–12 (2001).

195. Murai, K. K. & Pasquale, E. B. `Eph'ective signaling: forward, reverse and crosstalk. *J Cell Sci* **116**, 2823–2832 (2003).

196. Juliano, R. L. Signal Transduction by Cell Adhesion Receptors and the Cytoskeleton: Functions of Integrins, Cadherins, Selectins, and Immunoglobulin-Superfamily Members. *Annu Rev Pharmacol Toxicol* **42**, 283–323 (2002).

197. Arvanitis, D. N. *et al.* Ephrin B1 maintains apical adhesion of neural progenitors. *Development* **140**, 2082–2092 (2013).

198. Kadison, S. R., Mäkinen, T., Klein, R., Henkemeyer, M. & Kaprielian, Z. EphB Receptors and Ephrin-B3 Regulate Axon Guidance at the Ventral Midline of the Embryonic Mouse Spinal Cord. *The Journal of Neuroscience* **26**, 8909–8914 (2006).

199. Dufour, A. *et al.* Area Specificity and Topography of Thalamocortical Projections Are Controlled by ephrin/Eph Genes. *Neuron* **39**, 453–465 (2003).

200. Südhof, T. C. Synaptic Neurexin Complexes: A Molecular Code for the Logic of Neural Circuits. *Cell* **171**, 745–769 (2017).

201. Chen, L. Y., Jiang, M., Zhang, B., Gokce, O. & Südhof, T. C. Conditional Deletion of All Neurexins Defines Diversity of Essential Synaptic Organizer Functions for Neurexins. *Neuron* **94**, 611-625.e4 (2017).

202. Zhang, C. *et al.* Neurexins Physically and Functionally Interact with GABAA Receptors. *Neuron* **66**, 403–416 (2010).

203. Hennou, S., Khalilov, I., Diabira, D., Ben-Ari, Y. & Gozlan, H. Early sequential formation of functional GABA $_A$ and glutamatergic synapses on CA1 interneurons of the rat foetal hippocampus. *European Journal of Neuroscience* **16**, 197–208 (2002).

204. Südhof, T. C. Synaptic Neurexin Complexes: A Molecular Code for the Logic of Neural Circuits. *Cell* **171**, 745–769 (2017).

205. Zhang, S. *et al.* The protective effect and potential mechanism of NRXN1 on learning and memory in ADHD rat models. *Exp Neurol* **344**, (2021).

206. Kim, J.-I. *et al.* Human assembloids reveal the consequences of CACNA1G gene variants in the thalamocortical pathway. (2023) doi:10.1101/2023.03.15.530726.

207. Chemin, J. *et al.* De novo mutation screening in childhood-onset cerebellar atrophy identifies gain-of-function mutations in the CACNA1G calcium channel gene. *Brain* **141**, 1998–2013 (2018).

208. Yau, H.-J. Neural Development of the Neuregulin Receptor ErbB4 in the Cerebral Cortex and the Hippocampus: Preferential Expression by Interneurons Tangentially Migrating from the Ganglionic Eminences. *Cerebral Cortex* **13**, 252–264 (2003).

209. Wen, L. *et al.* Neuregulin 1 regulates pyramidal neuron activity via ErbB4 in parvalbumin-positive interneurons. *Proceedings of the National Academy of Sciences* **107**, 1211–1216 (2010).

210. Fazzari, P. *et al.* Control of cortical GABA circuitry development by Nrg1 and ErbB4 signalling. *Nature* **464**, 1376–1380 (2010).

211. Arain, F. M., Boyd, K. L. & Gallagher, M. J. Decreased viability and absence-like epilepsy in mice lacking or deficient in the GABAA receptor α1 subunit. *Epilepsia* **53**, e161–e165 (2012).

212. Nilsson, S. R. O. *et al.* A mouse model of the 15q13.3 microdeletion syndrome shows prefrontal neurophysiological dysfunctions and attentional impairment. *Psychopharmacology (Berl)* **233**, 2151–2163 (2016).

213. Fisahn, A., Neddens, J., Yan, L. & Buonanno, A. Neuregulin-1 Modulates Hippocampal Gamma Oscillations: Implications for Schizophrenia. *Cerebral Cortex* **19**, 612–618 (2009).

214. Thelin, J. *et al.* The translationally relevant mouse model of the 15q13.3 microdeletion syndrome reveals deficits in neuronal spike firing matching clinical neurophysiological biomarkers seen in schizophrenia. *Acta Physiologica* **220**, 124–136 (2017).

215. Patzke, C., Acuna, C., Giam, L. R., Wernig, M. & Südhof, T. C. Conditional deletion of *L1CAM* in human neurons impairs both axonal and dendritic arborization and action potential generation. *Journal of Experimental Medicine* **213**, 499–515 (2016).

216. Fleck, J. S. *et al.* Inferring and perturbing cell fate regulomes in human brain organoids. *Nature* (2022) doi:10.1038/s41586-022-05279-8.

217. van Dijk, D. *et al.* Recovering Gene Interactions from Single-Cell Data Using Data Diffusion. *Cell* **174**, 716-729.e27 (2018).

218. Jiang, R., Sun, T., Song, D. & Li, J. J. Statistics or biology: the zero-inflation controversy about scRNA-seq data. *Genome Biology* vol. 23 Preprint at https://doi.org/10.1186/s13059-022-02601-5 (2022).

219. Pietropaolo, S. & Provenzano, G. Editorial: Targeting Excitation-Inhibition Imbalance in Neurodevelopmental and Autism Spectrum Disorders. *Front Neurosci* **16**, (2022).

220. Foss-Feig, J. H. *et al.* Searching for Cross-Diagnostic Convergence: Neural Mechanisms Governing Excitation and Inhibition Balance in Schizophrenia and Autism Spectrum Disorders. *Biol Psychiatry* **81**, 848–861 (2017).

221. Zhang, S. *et al.* Network Effects of the 15q13.3 Microdeletion on the Transcriptome and Epigenome in Human-Induced Neurons. *Biol Psychiatry* **89**, 497–509 (2021).

222. Unda, B. K. *et al.* Impaired OTUD7A-dependent Ankyrin regulation mediates neuronal dysfunction in mouse and human models of the 15q13.3 microdeletion syndrome. *Mol Psychiatry* **28**, 1747–1769 (2023).

223. Lee, E., Lee, J. & Kim, E. Excitation/Inhibition Imbalance in Animal Models of Autism Spectrum Disorders. *Biol Psychiatry* **81**, 838–847 (2017).

224. Nowick, K., Hamilton, A. T., Zhang, H. & Stubbs, L. Rapid Sequence and Expression Divergence Suggest Selection for Novel Function in Primate-Specific KRAB-ZNF Genes. *Mol Biol Evol* **27**, 2606–2617 (2010).

225. Farmiloe, G., Lodewijk, G. A., Robben, S. F., van Bree, E. J. & Jacobs, F. M. J. Widespread correlation of KRAB zinc finger protein binding with brain-developmental gene expression patterns. *Philosophical Transactions of the Royal Society B: Biological Sciences* **375**, 20190333 (2020).

226. Imbeault, M., Helleboid, P. Y. & Trono, D. KRAB zinc-finger proteins contribute to the evolution of gene regulatory networks. *Nature* **543**, 550–554 (2017).

227. Al-Naama, N., Mackeh, R. & Kino, T. C2H2-Type Zinc Finger Proteins in Brain Development, Neurodevelopmental, and Other Neuropsychiatric Disorders: Systematic Literature-Based Analysis. *Front Neurol* **11**, (2020).

228. Nowick, K. *et al.* Gain, Loss and Divergence in Primate Zinc-Finger Genes: A Rich Resource for Evolution of Gene Regulatory Differences between Species. *PLoS One* **6**, e21553 (2011).

229. Cusanovich, D. A. *et al.* A Single-Cell Atlas of In Vivo Mammalian Chromatin Accessibility. *Cell* **174**, 1309-1324.e18 (2018).

230. Pontis, J. *et al.* Hominoid-Specific Transposable Elements and KZFPs Facilitate Human Embryonic Genome Activation and Control Transcription in Naive Human ESCs. *Cell Stem Cell* **24**, 724-735.e5 (2019).

231. Bröckelmann, P. J., de Jong, M. R. W. & Jachimowicz, R. D. Targeting DNA Repair, Cell Cycle, and Tumor Microenvironment in B Cell Lymphoma. *Cells* **9**, 2287 (2020).

232. Huang, R. & Zhou, P. K. DNA damage repair: historical perspectives, mechanistic pathways and clinical translation for targeted cancer therapy. *Signal Transduction and Targeted Therapy* vol. 6 Preprint at https://doi.org/10.1038/s41392-021-00648-7 (2021).

233. Park, K., Debyser, Z., Tabor, S., Richardson, C. C. & Griffith, J. D. Formation of a DNA loop at the replication fork generated by bacteriophage T7 replication proteins. *Journal of Biological Chemistry* **273**, 5260–5270 (1998).

234. Piazza, A. *et al.* Dynamic Processing of Displacement Loops during Recombinational DNA Repair. *Mol Cell* **73**, 1255-1266.e4 (2019).

235. Wishart, M. J. & Dixon, J. E. PTEN and myotubularin phosphatases: from 3-phosphoinositide dephosphorylation to disease. *Trends Cell Biol* **12**, 579–585 (2002).

236. Paulsen, R. D. *et al.* A Genome-wide siRNA Screen Reveals Diverse Cellular Processes and Pathways that Mediate Genome Stability. *Mol Cell* **35**, 228–239 (2009).

237. Lachaud, C. *et al.* Ubiquitinated Fancd2 recruits Fan1 to stalled replication forks to prevent genome instability. *Science (1979)* **351**, 846–849 (2016).

238. Goold, R. *et al.* FAN1 modifies Huntington's disease progression by stabilizing the expanded *HTT* CAG repeat. *Hum Mol Genet* **28**, 650–661 (2019).

239. Chaudhury, I., Stroik, D. R. & Sobeck, A. FANCD2-Controlled Chromatin Access of the Fanconi-Associated Nuclease FAN1 Is Crucial for the Recovery of Stalled Replication Forks. *Mol Cell Biol* **34**, 3939–3954 (2014).

240. Hamperl, S., Bocek, M. J., Saldivar, J. C., Swigut, T. & Cimprich, K. A. Transcription-Replication Conflict Orientation Modulates R-Loop Levels and Activates Distinct DNA Damage Responses. *Cell* **170**, 774-786.e19 (2017).

241. Wang, M. *et al.* Increased Neural Progenitor Proliferation in a hiPSC Model of Autism Induces Replication Stress-Associated Genome Instability. *Cell Stem Cell* **26**, 221-233.e6 (2020).

242. Foran, E. & Trotti, D. Glutamate Transporters and the Excitotoxic Path to Motor Neuron Degeneration in Amyotrophic Lateral Sclerosis. *Antioxid Redox Signal* **11**, 1587–1602 (2009).

243. Konopka, A. & Atkin, J. D. The Role of DNA Damage in Neural Plasticity in Physiology and Neurodegeneration. *Front Cell Neurosci* **16**, (2022).

244. Kruman, I. I. *et al. Homocysteine Elicits a DNA Damage Response in Neurons That Promotes Apoptosis and Hypersensitivity to Excitotoxicity*. (2000).

245. Didier, M. *et al. DNA Strand Breaks Induced by Sustained Glutamate Excitotoxicity in Primary Neuronal Cultures*. *The Journal of Neuroscience* vol. 16 (1996).

246. Wang, Y. & Qin, Z. Molecular and cellular mechanisms of excitotoxic neuronal death. *Apoptosis* **15**, 1382–1402 (2010).

247. Pucilowska, J. *et al.* The 16p11.2 deletion mouse model of autism exhibits altered cortical progenitor proliferation and brain cytoarchitecture linked to the ERK MAPK pathway. *Journal of Neuroscience* **35**, 3190–3200 (2015).

248. Katic, J. *et al.* Interaction of the Cell Adhesion Molecule CHL1 with Vitronectin, Integrins, and the Plasminogen Activator Inhibitor-2 Promotes CHL1-Induced Neurite Outgrowth and Neuronal Migration. *The Journal of Neuroscience* **34**, 14606–14623 (2014).

249. Delogu, A. *et al.* Subcortical Visual Shell Nuclei Targeted by ipRGCs Develop from a Sox14+-GABAergic Progenitor and Require Sox14 to Regulate Daily Activity Rhythms. *Neuron* **75**, 648–662 (2012).

250. Makrides, N. *et al.* Sequential role of SOXB2 factors in GABAergic neuron specification of the dorsal midbrain. *Front Mol Neurosci* **11**, (2018).

251. Stedehouder, J. *et al.* Fast-spiking Parvalbumin Interneurons are Frequently Myelinated in the Cerebral Cortex of Mice and Humans. *Cerebral Cortex* **27**, 5001–5013 (2017).

252. Dubey, M. *et al.* Myelination synchronizes cortical oscillations by consolidating parvalbumin-mediated phasic inhibition. *Elife* **11**, (2022).

253. Naruse, M., Ishizaki, Y., Ikenaka, K., Tanaka, A. & Hitoshi, S. Origin of oligodendrocytes in mammalian forebrains: a revised perspective. *The Journal of Physiological Sciences* **67**, 63–70 (2017).

254. Wu, Q. & Maniatis, T. A Striking Organization of a Large Family of Human Neural Cadherin-like Cell Adhesion Genes. *Cell* **97**, 779–790 (1999).

255. Anazi, S. *et al. Expanding the genetic heterogeneity of intellectual disability*.

256. Apically localized PANX1 impacts neuroepithelial expansion in human cerebral organoids. doi:10.1101/2023.07.28.550996.

257. Jensen, K. B. & Little, M. H. Organoids are not organs: Sources of variation and misinformation in organoid biology. *Stem Cell Reports* vol. 18 1255–1270 Preprint at https://doi.org/10.1016/j.stemcr.2023.05.009 (2023).

258. Fürth, D. *et al.* An interactive framework for whole-brain maps at cellular resolution. *Nat Neurosci* **21**, 139–149 (2018).

259. Garma, L. *et al.* Interneuron diversity in the human dorsal striatum. (2023) doi:10.21203/rs.3.rs-2921627/v1.

260. Bulfone, A. *et al.* T-Brain-1: A homolog of Brachyury whose expression defines molecularly distinct domains within the cerebral cortex. *Neuron* **15**, 63–78 (1995).

261. Yook, C. *et al.* A TBR1-K228E Mutation Induces Tbr1 Upregulation, Altered Cortical Distribution of Interneurons, Increased Inhibitory Synaptic Transmission, and Autistic-Like Behavioral Deficits in Mice. *Front Mol Neurosci* **12**, (2019).

262. Luccardini, C. *et al.* N-cadherin sustains motility and polarity of future cortical interneurons during tangential migration. *Journal of Neuroscience* **33**, 18149–18160 (2013).

263. Martinez-Garay, I. *et al.* Cadherin 2/4 signaling via PTP1B and catenins is crucial for nucleokinesis during radial neuronal migration in the neocortex. *Development (Cambridge)* **143**, 2121–2134 (2016).

264. László, Z. I. *et al.* N-cadherin (Cdh2) Maintains Migration and Postmitotic Survival of Cortical Interneuron Precursors in a Cell-Type-Specific Manner. *Cerebral Cortex* **30**, 1318–1329 (2020).

265. Luccardini, C., Leclech, C., Viou, L., Rio, J. P. & Métin, C. Cortical interneurons migrating on a pure substrate of N-cadherin exhibit fast synchronous centrosomal and nuclear movements and reduced ciliogenesis. *Front Cell Neurosci* **9**, (2015).

266. Ying, G. *et al.* The Protocadherin Gene Celsr3 Is Required for Interneuron Migration in the Mouse Forebrain . *Mol Cell Biol* **29**, 3045–3061 (2009).

267. Lele, Z. *et al. OPEN ACCESS EDITED BY Modifying PCDHHH levels aaects cortical interneuron migration*.

268. Flames, N. *et al.* Short- and Long-Range Attraction of Cortical GABAergic Interneurons by Neuregulin-1. *Neuron* **44**, 251–261 (2004).

269. Poluch, S. & Juliano, S. L. A normal radial glial scaffold is necessary for migration of interneurons during neocortical development. *Glia* **55**, 822–830 (2007).

270. Yuan, W. *et al.* The Mouse SLIT Family: Secreted Ligands for ROBO Expressed in Patterns That Suggest a Role in Morphogenesis and Axon Guidance. *Dev Biol* **212**, 290–306 (1999).

271. Bagri, A. *et al.* Slit Proteins Prevent Midline Crossing and Determine the Dorsoventral Position of Major Axonal Pathways in the Mammalian Forebrain. *Neuron* **33**, 233–248 (2002).

272. Marillat, V. *et al.* Spatiotemporal expression patterns of *slit* and *robo* genes in the rat brain. *Journal of Comparative Neurology* **442**, 130–155 (2002).

273. Andrews, W. *et al.* Robo1 regulates the development of major axon tracts and interneuron migration in the forebrain. *Development* **133**, 2243–2252 (2006).

274. Zhu, Y., Hua-Shun Li, §, Zhou, L., Wu, J. Y. & Rao, Y. *Cellular and Molecular Guidance of GABAergic Neuronal Migration from an Extracortical Origin to the Neocortex*. *Neuron* vol. 23 (1999).

275. Marín, O. *et al.* Directional guidance of interneuron migration to the cerebral cortex relies on subcortical Slit1/2-independent repulsion and cortical attraction. *Development* vol. 130 1889–1901 Preprint at https://doi.org/10.1242/dev.00417 (2003).

276. Barber, M. *et al.* The Role of Robo3 in the Development of Cortical Interneurons. *Cerebral Cortex* **19**, i22–i31 (2009).

277. *On GABAergic interneuron diversity and maturation*.

278. Velmeshev, D. *et al.* Molecular diversity and lineage commitment of human interneuron progenitors. doi:10.1101/2021.05.13.444045.

279. Valcanis, H. & Tan, S.-S. *Layer Specification of Transplanted Interneurons in Developing Mouse Neocortex*. (2003).

280. Quattrocolo, G., Fishell, G. & Petros, T. J. Heterotopic Transplantations Reveal Environmental Influences on Interneuron Diversity and Maturation. *Cell Rep* **21**, 721–731 (2017).

281. Shao, Z. *et al.* Dysregulated protocadherin-pathway activity as an intrinsic defect in induced pluripotent stem cell–derived cortical interneurons from subjects with schizophrenia. *Nat Neurosci* **22**, 229–242 (2019).

282. Mancini, M., Bassani, S. & Passafaro, M. Right Place at the Right Time: How Changes in Protocadherins Affect Synaptic Connections Contributing to the Etiology of Neurodevelopmental Disorders. *Cells* vol. 9 Preprint at https://doi.org/10.3390/cells9122711 (2020).

283. Lefebvre, J. L., Kostadinov, D., Chen, W. V., Maniatis, T. & Sanes, J. R. Protocadherins mediate dendritic self-avoidance in the mammalian nervous system. *Nature* **488**, 517–521 (2012).

284. Dumitriu, D., Cossart, R., Huang, J. & Yuste, R. Correlation Between Axonal Morphologies and Synaptic Input Kinetics of Interneurons from Mouse Visual Cortex. *Cerebral Cortex* **17**, 81–91 (2007).

285. De Marco García, N. V., Karayannis, T. & Fishell, G. Neuronal activity is required for the development of specific cortical interneuron subtypes. *Nature* **472**, 351–355 (2011).

286. Que, L., Lukacsovich, D., Luo, W. & Földy, C. Transcriptional and morphological profiling of parvalbumin interneuron subpopulations in the mouse hippocampus. *Nat Commun* **12**, 108 (2021).

287. Miyoshi, G., Butt, S. J. B., Takebayashi, H. & Fishell, G. Physiologically distinct temporal cohorts of cortical interneurons arise from telencephalic Olig2-expressing precursors. *Journal of Neuroscience* **27**, 7786–7798 (2007).

288. Miyoshi, G. *et al.* Genetic fate mapping reveals that the caudal ganglionic eminence produces a large and diverse population of superficial cortical interneurons. *Journal of Neuroscience* **30**, 1582–1594 (2010).

289. Bagley, J. A., Reumann, D., Bian, S., Lévi-Strauss, J. & Knoblich, J. A. Fused cerebral organoids model interactions between brain regions. *Nat Methods* **14**, (2017).

290. Unda, B. K. *et al.* Impaired OTUD7A-dependent Ankyrin regulation mediates neuronal dysfunction in mouse and human models of the 15q13.3 microdeletion syndrome. *Mol Psychiatry* (2023) doi:10.1038/s41380-022-01937-5.

291. Zhao, W., Johnston, K. G., Ren, H., Xu, X. & Nie, Q. Inferring neuron-neuron communications from single-cell transcriptomics through NeuronChat. *Nat Commun* **14**, (2023).

292. He, Z. *et al.* Lineage recording in human cerebral organoids. *Nat Methods* **19**, 90–99 (2022).

293. Method of the Year 2022: long-read sequencing. *Nat Methods* **20**, 1–1 (2023).

294. Liu, Q., Zhang, P., Wang, D., Gu, W. & Wang, K. Interrogating the ‘unsequenceable’ genomic trinucleotide repeat disorders by long-read sequencing. *Genome Med* **9**, (2017).

295. Derbis, M. *et al.* Short antisense oligonucleotides alleviate the pleiotropic toxicity of RNA harboring expanded CGG repeats. *Nat Commun* **12**, (2021).

296. Birey, F. & Pașca, S. P. Imaging neuronal migration and network activity in human forebrain assembloids. *STAR Protoc* **3**, (2022).

297. Zheng, X. *et al.* Human iPSC-derived midbrain organoids functionally integrate into striatum circuits and restore motor function in a mouse model of Parkinson's disease. *Theranostics* **13**, 2673–2692 (2023).

298. Jgamadze, D. *et al.* Structural and functional integration of human forebrain organoids with the injured adult rat visual system. *Cell Stem Cell* **30**, 137-152.e7 (2023).

299. Ciabatti, E., González-Rueda, A., Mariotti, L., Morgese, F. & Tripodi, M. Life-Long Genetic and Functional Access to Neural Circuits Using Self-Inactivating Rabies Virus. *Cell* **170**, 382-392.e14 (2017).

300. Tran-Van-Minh, A., Ye, Z. & Rancz, E. Quantitative analysis of rabies virus-based synaptic connectivity tracing. *PLoS One* **18**, e0278053 (2023).