DISSECTING THE COMPLEX LANDSCAPE OF NEURODEVELOPMENTAL DISORDERS USING HUMAN PLURIPOTENT STEM CELLS

A Dissertation

Presented to the Faculty of the Weill Cornell Graduate School

Of Medical Sciences

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

By

Gustav Y. Cederquist

December 2018

DISSECTING THE COMPLEX LANDSCAPE OF NEURODEVELOPMENTAL DISORDERS USING HUMAN PLURIPOTENT STEM CELLS

Gustav Y. Cederquist, Ph.D.

Cornell University 2018

Neurodevelopmental disorders (NDDs), which affect 1 in 6 children in the United States, are a group of disabilities caused by abnormal growth and function of the brain. Despite this prevalence, elucidating causative disease mechanisms has remained challenging. This is in large part due to current limitations in experimental modeling of NDDs. Two factors in particular that present challenges for disease modeling are 1) the cellular diversity of brain development and 2) the genetic diversity of disease risk.

Human pluripotent stem cells (hPSCs) have great potential for modeling NDDs, and may be well suited to address the challenges of cellular and genetic diversity for the following reasons. First, hPSCs can generate diverse cell types of the human body, thus providing unprecedented access to disease tissue. Second, hPSCs can organize cellular diversity into functional tissue structures through the use of three-dimensional culture systems called organoids. Third, hPSCs can be engineered to harbor any type of gene mutation, including both monogenic and polygenic genetic risk variants.

In this dissertation, I explore how the unique properties of hPSCs can be harnessed to dissect the complex cellular and genetic landscapes of NDD risk. First, I collaborate to develop a forebrain organoid culture model of ZIKA virus induced microcephaly and find that ZIKA virus infection inhibits growth and degrades the local architecture of forebrain progenitor zones. We further

identify small molecules that inhibit ZIKA virus infection of SOX2+ forebrain progenitors. While organoids recapitulate the cellular diversity and microarchitecture of progenitor zones, and have thus been successful at modeling disorders of tissue growth, they do not recapitulate the overall regional organization of the brain, in which each discrete brain area arise in a stereotyped anatomical location. I therefore next developed a method to trigger spatial patterning in organoids, thus generating topographically organized brain tissue that contains at least five major forebrain subdivisions. Using this system as a tool to study drug toxicity during fetal brain development, I find that cholesterol synthesis-inhibiting drugs, including statins, perturb sonic hedgehog-dependent spatial patterning and growth.

Finally, I develop an hPSC-based multiplex analysis platform for the simultaneous functional characterization of many disease-associated mutations. This approach is applied to model autism, a disorder of social isolation in which hundreds of distinct mutations contribute to diverse clinical outcomes. Multiplex analysis of 27 high-confidence, early developmental, *de novo* autism mutations in hPSC-derived prefrontal cortex tissue identifies that autism mutations commonly perturb cortical neurogenesis and WNT signaling. These phenotypes were used to segregate mutations into distinct subgroups. Autism patients that fall into distinct subgroups exhibit average differences in clinical parameters of communication skills and language development. Taken together, these findings demonstrate that hPSCs are a powerful and general platform to disentangle primary disease pathways from the complex cellular and genetic environments that give rise to NDDs.

BIOGRAPHICAL SKETCH

Gustav Y. Cederquist was born in Farmington, Connecticut in 1987. After completing schoolwork from Conestoga High School in Berwyn, Pennsylvania in 2005, he entered Harvard College. At Harvard, Gustav concentrated in Molecular and Cellular Biology, and performed research in neurodevelopment. He received a Bachelor of Arts in 2009. During the following three years Gustav was employed as a research technician, performing research in cortical development and human neurogenetics. In 2012, Gustav entered the Tri-Institutional Weill Cornell/Rockefeller/Sloan Kettering MD-PhD program in New York City, and joined his thesis laboratory in 2014.

To my family

ACKNOWLEDGMENTS

This work would not have been possible without the funding support of Tri-Institutional Medical Scientist Training Program or the Ruth L. Kirschstein National Research Service Award Fellowship. I am especially thankful to Dr. Lorenz Studer, Director of the Center for Stem Cell Biology and my thesis mentor, whose guidance and teaching has been invaluable. Before starting my thesis research, I had hoped that ultimately I would look back on the experience with fondness, and I can confidently say that Dr. Studer has permanently shaped my curiosity, desire, and will to pursue biomedical science.

It has been a privilege and a pleasure to receive guidance from each of the members of my dissertation committee, Dr. Betsy Ross, Dr. Cori Bargmann, and Dr. Marco Seandel, whose advice has been formative both professionally and personally. I am grateful to the amazing support provided my MD-PhD Program, and by the all the colleagues I have been able to work with over the years. I especially want to thank Dr. Jeffery Macklis, Dr. Elizabeth Engle, Dr. Eiman Azim and Dr. Alex Poulopolous whose early mentorship kindled my interest in science.

Most important of all, thank you to my loving family, my wife, my sister, my mother and father, who are supportive in ways I still cannot believe and who are a constant and everlasting source of joy in my life.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ADI-R	Autism Diagnostic Interview - Revised
ADHD	Attention Deficit Hyperactivity Disorder
AQ	Amodiaquine Dihydrochloride Dihydrate
BMP	Bone Morphogenetic Protein (Gene Family)
Cas	Castor (Gene)
CCND2	Cyclin D2 (Gene)
CHIR99021	Chemical Inhibitor of GSK3, WNT/β-catenin activator
CRISPR	Clustered Regularly Interspersed Short Palindromic Repeats
DD	Developmental Delay
ddPCR	Droplet Digital PCR
DISP1	Dispatched 1 (Gene)
E8/E6	Chemically defined media containing 8 or 6 components
FACS	Fluorescence Activated Cell Sorting
FDR	False Discovery Rate
FGF	Fibroblast Growth Factor (Gene Family)
FOXG1	Forkhead Box Protein G1 (Gene)
GFP	Green Fluorescent Protein
HH	Hippeastrine Hydrobromide (HH)
hNPC	Human Neural Precursor Cell
HPE	Holoprosencephaly
hPSC	Human Pluripotent Stem Cell
Hb	Hunchback (Gene)
Hu	Humeral (Gene)
ID	Intellectual Disability

LDN193189	Chemical Inhibitor of BMP Signaling
IPC	Intermediate Progenitor Cell
iPSC	Induced Pluripotent Stem Cell
iSHH	Inducible SHH expressing hPSC Line
Kr	Kruppel (Gene)
MIX30	Isogenic autism library containing 30 hPSC lines
NDD	Neurodevelopmental Disorder
NEUROD1	Neuronal Differentiation 1 (Gene)
NGN1/2	Neurogenin 1 or Neurogenin 2 (Gene)
oRG	Outer Radial Glia
PCW	Post-Conception Week
Pdm	Pou-domain protein (Gene Family)
PRC	Polycomb Repressive Complex
PTCH1	Patched 1 (Gene)
RA	Retinoic Acid (Gene)
ROI	Region of Interest
SALL1	Spalt-like Transcription Factor 1 (Gene)
SB431542	Chemical Inhibitor of TGFβ/Activin/Nodal Signaling
SHH	Sonic Hedgehog (Gene)
SOX5	SRY-related HMG-box 5 (Gene)
Svp	Seven-up (Gene)
TALEN	TAL Effector Nuclease
TGFβ	Transforming Growth Factor Beta
XAV939	Chemical Inhibitor of WNT/β-catenin via Axin Degradation
Y-27632	Chemical Inhibitor or ROCK pathway
ZIKV	Zika Virus

ZIKV E	Zika Virus Envelope Protein

CHAPTER ONE: INTRODUCTION TO HPSC-BASED MODELING OF NEURODEVELOPMENTAL DISORDERS

The remarkable cellular heterogeneity and processing capabilities of the human brain originate from a single cell, the zygote. The zygote generates neural tissue that is progressively compartmentalized into distinct regional subdivisions. Each subdivision behaves as a mini ecosystem, giving rise to highly specialized repertoires of cell types that assemble into local tissue structures. This protracted developmental journey can be derailed at many points along the way, leading to a group of disorders known as neurodevelopmental disorders (NDDs). Identifying effective therapeutic interventions for NDDs will depend on robust experimental platforms that capture the complexity of brain development and the many ways it can go awry.

Spatial patterning generates regional diversity in the forebrain

The diverse regions of the vertebrate forebrain derive from primordial tissue at the anterior end of the neural tube, known as the prosencephalon. Undifferentiated neural tissue is progressively subdivided into discrete domains by the concerted activity of secreted morphogens (Rallu et al., 2002a). Morphogens are released by small groups of cells called “organizers,” and form gradients of signaling activity across the tissue (Anderson and Stern, 2016). Signaling activity is high near the organizer and diminishes in a distance-dependent manner from the source. Since signaling activity is a function of distance from the organizer, morphogens convey positional information. In addition, morphogens instruct receiving cells to acquire cell identities commensurate with the level of local signaling activity. Thus, morphogens link positional and cell identity information to generate a reliable spatial pattern that dictates the topographic organization of forebrain subdomains (Wolpert, 1969) (**Figure 1.1a**).

Early in its development the forebrain is patterned along its anterior-posterior axis into the telencephalon and diencephalon. The telencephalon is located in the anterior forebrain and is defined by expression of the transcription factor *FOXG1* (Hatini et al., 1994). The diencephalon is located in the posterior forebrain and gives rise to the pretectum, thalamus, and prethalamus. Anterior-posterior patterning is primarily regulated by two signaling gradients. First, WNT/β-catenin signaling activity forms a posterior-high to anterior-low gradient that is generated at the midbrain-hindbrain boundary and is further sculpted by an anterior source of secreted WNT-inhibitors DKK1 and SFRP1 (Shinya et al., 2000). WNT/β-catenin signaling

promotes diencephalic identities, while its inhibition promotes telencephalic fates (Cavodeassi, 2014). Second, FGF8 is released from the anterior neural ridge, rostral to the forebrain, and induces *FOXG1* expression to specify anterior forebrain regions including the telencephalon and optic vesicles (Shimamura and Rubenstein, 1997) (**Figure 1.1b**).

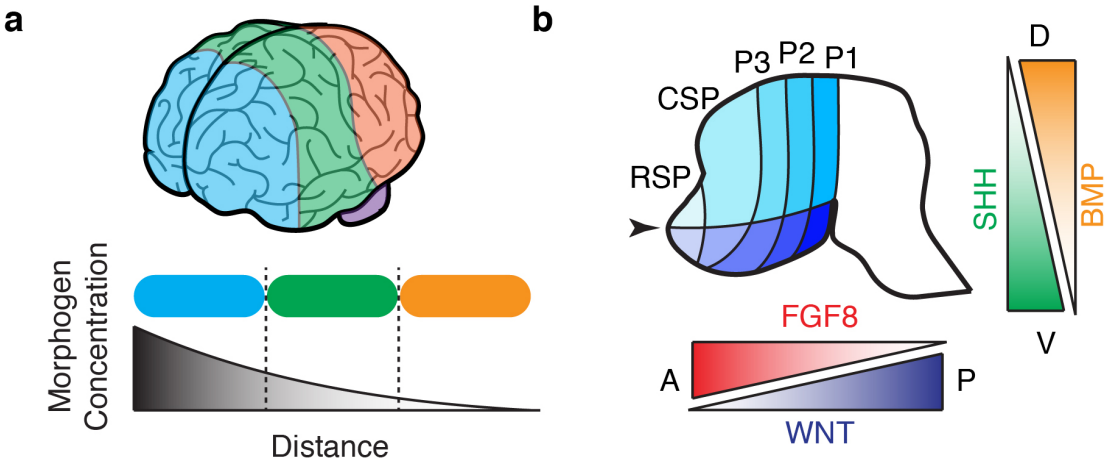


Figure 1.1. Regional patterning of the forebrain
(a) Morphogens establish concentration gradients that specify discrete positional identities in a distant dependent manner. Regional identities are illustrated as cyan, green, and orange territories. **(b)** Intersecting FGF, WNT, SHH, and BMP signaling gradients cooperate to specify discrete forebrain regions that reside within a Cartesian coordinate system of positional information. Dorso-ventral and antero-posterior positional axes define the Cartesian plane. This idea has been formalized by the prosomeric model, which divides the forebrain into transverse and longitudinal segments. P1, P2, and P3 domains give rise to diencephalic regions. The caudal secondary prosencephalon (CSP) and rostral secondary prosencephalon (RSP) give rise to telencephalic territories including the cerebral cortex and hypothalamus. Each transverse segment further divided into the dorso-ventral plane into alar (dorsal) and basal (ventral) regions, demarcated by the arrowhead. A, anterior; D, Dorsal; P, posterior; V, ventral.

The telencephalon and diencephalon are further subdivided along the dorso-ventral axis. Dorsally the telencephalon gives rise to cerebral cortex and hippocampus, while ventrally the telencephalon gives rise to structures of the basal ganglia and diverse populations of interneurons. Sonic hedgehog (SHH)

is released from the ventral midline forming a ventral-high to dorsal-low gradient (Echelard et al., 1993). Gli3 is a transcription factor that promotes dorsal forebrain identity and is initially expressed throughout the telencephalon. SHH represses Gli3 expression ventrally, thereby permitting the acquisition of ventral cell fates (Aoto et al., 2002; Rallu et al., 2002b). Bone morphogenetic protein (BMP) signals have corresponding dorsalizing activity from the roof plate (Monuki et al., 2001) (**Figure 1.1b**).

The intersection of multiple signaling gradients along antero-posterior, dorso-ventral, and medio-lateral axes imposes a grid-like organization in which the position of each forebrain subdomain can be described by its position in a Cartesian-coordinate system. This idea has been formalized by the prosomeric model, which divides the forebrain into five transverse segments called neuromeres, which are intersected by multiple longitudinal segments (Puelles et al., 2013) (**Figure 1.1b**).

This grid-like organization is observed in other developmental contexts, such as patterning of the cerebral cortex. Interestingly the cerebral cortex uses the same specific molecular cues to achieve areal patterning. Just as FGF8 specifies the anterior subdivision of the forebrain, the telencephalon, FGF8 also specifies is the anterior subdivision of the cerebral cortex, the prefrontal cortex (PFC) (Fukuchi-Shimogori and Grove, 2001). Just as WNT/βcatenin signaling promotes posterior forebrain identity, the diencephalon, WNT/βcatenin signaling also specifies the posterior domain of the cortex, the hippocampus (Lee et al., 2000). That the same molecular cues are re-used in sequential developmental contexts indicates that the cellular competence (i.e. internal state) of neural progenitors changes over time. Why are such changes important and how are they accomplished?

Temporal patterning regulates appropriate generation of cell type diversity

Each forebrain subdomain gives rise to a diverse repertoire of cell types that must emerge at the correct developmental time for proper circuit assembly. For example, the neocortex (a subdomain of the dorsal telencephalon) is a six-layered structure that develops in an “inside-out” manner whereby neurons of the deepest layer (layer VI) are born first, followed by neurons of sequentially more superficial layers (Angevine and Sidman, 1961). Broadly speaking, deep layer neurons (in layers V and VI) project to subcortical targets in the spinal cord and thalamus, while later-born superficial layer neurons (in layers II and III) project to intracortical targets.

The temporal pattern of neuronal subtype generation depends on two factors (1) temporal identity and (2) cellular competence (Kohwi and Doe, 2013). A substantial body of work in fly development has shed light on how these two factors pace the developmental clock. Neuroblasts of the drosophilia ventral nerve cord undergo sequential asymmetric divisions to give rise to a stereotyped set of neuronal progeny. The NB7-1 neuroblast gives rise to five distinct motoneurons (U1-U5) over five cell divisions. As the NB7-1 neuroblast transits through these five cell divisions, it sequentially expresses a set of transcription factors (*Hunchback* (*Hb*) > *Kruppel* (*Kr*) > *Pou-domain protein 1 and 2* (*Pdm*) > *Castor* (*Cas*)) that specify the identity of neuronal progeny (Kohwi and Doe, 2013) (**Figure 1.2a**). Thus, these transcription factors specify temporal identity of the progenitor. Feed-forward transcriptional mechanisms and an auxiliary COUP-family transcription factor Seven-up (*Svp*) regulate switching of temporal identity factors (Kanai et al., 2005).

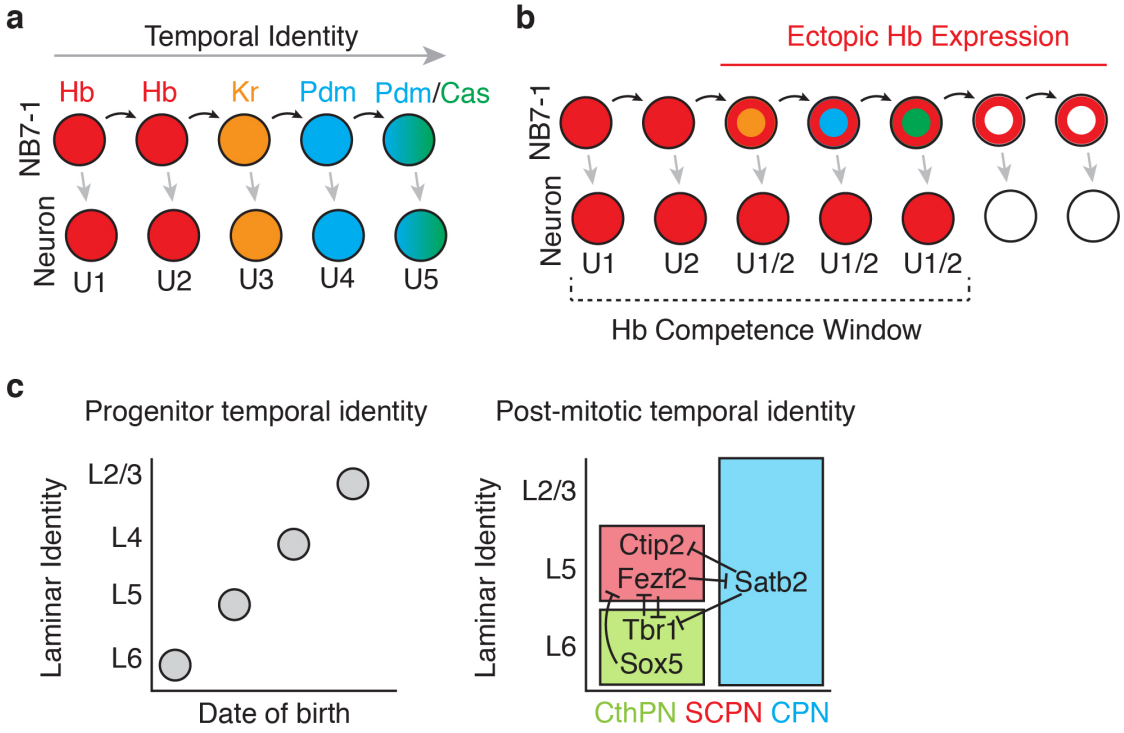


Figure 1.2. Specification of temporal identity in drosophila and mammalian nervous systems.
(a) The drosophila NB7-1 neuroblast undergoes sequential asymmetric divisions, in which a sequence of transcription factors (Hb > Kr > Pdm > Cas) defines temporal identity. Neuronal identity (U1-5) is determined by the temporal identity of the NB7-1 neuroblast. **(b)** Hb specifies U1 or U2 fates during the first two NB7-1 divisions. Misexpression of Hb during the third to fifth NB7-1 division is sufficient to generate ectopic U1/2 neurons, indicating that the Hb competence window extends past Hb temporal identity. **(c)** Temporal patterning of mammalian cortical progenitors is illustrated by the observation that cortical neurons acquire laminar fates based on the developmental age at which progenitors exit cell cycle (date of birth). Temporal patterning in the mammalian cortex continues through post-mitotic development, as subtype identity, defined by axonal connectivity, is determined through a network of cross-repressive transcription factor interactions. Adapted from (Kohwi et al., 2013) (a,b) and (Custo Greig et al., 2013) (c). Cas, Castor; Hb, Hunchback; Kr, Kruppel; L, Layer; Pdm, Pou-domain containing.

Although *Hb* specifies U1 and U2 motoneurons during the first two NB7-1 cell divisions, misexpression of *Hb* through the fifth cell division can specify ectopic U1 or U2 fates. Thus, the U1/U2 competence window extends past the time when U1/U2 fates are normally generated (Pearson and Doe, 2003) (**Figure 1.2b**). This indicates that temporal identity (marked by

expression of *Hb*) is distinct from cellular competence (the time window in which NB7-1 is able to generate *Hb*-dependent progeny). Progenitor competence of NB7-1 neuroblast was shown to be regulated by repositioning of *hb* gene locus to the nuclear lamina for permanent silencing, after the fifth NB7-1 cell division (Kohwi et al., 2013). Loss of the Polycomb Repressive Complex (PRC) function has also been shown to extend the window in which *Kr* could induce ectopic *Kr*-specified neuronal fates (Touma et al., 2012). Thus, genome and chromatin organization are critical regulators of progenitor competence.

The temporal identity of mammalian cortical progenitors was first inferred in the landmark study by Angevine and Sidman that established the "inside-out" pattern of cortical histogenesis, by demonstrating that the age at which cortical progenitors give rise to neuronal progeny determines laminar identity (Angevine and Sidman, 1961) (**Figure 1.2c**). Experiments using mouse pluripotent stem cells, over 50 years later, demonstrated that the correspondence between neuronal birth date and laminar identity is largely preserved *in vitro* (Gaspard et al., 2008). Though these studies reveal an intrinsic mechanism of temporal identity specification, few specific factors that establish temporal identity in the mammalian cortical progenitors have been identified. One rare example is *Ikaros*, a transcription factor that is evolutionarily related to the Hunchback family. Ikaros is normally expressed in early cortical progenitors, and promotes early fates even when ectopically expressed in late progenitors (Alsio et al., 2013).

Interestingly it appears that much of the temporal identity of cortical subtypes is established after the progenitor stage, during post-mitotic development. Five transcription factors with cross-repressive interactions,

Fezf2, Sox5, Tbr1, Ctip2, and Satb2, are expressed during early post-mitotic development to ensure the appropriate sequential generation of cortical subtypes (Custo Greig et al., 2013). For example, Sox5 is expressed in layer VI neurons and represses the expression of Fezf2 and Ctip2, which promote layer V identities (Kwan et al., 2008). Satb2 is expressed in intracortical projection neurons, which occupy both deep and upper layers, and represses Ctip2 expression in layer V (Alcamo et al., 2008) (**Figure 1.2c**).

The transit of cortical progenitors through competence states was first demonstrated in a series of elegant isochronic and heterochronic transplantation experiments in ferret cortex (Desai and McConnell, 2000; Frantz and McConnell, 1996; McConnell, 1985, 1988). These studies found that early cortical progenitors can generate both early and late laminar subtypes when transplanted into embryonic (early) or postnatal (late) environments. On the other hand, later cortical progenitors progressively lose the ability to generate early subtypes, regardless of host environment.

The molecular underpinnings of changes in cellular competence are well illustrated by the response of cortical progenitors to WNT/β–catenin signaling. Early in cortical development, the WNT/β-catenin pathway is highly activated in cortical stem cells and promotes stem cell proliferation (Chenn and Walsh, 2002). Later in development, as stem cells mature into intermediate progenitor cells (IPCs) the WNT/β-catenin pathway promotes differentiation of IPCs into neurons (Hirabayashi et al., 2004; Munji et al., 2011). The dynamic response of cortical progenitors to WNT/β-catenin signaling is at least partially explained by an evolving epigenetic landscape during development. Upon WNT stimulation, β-CATENIN translocates into the nucleus where it activates transcriptional targets. β-CATENIN can activate

both proliferative genes, such as *Myc* and *Cyclin D2* (*Ccnd2*), and it can also activate pro-neurogenic genes, such as *Neurogenin 1 (Ngn1)*, *Neurogenin 2 (Ngn2)*, and *Neuronal differentiation 1 (NeuroD1)* (Hirabayashi et al., 2004; Hirabayashi et al., 2009; Kuwabara et al., 2009). It has been shown that the ability of β-CATENIN to activate the proneurogenic gene *NGN1* depends on a permissive chromatin state in which PRC proteins are not bound to the *NGN1* promoter (Hirabayashi et al., 2009) (**Figure 1.3**). Thus both temporal identity and cellular competence are regulated at the level of chromatin organization.

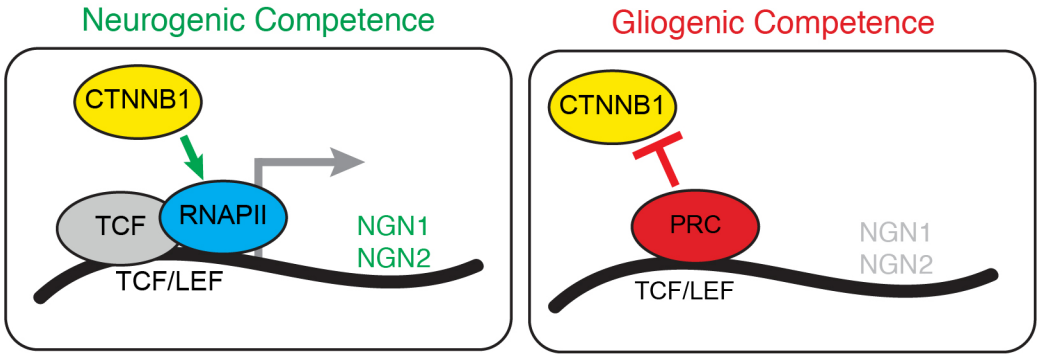


Figure 1.3. Polycomb regulates cellular competence state.
Cortical progenitors transit through competence states, in which they have the potential to generate a restricted set of cellular fates. WNT/β-catenin can promotes neurogenesis during the neurogenic competence window. Here, β-catenin (CTNNB1) can access the TCF/LEF element in the NGN1 and NGN2 promoter region to promote transcription. Later in development, the polycomb repressive complex (PRC) binds the NGN1 and NGN2 promoter region and prevents β-catenin binding. Thus cellular competence is determined by chromatin organization.

NDDs result from aberrant spatial and temporal forebrain patterning

NDDs result from aberrations in numerous developmental processes that affect brain growth, patterning, and maturation. The diversity of underlying etiologies is reflected in the wide range of NDD clinical manifestations, which range from structural brain abnormalities to psychiatric disease.

Holoprosencephaly (HPE) is the most common structural forebrain abnormality at birth and comprises a spectrum of malformations caused by abnormal midline patterning. In the most severe cases, HPE results in a single cerebral vesicle and cyclopia. Intermediate severity HPE can be characterized by near-set eyes and a proboscis, while mild cases may present with agenesis of the corpus callosum or a single central incisor (Gongal et al., 2011).

SHH is a key patterning factor that specifies ventral midline structures such as the hypothalamus, and thus bisecting the eye-field into a bilateral structure (Chiang et al., 1996; Li et al., 1997) (**Figure 1.4a**). Putative loss-of-function mutations in the *SHH* gene, or in key components of the SHH pathway, including genes required for SHH secretion (*DISPATCHED 1 (DISP1*) or SHH signal transduction (*PATCHED 1 (PTCH1)*), lead to HPE (Roessler and Muenke, 2010).

Epilepsy is a neurological condition of abnormal brain activity in which there is an imbalance of excitation and inhibition. In the cortex, excitatory and inhibitory neurons are born from dorsal and ventral telencephalic regions, respectively (Anderson et al., 1997; Anderson et al., 2002). *FOXG1*, in addition to its role in telencephalic specification, also regulates dorsal-ventral patterning (Manuel et al., 2010; Martynoga et al., 2005). FOXG1 function is required for the balanced generation of dorsal and ventral territories and for excitatory and inhibitory cell types (Danesin et al., 2009). *FOXG1* haploinsufficient alleles in humans lead to epilepsy (Seltzer et al., 2014), in addition to a number of other neurologic symptoms.

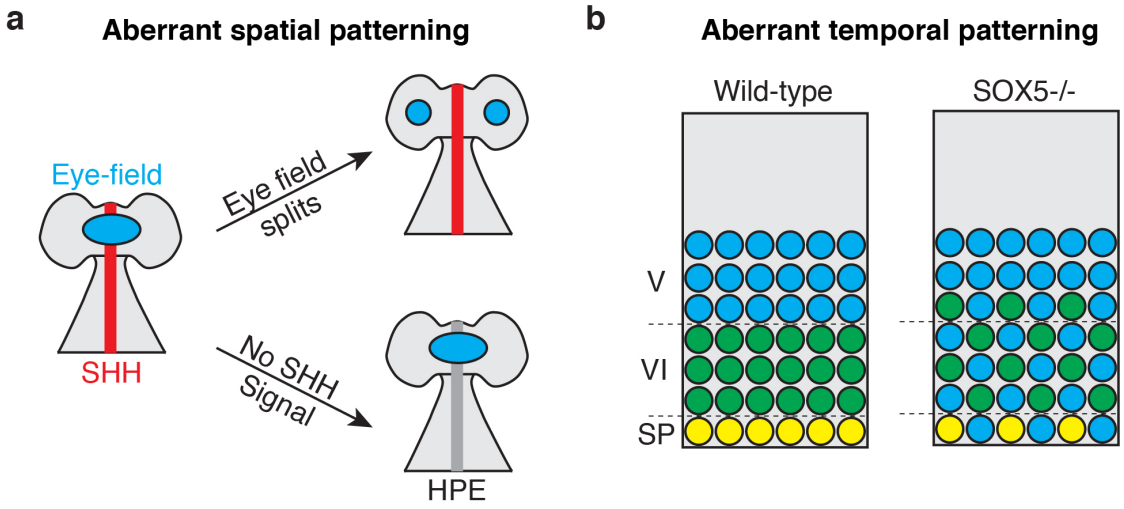


Figure 1.4. Aberrant spatial and temporal patterning leads to NDDs.
(a) SHH is expressed at the ventral midline of the neural tube. By specifying midline identities and promoting midline growth, SHH bisects a single midline eye-field into two eye fields, necessary for the bilateral symmetric positioning of most vertebrate eyes. In the absence of SHH function, the single midline eye field fails to split, as is observed in severe cases of holoprosencephaly (HPE). **(b)** Distinct cortical neuron subtypes, which reside in the subplate (SP), layer VI, and layer V are generated in a deep-to-superficial sequential fashion. This temporal pattern is disrupted in the absence of SOX5 function. Layer V neuronal identities are prematurely generated during the time when subplate and layer VI neurons are normally generated.

From the examples of *SHH* and *FOXG1* mutations, it is clear that NDDs can result from abnormal spatial patterning. Similarly, NDDs can also result from abnormal temporal patterning. SALL1 is a transcription factor that regulates multiple temporal transitions during cortical development, including the transition from cortical stem cell to intermediate progenitor cell (IPC) and the transition of IPC to neuron (Harrison et al., 2012). Mutations in *SALL1* lead to Townes-Brock syndrome, in which a subset of patients exhibit neurological symptoms that include intellectual disability (ID), developmental delay (DD), attention deficit hyperactivity disorder (ADHD), corpus callosum hypoplasia, and seizures. SOX5 is a second transcription factor that regulates temporal identity in the developing cortex. SOX5 controls the sequential generation of corticofugal subtypes by repressing the premature differentiation of later-born

subtypes (Lai et al., 2008) (**Figure 1.4b**). *SOX5* haploinsufficiency leads a number of neurologic impairments including speech delay, ID, and behavioral abnormalities (Lamb et al., 2012).

Uses and limitations of hPSCs in NDD modeling

hPSCs provide a promising approach for modeling NDD pathogenesis as they recapitulate many of the neurodevelopmental processes observed *in vivo*. As mentioned previously, PSC-based *in vitro* models recapitulate the sequential generation of cortical subtypes (Gaspard et al., 2008). A PSC-based model of Timothy Syndrome, which is caused by a mutation in an L-type Calcium Channel, found abnormalities in the proportion of cortical subtypes that are generated (Pasca et al., 2011). This study demonstrates that the utility of PSC-based models to explore how human mutations can disrupt temporal patterning of cortical development.

Remarkably, self-organizing hPSC-based three-dimensional culture models called organoids can recapitulate the spatial organization of the cortical neuroepithelium, including many of the finer cytoarchitectural details (Kadoshima et al., 2013; Lancaster et al., 2013). Broadly, the neuroepithelium exhibits distinct laminar domains. An inner zone near the ventricular surface contains proliferating cells while an outer zone contains post-mitotic neurons. An organoid model using induced PSC (iPSCs) derived from a patient with microcephaly and who harbored a mutation in the cytoskeletal gene CDK5RAP2 exhibited premature depletion of proliferative cells and early differentiation into neurons(Lancaster et al., 2013). This study demonstrates

how an hPSC-based system could be used to explore defective spatial patterning in human cortical disease.

Importantly, hPSCs retain human-specific features of development that are not captured by animal models, expanding the capabilities for disease modeling. For example, human cortex contains a second proliferative zone populated by outer radial glia (oRG), which are not found in mice(Lancaster et al., 2013). oRGs behave as cortical stem cells and contribute to the exceptional evolutionary expansion of the cortical neuroepithelium (Lui et al., 2011). A recent study used organoids to model Miller-Dieker Syndrome, in which patients exhibit lissencephaly. The study found that oRGs exhibited developmental defects that were distinct from those found in other proliferative cell populations in the organoids (Bershteyn et al., 2017). In addition to oRGs, humans contain other unique cell-types such as Von Economo neurons, which are implicated in social behavior and may be vulnerable in frontotemporal dementia (Butti et al., 2013). Humans also exhibit unique endogenous retroelements, recently evolved genes and regulatory elements, and a protracted timeline of development, all of which may cooperate in the expression of human disease.

Complex or polygenic risk factors often contribute to common NDDs, such as autism, epilepsy, and schizophrenia. Such genetic architectures are difficult to recreate in animal models. PSC-based models are potentially useful in this regard. Since iPSCs will harbor the precise genetic variant or variants found in human donors, PSC models can capture complex or polygenic risk alleles. For example, iPSCs derived neurons from schizophrenia patients, who did not have a known genetic etiology, were found to exhibit reduced synaptic connectivity (Brennand et al., 2011).

Despite the potential of hPSCs for modeling complex disease, there are several important limitations. These limitations relate to the complexity of tissue structure and the diversity of genetic variants that can be modeled with hPSCs. For example, the human brain is comprised of many individual regions that must occupy a stereotyped anatomical position to support appropriate brain wiring. While organoids can recapitulate some of the cytoarchitectural features of the brain, they lack this topographic organization.

In addition, hundreds or potentially thousands of distinct genetic variants can contribute to common NDDs such as autism and schizophrenia (Krumm et al., 2014; Schizophrenia Working Group of the Psychiatric Genomics, 2014). While hPSCs have successfully been used to characterize individual variants (Chailangkarn et al., 2016; Shcheglovitov et al., 2013), the laborious nature of hPSC-based disease modeling contrasted with an immense genetic heterogeneity of disease permits only incremental advances in functional characterization of risk alleles. Furthermore, while iPSCs can harbor disease-risk variants, they also likely harbor many variants unrelated to disease. It will thus be necessary to disentangle which variants contribute to disease and which are simply passengers. In this dissertation, I attempt to develop novel hPSC-based methods to enable to study of complex cellular and genetic landscapes associated with many NDDs.

CHAPTER TWO: HIGH-CONTENT SCREENING IN HPSC-NEURAL PROGENITORS IDENTIFIES DRUG CANDIDATES THAT INHIBIT ZIKA VIRUS INFECTION IN FETAL-LIKE ORGANOIDS AND ADULT BRAIN

Abstract

Zika virus (ZIKV) infects fetal and adult human brain, and is associated with serious neurological complications. To date, no therapeutic treatment is available to treat ZIKV infected patients. We performed a high content chemical screen using human pluripotent stem cell (hPSC) derived cortical neuron progenitor cells (hNPCs). Two compounds, hippeastrine hydrobromide (HH) and amodiaquine dihydrochloride dihydrate (AQ), were discovered to inhibit ZIKV infection in hNPCs. HH was further validated to inhibit ZIKV infection and to rescue ZIKV-induced growth and differentiation defects in hNPCs and human fetal-like forebrain organoids. Finally, HH and AQ inhibit ZIKV infection in adult mouse brain *in vivo*. Strikingly, HH suppresses viral propagation when administered to adult mice with active ZIKV infection, highlighting its therapeutic potential. This stem cell-based screen followed by

validation using human forebrain organoids and a mouse model identifies drug candidates for the treatment of ZIKV infection and ZIKV-related neurological complications in fetal and adult patients.

Introduction

ZIKV infection is associated with an increased incidence of congenital defects, most prominently microcephaly (Mlakar et al., 2016; Rasmussen et al., 2016). In adults, ZIKV can lead to serious neurological complications, such as Guillain-Barré syndrome, meningoencephalitis and myelitis (Cao-Lormeau et al., 2016; Carteaux et al., 2016; Mecharles et al., 2016; Parra et al., 2016). Recent studies using hPSC-derived hNPC and forebrain organoids demonstrated that ZIKV replication persists in human fetal neural progenitors for weeks, leading to cell death and attenuation of cortical growth (Garcez et al., 2016; Hanners et al., 2016; Qian et al., 2016; Tang et al., 2016). Furthermore, ZIKV can infect neural stem cells in the murine adult brain (Li et al., 2016). These data support a direct link between ZIKV infection and congenital and adult neurologic symptoms.

There is therefore an urgent need to develop cost-effective therapies for ZIKV patients. Toward this goal, two high content chemical screens have reported the identification of drug candidates inhibiting ZIKV (Barrows et al., 2016; Xu et al., 2016). Although these studies validate the screen platform to identify ZIKV inhibitors, none of the hit compounds have been evaluated for therapeutic use in ZIKV-infected animals. Here, we performed a high content screen using hNPCs and identified two compounds that can inhibit ZIKV infection. One of these drug candidates, hippeastrine hydrobromide (HH)

suppresses ZIKV infection in the brain of mice carrying replicating ZIKV, validating the therapeutic potential of the identified drug candidates.

Results

A high content chemical screen using hNPCs identifies anti-ZIKV compounds. To identify anti-ZIKV compounds, we screened the Prestwick library that includes 1120 FDA-approved drug and drug candidates, selected for known bioavailability and safety in humans. hNPCs were treated with each compound from the library or vehicle control (DMSO) for 1 hour, followed by addition of ZIKV MR766 (MOI=0.125) for 2 hours. After an additional 3 days of culture, cells were stained with antibodies against the ZIKV envelope protein (ZIKV E) and a cell proliferation marker Ki67 (**Figure 2.1 a**). Quantifications were performed by an automated imaging and analysis system. Positive hits were defined as compounds that increased total cell number by greater than 1 fold and suppressed ZIKA infection to less than 20% of control. Based on the fold change of the percentage of ZIKV infection, the signal-to-basal (S/B) ratio is 4.2 and the coefficient of variation (CV) is 6.9%. The Z factor is 0.59 indicating a robust screening system.

Using these criteria, 9 primary hit compounds were confirmed to significantly inhibit ZIKV infection without affecting cell proliferation at 10 μM. Two hit compounds that showed the highest efficacy, HH (IC_{50}=1.95 μM), and Amodiaquine dihydrochloride dihydrate (AQ, IC_{50}=2.28 μM), were chosen for further study (**Figure 2.1 b,c**). Both compounds effectively inhibited ZIKV infection and blocked ZIKV induced growth arrest or apoptosis compared to DMSO controls (**Figure 2.1 d-h**).

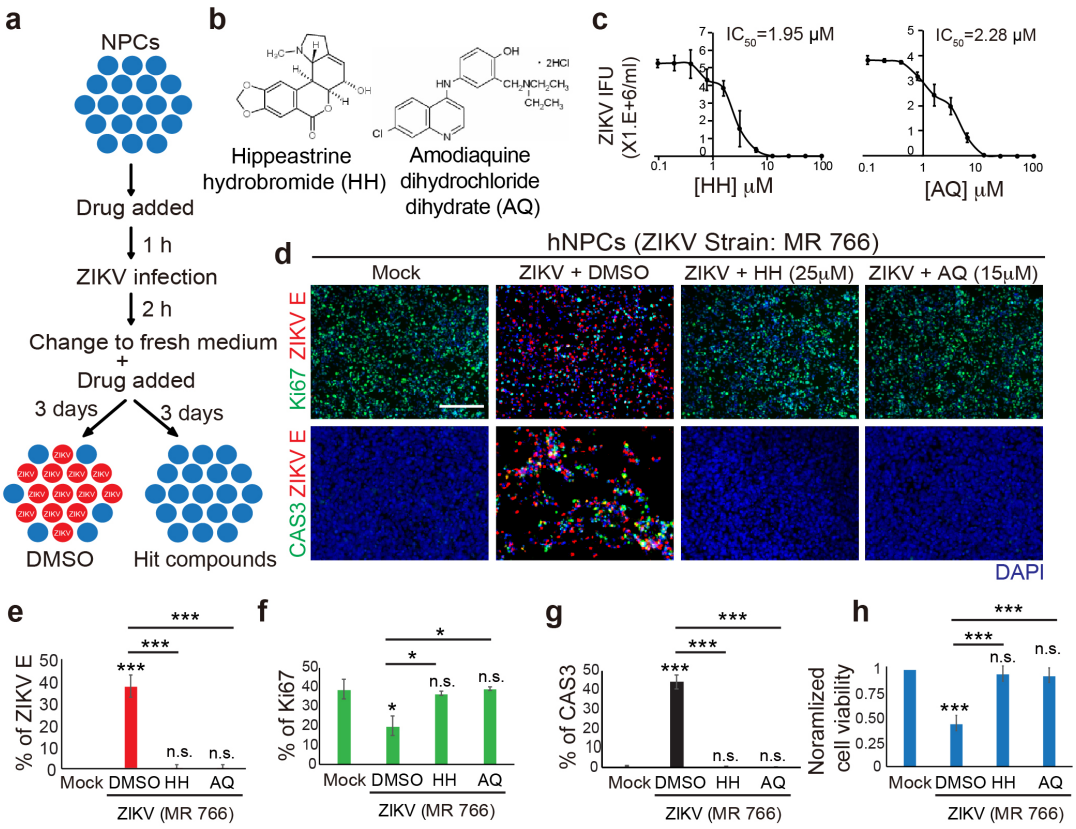


Figure 2.1. A high content chemical screen identifies anti-ZIKV compounds in hNPCs.
(A) Scheme of high content chemical screen. **(B)** Chemical structure of hit compounds. **(C)** Inhibitory curve of HH or AQ in hNPCs **(n=3**) based on infectious particles in supernatant as determined by the Vero reinfection assay. **(D-H)** Immunocytochemistry analysis (D), the quantification of the percentage of ZIKV E^+ cells (E), proliferation rate (F), cell apoptosis rate (G) and the normalized cell viability (H) of hNPCs at 72 h after ZIKV (MR766) infection. Scale bars, 100 μm. (n=6). Experiments and analysis performed by: Ting Zhou, Lei Tan, Shuibing Chen.

HH and AQ also significantly suppressed the production of ZIKV viral RNA (vRNA) (**Figure 2.2a**), and production of infectious ZIKV (**Figure 2.2 b,c**). Furthermore, transcript profiling by RNA-seq analysis suggested that HH or AQ treatment reversed the transcriptional changes induced by ZIKV infection (**Figure 2.2 d**). Finally, 25 μM HH or 15 μM AQ effectively inhibited ZIKV infection of another ZIKV strain, PRVABC59 (**Figure 2.2 e,f**).

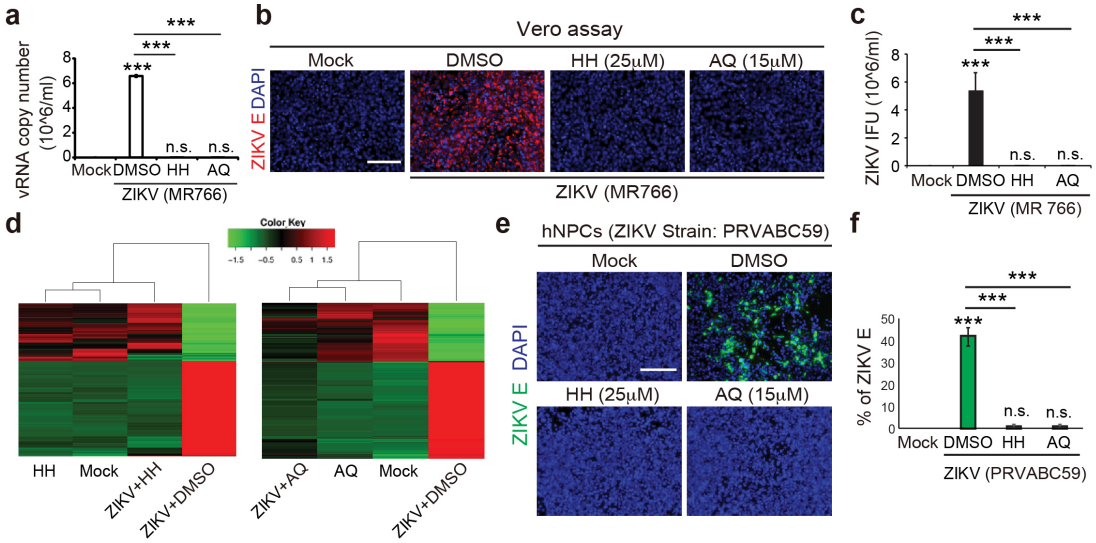


Figure 2.2. Additional preliminary characterization of HH and AQ anti-ZIKV activity.
(a) qRT-PCR analysis to monitor the total ZIKV vRNA in the supernatant of hNPC cultures at 72 h after ZIKV infection. (**b,c**) Vero cell reinfection assay (b) and the quantification of ZIKV infectious particles produced by ZIKV-infected hNPCs (c). Scale bars, 200 μm. (n=3**)**. (**d**) Heatmap of mock-infected hNPCs+DMSO (mock); mock-infected hNPCs+25 μM HH (HH), mock-infected hNPCs+15 μM AQ (AQ), ZIKV infected hNPCs+DMSO (ZIKV+DMSO); ZIKV-infected hNPCs+25 μM HH (ZIKV+HH), and ZIKV-infected hNPCs+15 μM AQ (ZIKV+AQ). Genes were selected from the up/down-regulated genes (fold change >4) in ZIKV-infected hNPCs+DMSO compared to mock-infected hNPCs+DMSO. (**e,f**) Immunocytochemistry analysis (e) and quantification of the percentage of ZIKV E^+ cells (f) of hNPCs at 72 h after ZIKV (PRVABC59, MOI=0.1) infection (n=3). Scale bars, 100 μm. Experiments and analysis performed by: Ting Zhou, Lei Tan, Shuibing Chen.

HH eliminates virus from ZIKV infected hNPCs. To determine whether the drug candidates are capable of eliminating ZIKV in previously infected hNPCs, HH or AQ was applied 2 hours post ZIKV infection (hpi, **Figure 2.3a**). Both compounds were capable of eliminating ZIKV infection with IC_{50} = 3.62 μM (HH, **Figure 2.3 b**) and IC_{50} = 2.81 μM (AQ). In addition, HH successfully rescued ZIKV-induced loss of cell viability with no obvious cytotoxicity (**Figure 2.3 c**). In contrast, AQ affected cell viability at 25 μM, suggesting that AQ can be toxic to hNPCs.

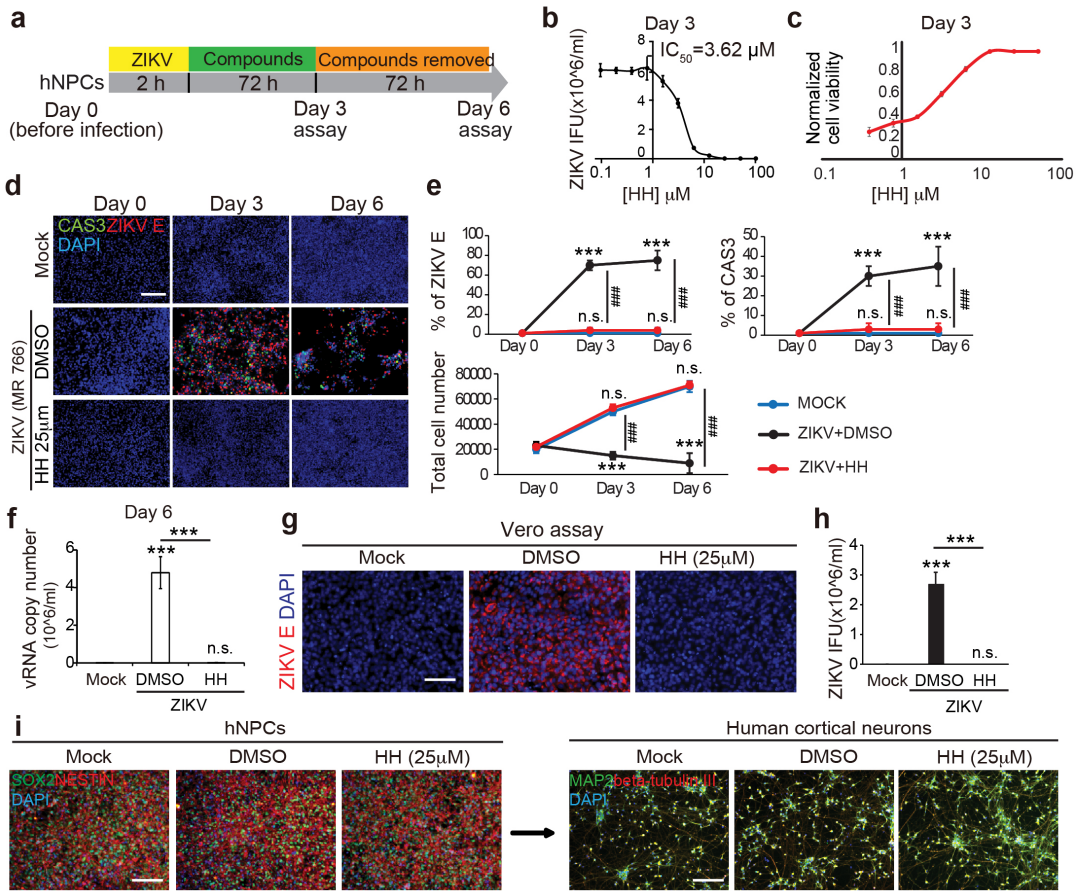


Figure 2.3. HH eliminates virus in ZIKV-infected hNPCs.
(a) Scheme to evaluate the elimination of ZIKV. **(b)** Inhibitory curve of HH calculated based on infectious particles in the supernatant as determined by the Vero reinfection assay (n=3**)**. **(c)** Effective curve of HH calculated based on normalized cell viability (n=3). **(d,e)** Immunocytochemistry analysis (d) and the quantification of the percentage of ZIKV E^+ cells, cell apoptosis rate, and total cell number (e) of hNPCs on day 3 and day 6. Scale bars, 100 μm. (n=3**)**.
(**f**) qRT-PCR analysis to monitor the (+) strand ZIKV vRNA in the supernatant of hNPC cultures on day 6. Scale bars, 200 μm. (n=3). (**g,h**) Vero cell reinfection assay (g) and the quantification of ZIKV infectious particles produced by ZIKV-infected hNPCs at day 6. Scale bars, 200 μm. (**i**) hNPCs with or without ZIKV MR766 infection were treated with 25 μM HH or DMSO for 3 days and then stained for hNPC markers: SOX2 (green) and NESTIN (red), and with DAPI (blue), or were differentiated to cortical neurons. Scale bars, 100 μm. (n=3**)**. Experiments and analysis performed by: Ting Zhou, Lei Tan, Shuibing Chen.

ZIKV-infected hNPCs were treated with HH 3 days, followed by 3 days culture in drug-free medium. HH suppressed ZIKV infection to undetectable levels by day 3 (**Figure 2.3 d**). By day 6, HH-treated hNPCs remained undetectable for

ZIKV infection (**Figure 2.3 d,e**). Consistent with these observations, HH treatment also rescued the cell death rate and total cell number at levels comparable to those of the mock infection conditions (**Figure 2.3 e**). In addition, neither ZIKV vRNA (**Figure 2.3 f**), nor production of infectious ZIKV was detected in day 6 supernatant of HH-treated conditions (**Figure 2.3 g,h**). Furthermore, hNPCs with or without ZIKV infection treated with HH for 3 days maintained the expression of neural progenitor markers, SOX2 and NESTIN, and were capable of differentiation into cortical neurons expressing MAP2 and beta-tubulin III (**Figure 2.3 i**). Together, these data strongly suggest that HH can eliminate ZIKV from infected hNPCs without affecting cell identity or differentiation capacity. In contrast to HH treatment, AQ treatment resulted in viral rebound after 3 days of growth in drug-free medium, suggesting that AQ only transiently suppresses ZIKV infection but fails to effectively eliminate virus from the cells.

To further evaluate the therapeutic potential of HH, qRT-PCR analysis was performed to monitor the dynamics of ZIKV vRNA expression (MR766 strain) in hNPCs. Both (+) and replicating (-) strand of ZIKV vRNAs could be detected at 24 hpi, and the vRNA level increased progressively at 36 and 48 hpi (**Figure 2.4 a**). HH treatment at 24, 36 or 48 hpi successfully suppressed ZIKV vRNA levels (**Figure 2.4 b**). Together, our data show that HH is capable of eliminating virus from hNPCs carrying replicating ZIKV virus.

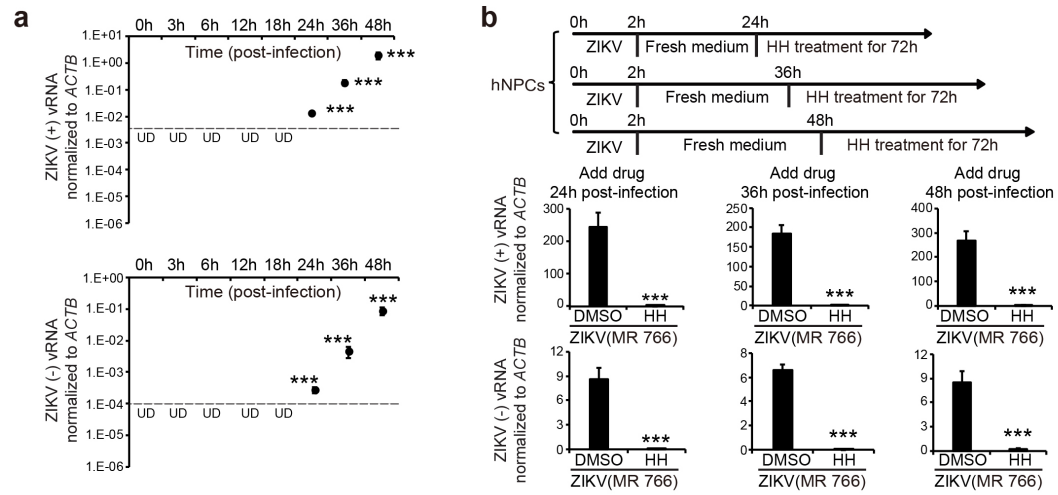


Figure 2.4. Kinetics of ZIKV infection and elimination.
(a) qRT-PCR analysis to quantify the (+) strand and replicating (-) strand ZIKV vRNAs in the hNPCs at different time points post infection. The dash line shows detection limit. *p* values were calculated by one-way repeated measures ANOVA with a Bonferroni test. ***p<0.001. UD: undetectable. **(b)** qRT-PCR analysis of hNPCs, on which 25 μM HH or DMSO were added at 24, 36 and 48 hpi (MR766) infection and maintained for additional 3 days. *p* values were calculated by unpaired two-tailed Student's t-test. *p<0.05, **p<0.01 and ***p<0.001, if not mentioned specifically. Experiments and analysis performed by: Ting Zhou, Lei Tan, Shuibing Chen.

HH inhibits ZIKV infection in a human fetal-like forebrain organoid model. hPSC-derived forebrain organoids provide a unique opportunity to study disorders of brain growth including microcephaly (Bershteyn and Kriegstein, 2013). Here, we systematically test the short-term and long-term effects of HH on day 20 (D20) forebrain organoids following ZIKV infection. D20 forebrain organoids were inoculated with ZIKV (MR766) for 24 h, followed by treatment with 25 μM HH for 3 days (D20+4, short-term) or 17 days (D20+18, long-term). Organoids were analyzed for ZIKV infection, cell proliferation, and apoptosis (**Figure 2.5 a**). Immunohistochemical analysis of SOX2 and ZIKV E showed that D20 organoids consist mainly of $SOX2^+$ NPCs, and more than 80% of the cells were infected by ZIKV at D20+4 (**Figure 2.5 b,c**), Strikingly, HH was effective at suppressing the percentage of ZIKV E^+

cells (**Figure 2.5 c**) and total ZIKV vRNA level in the supernatant (**Figure 2.5 d**) to undetectable levels. In addition, infection of D20 organoids with ZIKV reduced the relative number of Ki67^{+} proliferating cells, and increased the number of CAS3^{+} apoptotic cells (**Figure 2.5 e**). These phenotypes were completely rescued by HH (**Figure 2.5 f**).

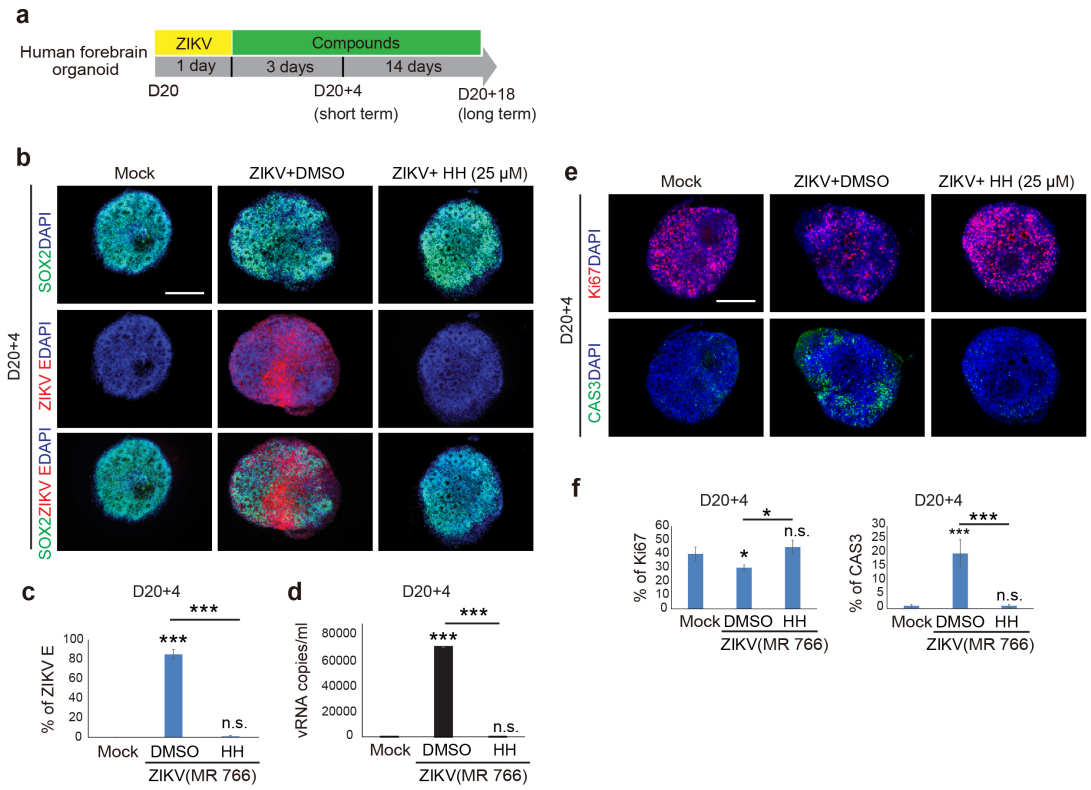


Figure 2.5. HH inhibits ZIKV infection in short-term forebrain organoid cultures **(a)** Scheme to evaluate HH using forebrain organoids. **(b,c)** Immunocytochemistry analysis (b) and the quantification of the percentage of ZIKV E^{+} cells (c) in mock, DMSO or 25 μM HH treated ZIKV-infected D20 organoids at D20+4 (n=3). Scale bars, 100 μm. (**d**) qRT-PCR analysis of (+) strand ZIKV vRNA in the supernatant of DMSO or 25 μM HH treated ZIKV-infected D20 organoids (n=3). **(e)** HH rescues ZIKV-related proliferation and apoptosis defects in D20+4 organoids. Scale bars, 100 μm. Experimental Design: Ting Zhou, Shuibing Chen, Lorenz Studer. Organoid Culture: Gustav Cederquist. ZIKV infection: Lei Tan. Data Analysis: Ting Zhou, Lei Tan, Gustav Cederquist, Lorenz Studer, Shuibing Chen.

We next assessed whether HH confers durable, long-term control of ZIKV infection and whether it can rescue microcephaly-related defects caused by ZIKV infection. D20 organoids were infected with ZIKV and after 18 additional days of growth, analyzed for ZIKV infection. HH treated organoids were devoid of any detectable ZIKV, while ZIKV infected organoids without drug treatment were strongly infiltrated by virus (**Figure 2.6 a**). Moreover, HH successfully rescued ZIKV-induced structural defects seen in long-term cultures. Similar to the mock condition, in the ZIKV-infected organoids treated with HH, there are clearly defined zones of progenitors ($SOX2^+$) and neurons ($TUJ1^+$), while in the ZIKV infected organoids, the progenitor and neuronal zones are intermixed and lack any discernible organization (**Figure 2.6 b**). To determine whether HH rescues microcephaly-related pathologies, organoid size was measured at 18 days post infection (dpi). At 18 dpi, the cross-sectional area of ZIKV infected organoids was significantly decreased compared to mock infected conditions, consistent with modeling ZIKV-associated congenital microcephaly (**Figure 2.6 c,d**). In contrast, ZIKV-infected organoids that were treated with 25 μM HH grew to a similar size as mock-infected organoids (**Figure 2.6 c,d**). Together, these data indicate that HH successfully inhibits ZIKV activity during long-term forebrain organoid growth and rescues crucial aspects associated with fetal brain development.

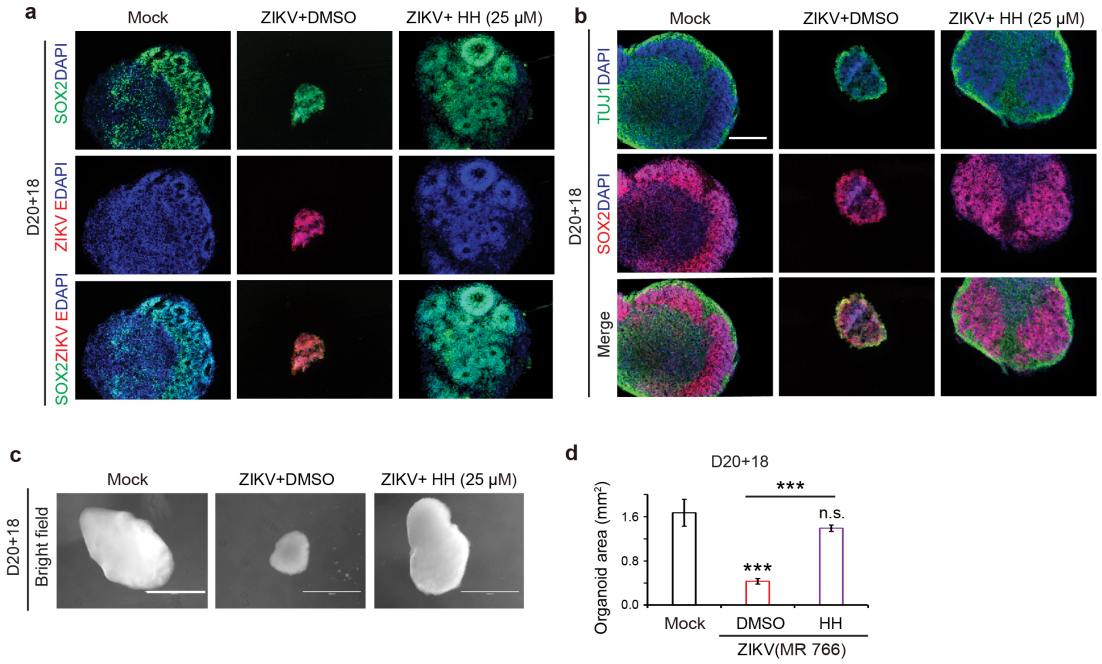


Figure 2.6. HH inhibits ZIKV infection in long-term forebrain organoid cultures. **(a)** D20+18 organoids are highly infiltrated by ZIKV (red), while HH suppressed any detectable ZIKV infection (n=3). Scale bars, 50 μm. **(b)** HH rescues ZIKV-induced structural defects in forebrain organoids. (n=3). Scale bars, 100 μm. (**c,d**) Brightfield images of whole organoids (c) and the quantification of average cross sectional area of D20+18 organoids (d). Values represent mean ± SEM (n=3). Scale bars, 1 mm. *p* values were calculated by unpaired two-tailed Student's t-test. **p*<0.05 and ****p*<0.001. Experimental Design: Ting Zhou, Shuibing Chen, Lorenz Studer. Organoid Culture: Gustav Cederquist. ZIKV infection: Lei Tan. Data Analysis: Ting Zhou, Lei Tan, Gustav Cederquist, Lorenz Studer, Shuibing Chen.

HH or AQ suppresses ZIKV infection *in vivo*. To evaluate the prophylactic activity of drug candidates *in vivo*, HH was administrated to 6-8 week old SCID-beige mice 12 h before they were infected with ZIKV (MR766). HH was injected at a dose of 100 mg/kg/day subcutaneously (**Figure 2.7 a**). The mice were euthanized at 7 dpi. ZIKV vRNA was quantified in multiple organs of ZIKV-infected mice treated with vehicle (**Figure 2.7 b**). The levels of (+) strand ZIKV vRNA in liver, spleen and kidney was low (**Figure 2.7 b**). However, strong ZIKV infection was detected in the brains of many of the infected mice

(**Figure 2.7 b**). Strikingly, HH treatment suppressed significantly suppressed levels of ZIKV vRNA in brain (**Figure 2.7 c**).

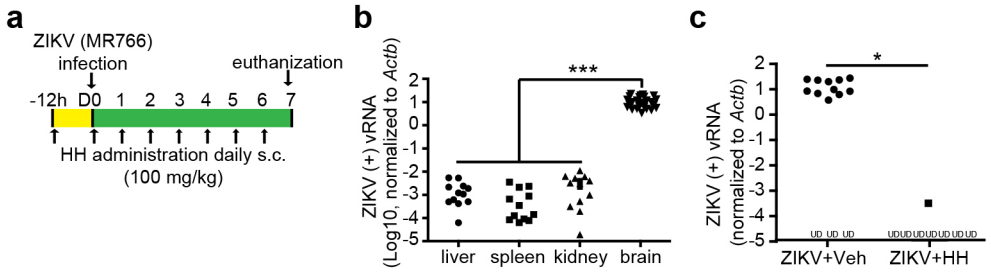
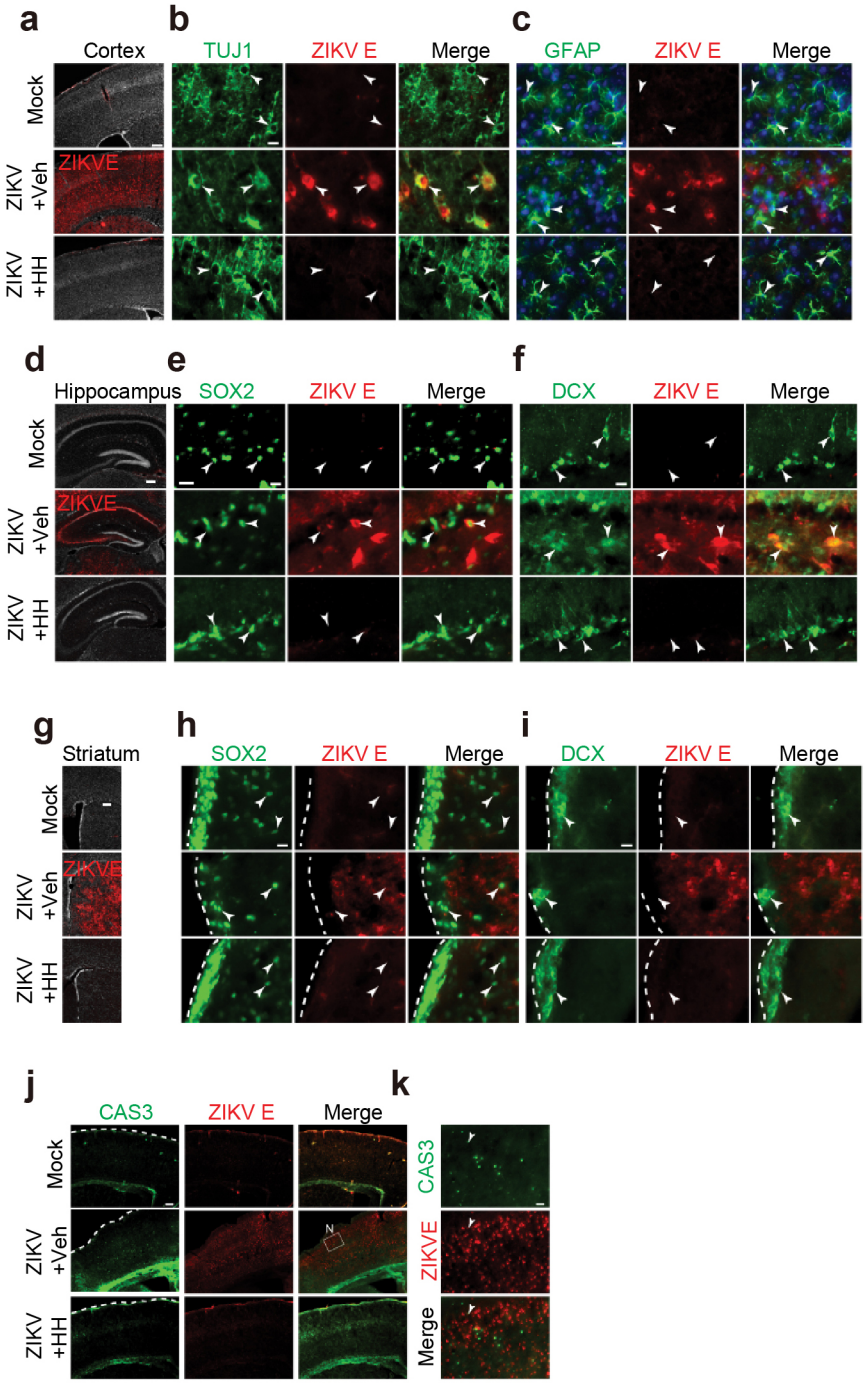


Figure 2.7. HH suppresses ZIKV infection *in vivo*.
(a) Scheme of prophylactic treatment of HH *in vivo*. **(b)** qRT-PCR analysis of (+) strand ZIKV vRNA in liver, spleen, kidney and brain of ZIKV infected mice. n=12 for liver, spleen and kidney and n=20 for brain. **(c)** qRT-PCR analysis of (+) strand ZIKV vRNA in brains of ZIKV-infected mice treated with vehicle (n=14) or HH (n=8). UD, undetectable. All experiments and analysis by Lei Tan, Yujie Fan, Shuibing Chen.

We next determined the brain structures and cell types infected by ZIKV in this *in vivo* system. Immunohistochemical analysis revealed that ZIKV broadly infects forebrain structures, with strong infiltration of the adult cortex (**Figure 2.8 a**). Interestingly, ZIKV was found in many post-mitotic TUJ1^{+} cortical neurons (**Figure 2.8 b**), but in far fewer, if any, GFAP^{+} cortical astrocytes (**Figure 2.8 c**). HH fully inhibited infection of adult cortex (**Figure 2.8 a-c**). ZIKV also strongly infiltrated hippocampal structures, including the cornu ammonis (CA1 and CA3) regions, suggesting infection of post-mitotic neurons (**Figure 2.8 d**). In addition, high magnification images revealed that ZIKV can infect SOX2^{+} neural progenitors (**Figure 2.8 e**) and DCX^{+} immature neurons (**Figure 2.8 f**). HH completely inhibited infection of the adult hippocampus, including immature cells (**Figure 2.8 d-f**). Finally, ZIKV strongly infiltrated the striatum (**Figure 2.8 g**). Interestingly, ZIVK appeared to localize outside the neurogenic subventricular zone (SVZ). High magnification imaging

Figure 2.8. Cellular identity of ZIKV infected cells in vivo.
(a-c) Immunohistochemical analysis of adult cortex using antibodies against ZIKV E (a), βIII-tubulin (b) and GFAP (c). Scale bar, 100 μm in (a) and 10 μm in (b,c). (**d-f**) Immunohistochemical analysis of adult hippocampus including the post-mitotic CA regions using antibodies against ZIKV E (d), SOX2 (e) and DCX (f). Scale bar, 100 μm in (d) and 10 μm in (e,f). (**g-i**) Immunohistochemical analysis of adult striatum using antibodies against ZIKV E (g), SOX2 (h) and DCX (i). Scale bar, 100 μm in (g) and 10 μm in (h,i). (**j,k**) Immunohistochemical analysis of cellular apoptosis in adult cortex antibodies against CAS3. Scale bar, 100 μm in (M) and 10 μm in (N). Experimental Design: Lei Tan, Shuibing Chen. Drug administration: Lei Tan. Histology: Gustav Cederquist. Data analysis: Lei Tan, Gustav Cederquist, Shuibing Chen.

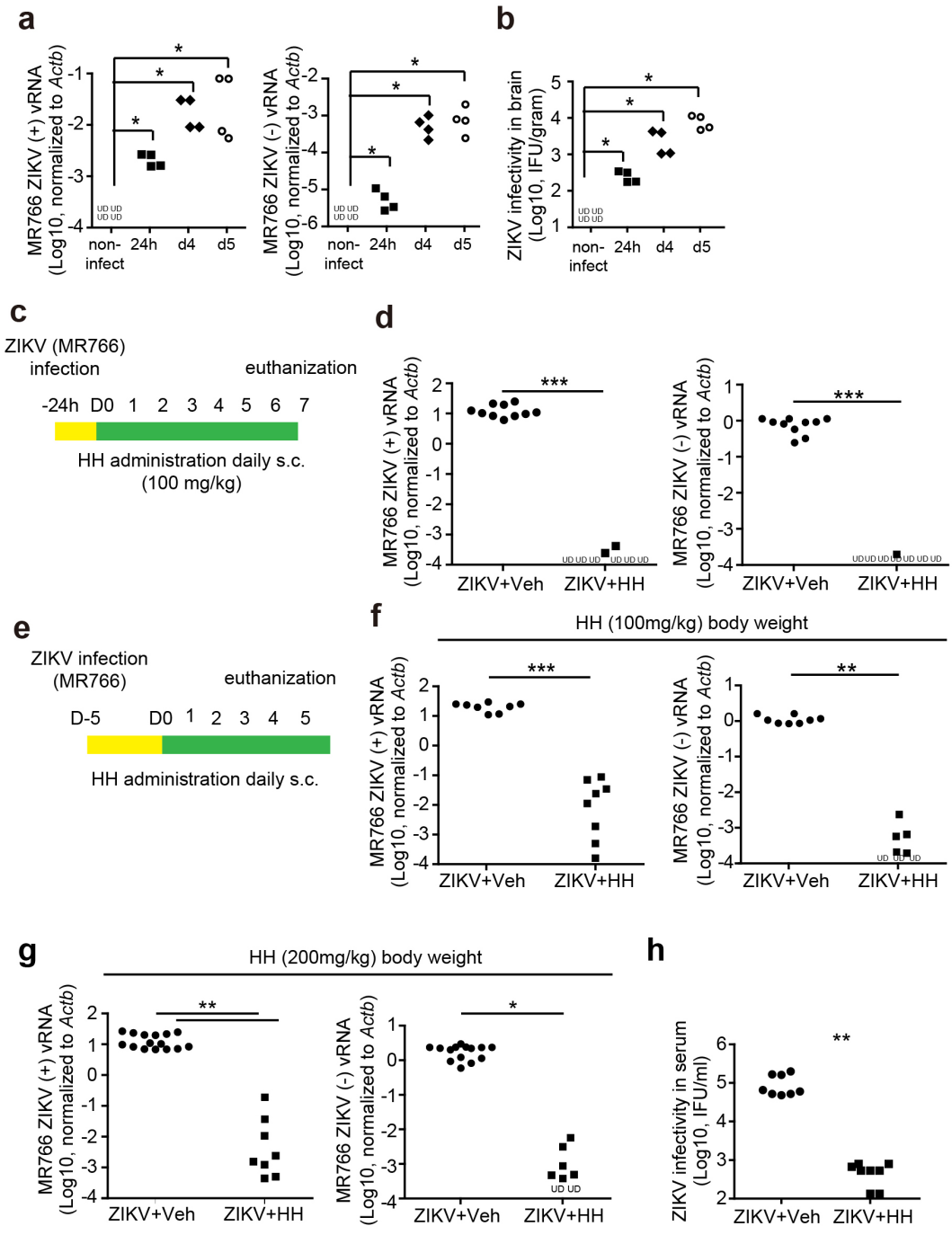

a
Cortex
Mock
ZIKV +Veh
ZIKV +HH
ZIKVE
b
TUJ1
ZIKV E
Merge
c
GFAP
ZIKV E
Merge
d
Hippocampus
e
SOX2
ZIKV E
Merge
f
DCX
ZIKV E
Merge
g
Striatum
h
SOX2
ZIKV E
Merge
i
DCX
ZIKV E
Merge
j
CAS3
ZIKV E
Merge
k
CAS3
ZIKVE
Merge

revealed few ZIKV-infected $SOX2^+$ neural progenitors (**Figure 2.8 h**) or DCX^+ immature neurons (**Figure 2.8 i**). HH fully inhibited infection of the adult striatum (**Figure 2.8 g-i**). Given the strong infiltration of adult brain structures by ZIKV, we analyzed whether ZIKV exerts detrimental cellular effects in the adult nervous system. Immunostaining for CAS3 revealed that ZIKV infection increases the number of apoptotic cells in the cortex (**Figure 2.8 j**). ZIKV induced apoptosis in both infected and non-infected cells (**Figure 2.8 k**). Treatment of mice with HH rescued cellular apoptosis associated with ZIKV infection in the cortex (**Figure 2.8 k**).

To evaluate the therapeutic potential of HH, we first monitored the kinetics of ZIKV infection in adult mouse organs. Increasing levels of (+) strand (**Figure 2.9 a, left**) and replicating (-) strand (**Figure 2.9 a, right**) vRNAs were detected in the brains of ZIKV-infected SCID-beige mice at 24 hpi, 4 and 5 dpi with MR766. Vero re-infection assay was used to confirm infectious particles in the brain (**Figure 2.9 b**). Based on this result, we decided to focus on evaluating HH's effect on mouse brain.

Based on the dynamics of viral replication, we tested the anti-ZIKV potential of HH at 24 hpi and 5 dpi, two time points when ZIKV vRNA is detected. At 24 hpi, the ZIKV-infected mice were randomly separated into two groups and treated with vehicle or HH (100 mg/kg/day subcutaneously) for 7 days (**Figure 2.9 c**). The (+) strand and (-) strand ZIKV vRNA were rarely detectable following HH treatment (**Figure 2.9 d**). In addition, HH treatment at 5 dpi, 4 days after the emergence of detectable ZIKV (**Figure 2.9 e)**,

Figure 2.9. Kinetics of ZIKV infection and treatment *in vivo*.
(**a**,**b**) Kinetics of ZIKV (MR766). Mice were infected with (1 x10^6 PFU) through intraperitoneal injection. Mice were euthanized at 24 hpi, 4 and 5 dpi. vRNA level was analyzed by qRT-PCR (a). Infectivity of viral particles was analyzed by Vero assay (b). n ≥ 4 mice at each time point. (**c**) Scheme of therapeutic intervention with HH 24 hpi (MR766) in (d). **(d)** qRT-PCR analysis of ZIKV vRNA in brains of mice treated with vehicle (n=10) or HH (n=8) for 5 days starting at 24 hpi. **(e)** Scheme of therapeutic intervention with HH 5 dpi (MR766) in (f-h). **(f)** qRT-PCR analysis of ZIKV vRNA in the brains of mice treated with vehicle (n=8) or 100 mg/kg HH (n=8) for 5 days starting at 5 dpi. **(g)** qRT-PCR analysis of ZIKV vRNA in the brains of mice treated with vehicle (n=15) or 200 mg/kg HH (n=8) for 5 days starting at 5 dpi. **(h)** Quantification of ZIKV infectious particles in serum of ZIKV-infected mice (MR766) treated with vehicle (n=8) or 200 mg/kg HH (n=8) for 5 days starting at 5 dpi. *p* values were calculated by one-way repeated measures ANOVA or two-way repeated measures ANOVA with a Bonferroni test. *p<0.05, **p<0.01, ***p<0.001. All experiments and analysis by Lei Tan, Yujie Fan, Shuibing Chen (Weill Cornell).


a
MR766 ZIKV (+) vRNA (Log10, normalized to Actb)
MR766 ZIKV (-) vRNA (Log10, normalized to Actb)
UD UD UD UD
non-infect
24h
d4
d5
b
ZIKV infectivity in brain (Log10, IFU/gram)
c
ZIKV (MR766) infection
euthanization
-24h D0 1 2 3 4 5 6 7
HH administration daily s.c. (100 mg/kg)
d
ZIKV+Veh
ZIKV+HH
e
ZIKV infection (MR766)
euthanization
D-5 D0 1 2 3 4 5
HH administration daily s.c.
f
HH (100mg/kg) body weight
g
HH (200mg/kg) body weight
h
ZIKV infectivity in serum (Log10, IFU/ml)

consistently decreased both (+) strand and replicating (-) strand vRNA levels when compared to the vehicle treated cohort (**Figure 2.9 f**). High dose HH treatment (200 mg/kg/day subcutaneously) applied at 5 dpi also suppressed (+) and (-) strand ZIKV vRNA in the brain (**Figure 2.9 g**). Importantly, high dose HH treatment starting at 5 dpi also significantly suppressed the infectious capacity of ZIKV in mouse serum (**Figure 2.9 h**). Together, these data demonstrated that in immuno-compromised adult hosts, ZIKV has the ability to achieve disseminated infection outside of progenitor zones, where it typically resides. Furthermore, HH has broad anti-ZIKV activity in mouse brain carrying replicating ZIKV, highlighting its therapeutic application.

Discussion.

We successfully carried out a high content chemical screen using an FDA-approved drug library and studied two drug candidates that have anti-ZIKV activity in hNPCs. Both HH and AQ can inhibit ZIKV infection of hNPCs. HH, but not AQ, eliminates ZIKV from hNPCs previously infected with ZIKV. As a result, HH can rescue microcephaly-related defects in human fetal-like forebrain organoid cultures, with little or no evidence of cellular toxicity. Finally, HH exhibits anti-ZIKV activity *in vivo* in the adult mouse brain. Recently, two compounds that have been reported in mouse models, NITD008 (Deng et al., 2016) and 7-deaza-2'-C-methyladenosine (Zmurko et al., 2016), have only been tested for their prophylactic activity, but not therapeutic application on animals with active ZIKV infection. In our study, we demonstrated that HH has strong *in vivo* anti-ZIKV activity, inhibiting virus production and spread in ZIKV-infected immuno-compromised SCID-beige mice.

HH was first isolated from the amaryllidaceae family plant species *Lycoris radiata*. HH was reported to exhibit anti-avian influenza H5N1 viral activity (He et al., 2013). In addition, a chemical analog of HH inhibits Hepatitis C virus, another flavivirus (Chen et al., 2015a). It is worth noting that HH shows no toxicity even at high doses (100 μM *in vitro* and 200 mg/kg *in vivo*). Given the safety and efficacy of HH to suppress ZIKV infection in hNPCs, human forebrain organoids, and adult mice, HH is a highly promising drug candidate for further clinical development to treat ZIKV-infected patients across the lifespan.

In addition to HH, we identified another FDA approved drug, AQ, also called Amodiaquine, as an anti-ZIKV compound. The WHO recommended AQ as an optional treatment of uncomplicated malaria (Olliaro and Mussano, 2003). AQ has activity against another flavivirus, Dengue virus, by inhibition of Dengue virus type 2 replication and infectivity (Boonyasuppayakorn et al., 2014). We found that AQ effectively inhibits ZIKV infection in NPCs *in vitro* and in a mouse model. However, AQ exhibited high cytotoxicity in brain organoid cultures, which raises safety concerns for using this drug during pregnancy.

Despite the ZIKV pandemic, very few options are available to treat this potentially devastating infection. Chemical-based drugs may provide a first response option to treat this epidemic. Our findings provide a strong drug candidate for the treatment of ZIKV infection, in addition to indicating novel targets for drug development against other flaviviruses, including West Nile Virus, Dengue Virus, and yellow fever virus, which all cause severe illness.

CHAPTER THREE: SPECIFICATION OF POSITIONAL IDENTITY IN FOREBRAIN ORGANOIDS

Abstract

Human brain organoids are self-organizing, three-dimensional neural tissues derived from pluripotent stem cells. While organoids recapitulate histological features of the human brain, they lack a reproducible topographic organization, which is fundamental to brain structure and function. During development, spatial topography is determined by the graded activity of signaling molecules, which are released from discrete signaling centers. We hypothesized that introduction of a signaling center into forebrain organoids would specify the positional identity of neural tissue in a distance-dependent manner. Here, we present a system to trigger a sonic hedgehog (SHH) protein gradient in developing forebrain organoids that enables ordered self-organization along dorso-ventral and antero-posterior positional axes. SHH-patterned forebrain organoids establish major subdivisions of the forebrain,

and strikingly, these subdivisions are positioned with *in vivo*-like topography. Consistent with its behavior *in vivo*, SHH exhibits long-range signaling activity in organoids. Finally, we use SHH-patterned cerebral organoids as a tool to study the potential teratogenic effects of statins during early fetal brain development. Together, our work identifies inductive signaling as an effective organizing strategy to recapitulate *in vivo* topography in human brain organoids.

Introduction

Human pluripotent stem cells (hPSCs) have the intrinsic capacity to self-organize into multicellular, organ-like structures called organoids (Kelava and Lancaster, 2016; Sasai et al., 2012). Brain organoids recapitulate the cellular diversity and micro-architectural features characteristic of discrete brain regions, providing unprecedented opportunities to model human brain development and disease (Lancaster and Knoblich, 2014b; Lancaster et al., 2013; Quadrato et al., 2016). However, in these organoids, individual brain regions are ordered randomly, non-reproducibly and lack the characteristic *in vivo* anterior-posterior, dorso-ventral, and medio-lateral positioning that supports the emergence of complex brain structure and function (Kelava and Lancaster, 2016; Lancaster et al., 2013; Renner et al., 2017). The absence of a defined topography is a crucial shortcoming of current brain organoid technologies. Region-specific organoids attempt to overcome positional heterogeneity by restricting cellular identities to a single brain area such as the forebrain, midbrain, or hindbrain (Jo et al., 2016; Kadoshima et al., 2013; Muguruma et al., 2015). More recently, several groups have created a dorso-ventral axis by fusing dorsal and ventral forebrain organoids. While these

strategies recapitulate some processes associated with long-range tissue interactions, such as cell migration, they fail to capture the full continuum of regional diversity encompassed by the human forebrain.

During development, topographic maps are generated by gradients of signaling activity across neural tissue, which allow cells to acquire discrete regional identities as a function of their specific position (O'Leary et al., 2007; Sagner and Briscoe, 2017). In contrast, brain organoids are generated either in the absence of extrinsic signaling molecules or with morphogens added to the culture medium resulting in uniform, rather than graded, signaling cues. Exposure to uniform extrinsic signals results in restricted, area-specific organoids that lack the regional complexity of the nervous system (Bagley et al., 2017; Eiraku et al., 2011; Jo et al., 2016; Kadoshima et al., 2013; Muguruma et al., 2015; Qian et al., 2016; Suga et al., 2011). Organoids devoid of any extrinsic signals exhibit a high degree of complexity, yet are often highly variable with respect to their regional identity and lack an ordered topographical organization (Lancaster and Knoblich, 2014a; Lancaster et al., 2013; Quadrato et al., 2017; Renner et al., 2017).

Results

Sonic Hedgehog (SHH) is a signaling factor whose graded expression specifies the spatial organization of discrete progenitor domains across the neuraxis (Jessell, 2000; Lupo et al., 2006). To test the hypothesis that introduction of a SHH signaling gradient into developing forebrain organoids could specify positional domains, we engineered an inducible SHH-expressing hPSC line (iSHH) that could be embedded at one pole of a developing

organoid (**Figure 3.1 a**). The iSHH line was generated by TALEN-mediated gene targeting into the *AAVS1* locus of an hPSC line with constitutive RFP expression (Fattahi et al., 2016; Gonzalez et al., 2014). A constitutively active reverse tetracycline transactivator is expressed from one allele of the *AAVS1* locus, while full-length human *SHH* is expressed from the other allele, under the control of a tetracycline response element (TRE) (**Figure 3.1 a**). This strategy resulted in titratable *SHH* expression (**Figure 3.1 b**) with commensurate level of post-translational palmitoylation (**Figure 3.1 c**). The iSHH line induced robust expression of the SHH target gene *FOXA2* during neural differentiation in the presence of doxycycline, indicating appropriate biological activity (**Figure 3.1 d**).

We next envisioned a strategy to trigger a SHH gradient in forebrain organoids (**Figure 3.2 a).** First, we developed a method to embed iSHH cells at one pole of an hPSC spheroid, mimicking a developmental organizer. 1,000 iSHH cells were seeded in low-attachment round bottom microwells and allowed to aggregate for 24 hours (day -2 to -1). The next day 10,000 wild-type H9 hPSC were seeded on top of the iSHH cells (day -1 to 0) and allowed to aggregate for another 24 hours, reproducibly resulting in a chimeric iSHH-H9 spheroid in which a small cluster of iSHH cells are embedded within a larger spheroid (**Figure 3.2 b**, >90% efficiency). We refer to these chimeric 3D hPSC cultures as SHH-spheroids.

For the specific induction of forebrain identity we devised a strategy that combines aspects of several previous organoid induction protocols. In brief, SHH-spheroids were cultured in the presence of inhibitors of TGFβ, BMP, and WNT (3-inhibitor protocol) for 6-8 days to promote specification of anterior forebrain identity (Kadoshima et al., 2013). After 6-8 days of differentiation,

spheroids were embedded in matrigel droplets to promote neuroepithelial organization and moved to an orbital shaker upon outgrowth of neuroepithelial buds, which was typically 4-6 days (Lancaster and Knoblich, 2014a; Lancaster et al., 2013) (**Figure 3.2 c**). Doxycycline (400ng/ml) was added to the differentiation medium starting at day 0 to induce an asymmetric SHH signal. Under those conditions, the iSHH-organizer typically remained positioned at one end of the developing organoid, though we observed shedding of small clusters of organizer cells into distal territories in 25% of instances (**Figure 3.2 d**).

Figure 3.1. Generation of SHH inducible hPSC line.
(a) Strategy to specify positional identity in forebrain organoids. SHH exhibits a ventral-high to dorsal-low gradient in the developing forebrain. A doxycycline-inducible SHH (iSHH) hPSC line was engineered using TALEN-mediated homologous recombination. The iSHH line could be positioned at one pole of developing forebrain organoids. **(b)** qRT-PCR analysis shows robust and titratable induction of SHH transcript expression in iSHH but not H9 hPSC line. Individual replicates plotted on bar graphs, mean ± S.D. One way ANOVA with Dunnett Test. *** P=0.0003, **** P=0.0001. N=3 hPSC cultures for each iSHH condition, N=4 hPSC cultures for each H9 condition. **(c)** Schematic of [I^{125}]Iodopalmitate labeling experiment (left). SHH produced from the iSHH line is palmitoylated at levels commensurate to the amount of SHH protein expressed. **(d)** qRT-PCR analysis of SHH and FOXA2, a SHH target gene, during neural differentiation shows that iSHH expresses biologically active protein, comparable to levels from hPSC-derived floorplate. Individual replicates plotted on bar graphs, mean ± S.D. One way ANOVA with Dunnett Test. ***P<0.001, **P<0.01. N=3 differentiations for all conditions**.** Generation and characterization of iSHH line: Gustav Cederquist, Jason Tchieu, Lorenz Studer. [I^{125}]Iodopalmitate labeling: James Asciolla, Marilyn Resh.

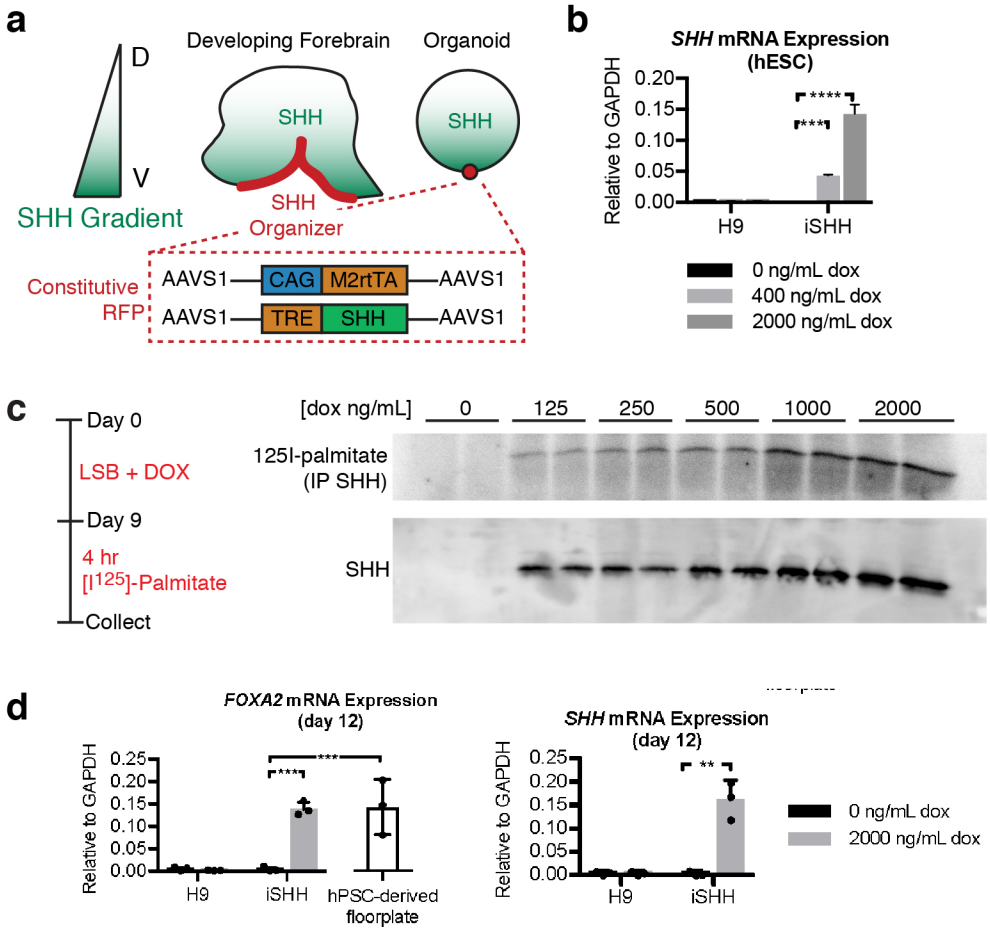
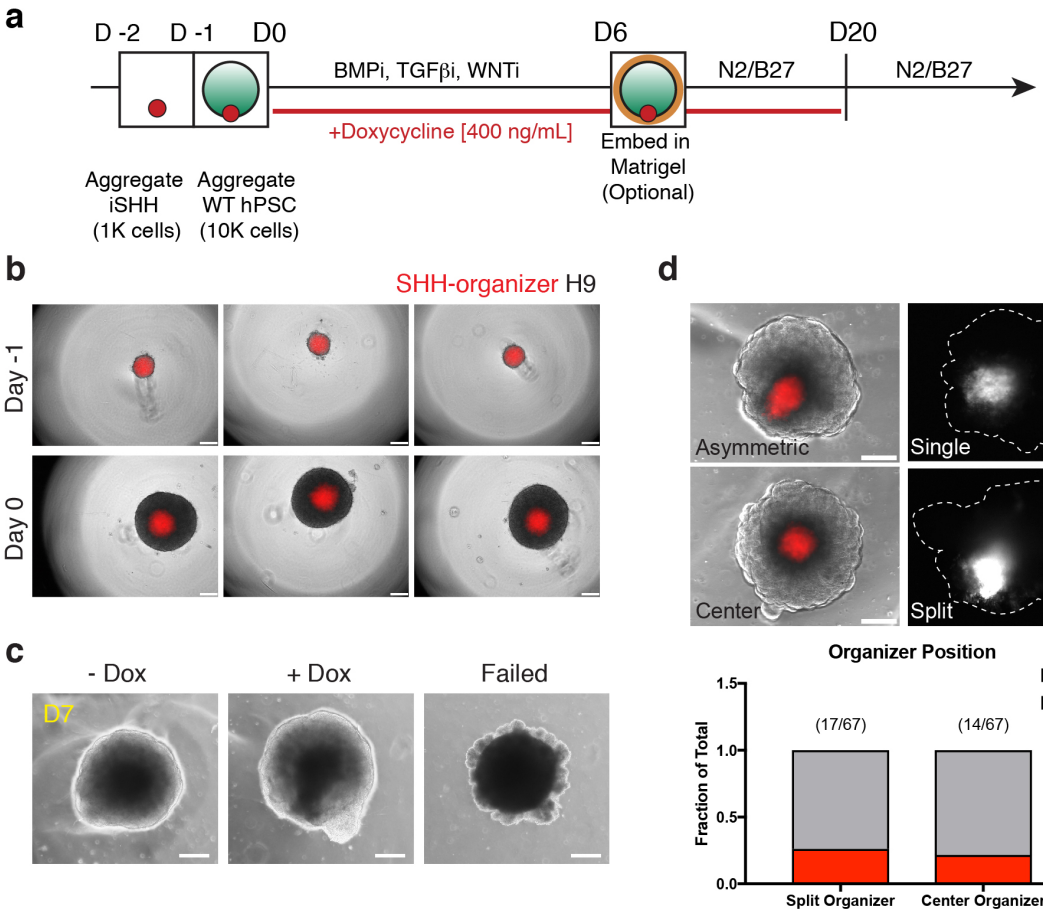

a
D
V
SHH Gradient
Developing Forebrain
Organoid
SHH
SHH
SHH
Organizer
Constitutive
RFP
AAVS1
CAG
M2rtTA
AAVS1
AAVS1
TRE
SHH
AAVS1
b
SHH mRNA Expression
(hESC)
Relative to GAPDH
0.20
0.15
0.10
0.05
0.00
****
***
H9
iSHH
0 ng/mL dox
400 ng/mL dox
2000 ng/mL dox
c
Day 0
LSB + DOX
Day 9
4 hr
[I125]-Palmitate
Collect
[dox ng/mL]
0
125
250
500
1000
2000
125I-palmitate
(IP SHH)
SHH
d
FOXA2 mRNA Expression
(day 12)
Relative to GAPDH
0.25
0.20
0.15
0.10
0.05
0.00
****
***
H9
iSHH
hPSC-derived
floorplate
SHH mRNA Expression
(day 12)
Relative to GAPDH
0.25
0.20
0.15
0.10
0.05
0.00
**
H9
iSHH
0 ng/mL dox
2000 ng/mL dox

Figure 3.2. Protocol to generate SHH-dependent topographic patterning of forebrain organoids

(a) Schematic of the method for generating forebrain organoids with a SHH protein gradient. Timeline and small molecules used for neural induction are shown. The specific timing of matrigel embedding can exhibit slight variation among differentiations. **(b)** Reproducibility of organizer plating and formation of SHH-H9 spheroids. 1,000 iSHH cells are plated in low-attachment microwells and allowed to aggregate overnight. 10,000 wildtype hPSCs are plated on top of the iSHH organizer cells. **(c)** Spheroids are embedded in matrigel, and the day of embedding is critical to efficient neuroepithelial growth. SHH organoids (no dox) embedded on day 5 exhibit no neuroepithelial growth, while SHH organoids (no dox) embedded on day 6 exhibit efficient neuroepithelial formation. **(d)** Typically, the iSHH organizer remains clustered at one pole during differentiation, though in ~25% of instances the organizer shed small clusters of cells into distal sites. Scale bars: 200μm. Experimental design and data analysis: Gustav Cederquist, Lorenz Studer; Organoid Culture: Gustav Cederquist.


a
D -2
D -1
D0
D6
D20
BMPi, TGFβi, WNTi
N2/B27
N2/B27
+Doxycycline [400 ng/mL]
Embed in Matrigel (Optional)
Aggregate iSHH (1K cells)
Aggregate WT hPSC (10K cells)
b
SHH-organizer H9
Day -1
Day 0
c
- Dox
+ Dox
Failed
D7
d
Asymmetric
Single
Center
Split
Organizer Position
No
Yes
(17/67)
(14/67)
Fraction of Total
1.5
1.0
0.5
0.0
Split Organizer
Center Organizer

To assess whether this strategy generates a gradient of SHH protein, we performed immunocytochemistry, using tyramide signal amplification, at day 4 of the differentiation. In the absence of doxycycline, no SHH protein was detectable. In the presence of doxycycline, high levels of SHH expression were detected in the region of the iSHH cells, and the abundance of SHH protein was reduced as a function of distance from the SHH-organizer (**Figure 3.3 a**).

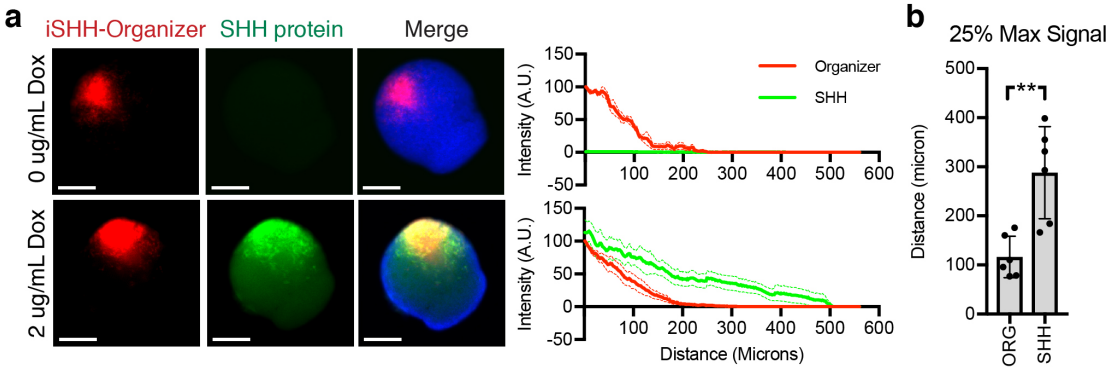


Figure 3.3. SHH protein gradient in 3D culture.
(a) Visualization of SHH protein gradient using tyramide signal amplification (TSA). A line scan was used to quantify SHH protein signal. Graph shows mean ± S.D. Intensity of SHH signal (y-axis, green) normalized to maximum iSHH fluorescence. Origin (X-axis) set to position of maximum iSHH fluorescence. **(b)** Quantification of the distance at which the iSHH-organizer signal (ORG) and SHH protein signal (SHH) reach 25% of maximum intensity demonstrates that the SHH protein gradient extends beyond the limit of the organizer cells. Student two-tailed t-test. N=6 spheroids, 2 batches. All experiments and analysis by Gustav Cederquist and Lorenz Studer.

Functional subdivisions within the mammalian forebrain are located at discrete positions within a Cartesian coordinate system, whose axes are arranged in the dorso-ventral, medio-lateral, and antero-posterior directions. Each of the dorso-ventral and antero-posterior domains can be identified by the expression of characteristic transcription factors. PAX6 is expressed in the dorsal forebrain and NKX2.1 in the ventral forebrain (**Figure 3.4 a**). FOXG1 is expressed in the anterior forebrain, or telencephalon, while it is absent from the posterior forebrain, or diencephalon. In contrast, OTX2, SIX3, and TCF7L2

are expressed in the diencephalon (**Figure 3.4 b,c**). When we differentiated H9 spheroids or SHH-spheroids in the absence of doxycycline they predominantly expressed PAX6 and FOXG1 within self-organized neuroepithelium, in agreement with previous findings(Kadoshima et al., 2013) (**Figure 3.4 a,b**). This indicates that the default identity of our 3D cultures is dorsal-anterior forebrain, which gives rise to the neocortex. SHH-spheroids grown in the presence of doxycycline (400 ng/ml) induced NKX2.1 near the organizer, while PAX6 expression was suppressed (**Figure 3.4 a**).

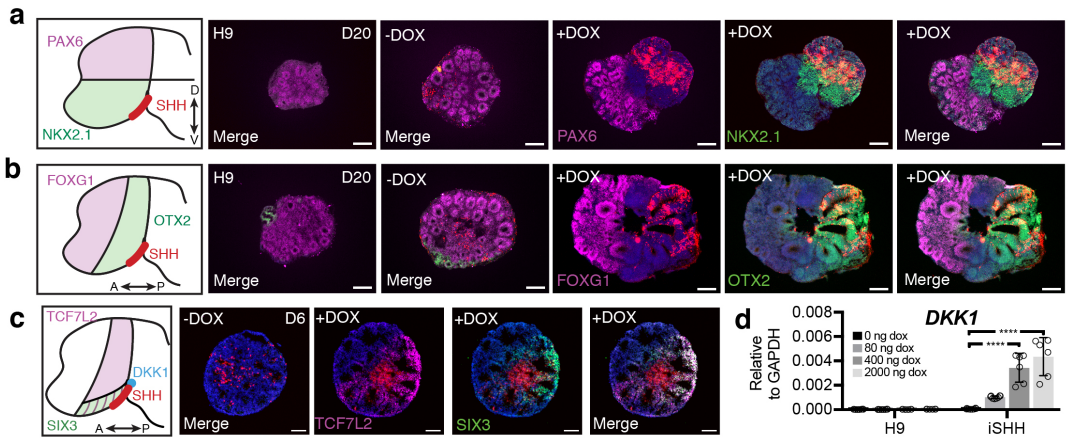


Figure 3.4. Specification of developmental axes in SHH-organoids.
(a) Schematic illustrating that PAX6 and NKX2.1 define dorsal and ventral forebrain, respectively. H9 organoids (H9) or SHH-organoids grown without doxycycline (-DOX) largely express PAX6 without detectable NKX2.1 expression. SHH-organoids grown in the presence of doxycycline (400 ng/ml, +DOX) suppress PAX6 and induce NKX2.1 near the iSHH organizer. H9, N=7 organoids; -DOX, N=8 organoids; +DOX, N=11 organoids. Samples are from at least 2 batches. **(b)** Schematic illustrating that FOXG1 and OTX2 define anterior and posterior forebrain domains. H9 organoids or SHH-organoids grown without doxycycline largely express FOXG1 with only small amounts of OTX2 expression. SHH-organoids grown in the presence of doxycycline (400 ng/ml) suppress FOXG1 and induce OTX2 near the iSHH organizer. H9, N=8 organoids; -DOX, N=8 organoids; +DOX, N=12 organoids. Samples are from at least 2 differentiations. **(c)** Schematic illustrating diencephalic expression of SIX3, TCF7L2, and DKK1. SHH-organoids grown in doxycycline (400 ng/ml) induce SIX3 and TCF7L2 near organizer tissue. N = 9, -DOX and +DOX. **(d)** Dose-dependent activation of DKK1 by SHH induction. N=6-8 organoids from 2 batches for all conditions. Individual replicates plotted on bar graphs, mean ± S.D. One-way ANOVA, $P < 0.0001$, followed by post-hoc Tukey test for multiple comparisons. **** $P < 0.0001$, ** $P < 0.01$. Representative images shown in all panels. Scale Bars: 200 μm. All experiments and analysis by Gustav Cederquist and Lorenz Studer.

Surprisingly, FOXG1 expression was also suppressed near the organizer, while OTX2 expression was maintained (**Figure 3.4 b**), suggesting that SHH can lead to posterior forebrain specification. We therefore examined the expression of TCF7L2, which is expressed in the diencephalon (Merchan et al., 2011; Shiraishi et al., 2017) and SIX3, which is expressed in optic recess and hypothalamus (Oliver et al., 1995). Both TCF7L2 and SIX3 were induced near organizer tissue (**Figure 3.4 c**), supporting a role for SHH in anterior-posterior patterning, as well as dorsoventral patterning. SHH exhibited dose-dependent induction of *DKK1* (**Figure 3.4 d**), a secreted antagonist that shapes the anterior-posterior WNT gradient *in vivo,* is commonly induced in regions of high WNT/β–catenin activity (Niida et al., 2004; Shinya et al., 2000), and is a critical regulator of anterior-posterior patterning (Houart et al., 2002; Kiecker and Niehrs, 2001). It is thus interesting to hypothesize that SHH might exert anterior-posterior patterning activity via regulation of WNT signaling(Maroof et al., 2013).

Given that a SHH protein gradient could specify distinct dorso-ventral and antero-posterior positional domains, we next sought to determine if the resulting forebrain subdivisions are aligned according to an anatomically appropriate topography. The position of the presumptive forebrain subdivisions *in vivo* can be defined by their distance from the SHH source in the hypothalamus. The neocortex (PAX6+/FOXG1+) represents the anterior and dorsal subdivision of the forebrain which is located most distal from the SHH source. Just ventral to the neocortex is the lateral ganglionic eminence (LGE) characterized by the co-expression expression of GSH2+/FOXG1+, followed by the medial ganglionic eminence (MGE) characterized by co-

expression of NKX2.1+/FOXG1+, the antero-dorsal hypothalamus characterized by expression of NKX2.2+, and finally the ventro-posterior hypothalamus characterized by the expression of NKX2.1+/FOXG1-, representing the brain region located most ventral and proximal to the SHH source (**Figure 3.5 a**).

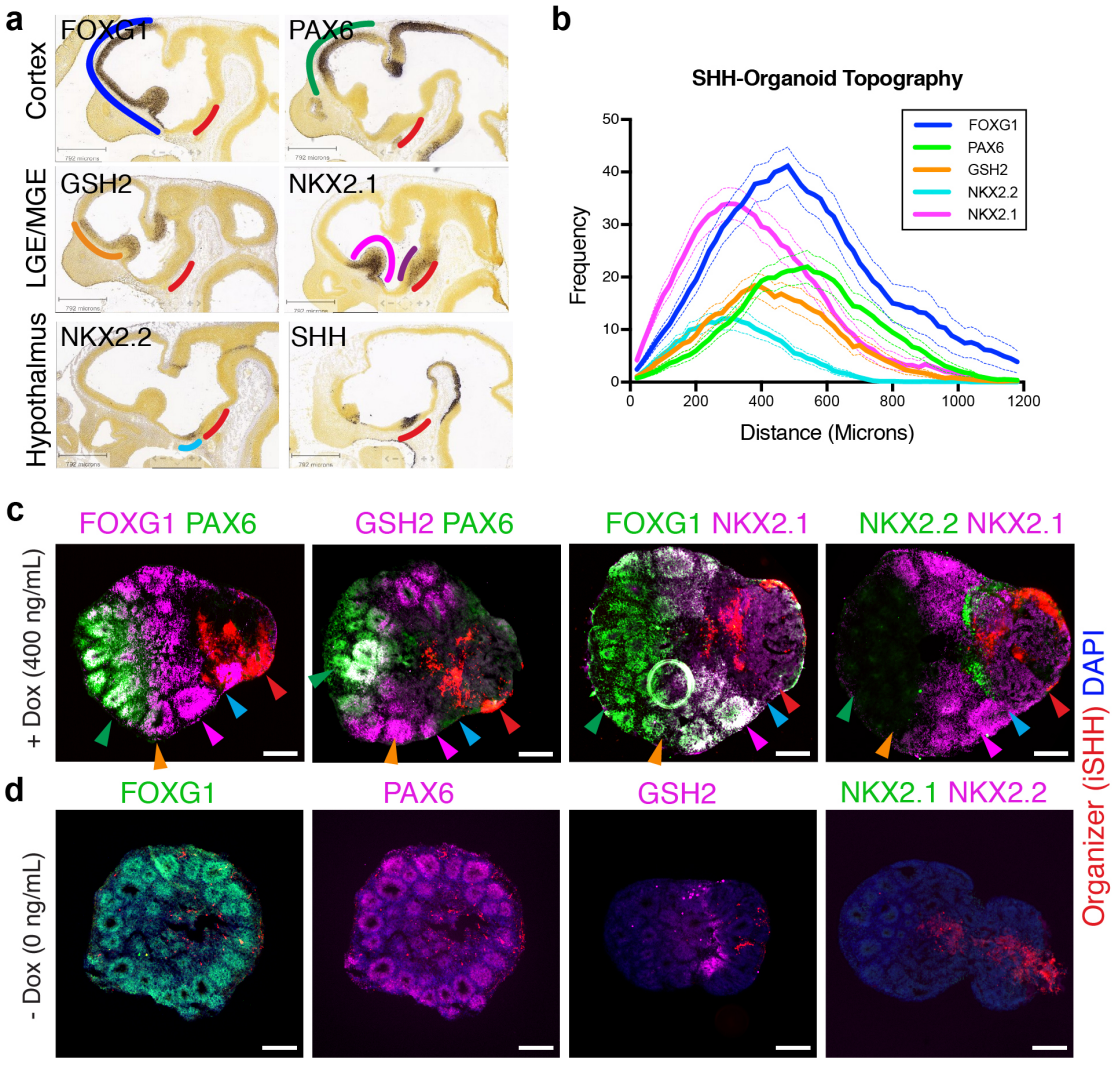
When differentiated in the presence of doxycycline (400 ng/ml), SHH-spheroids formed at least five topographically distinct presumptive forebrain domains. Quantification of the distance of forebrain domains from the SHH-organizer, identified by protein expression, revealed that the positioning of these domains precisely mimics the pattern observed *in vivo* (**Figure 3.5 b**). Topographic organization could be triggered in at least five additional hPSC or iPSC lines using the same H9-based iSHH organizer, though the specific growth rates and size of regional domains could differ (**Figure 3.6**). We next identified presumptive neocortical (FOXG1+/PAX6+), LGE (GSH2+), MGE (FOXG1+/NKX2.1+), anterior hypothalamic (NKX2.2+) and ventro-posterior hypothalamic (NKX2.1+/FOXG1-) regions by immunocytochemistry (**Figure 3.5 c**). The resulting topographic patterning is in contrast to the patterning exerted by bath application of SHH agonist to forebrain organoids, which results in subpallium-restricted identities (Bagley et al., 2017; Birey et al., 2017). SHH-spheroids differentiated in the absence of doxycycline expressed markers of dorsal-anterior forebrain, including PAX6 and FOXG1, with a small proportion of GSH2 expressing cells. These 3D cultures in the absence of doxycycline did not express the ventral identity genes NKX2.1 and NKX2.2 (**Figure 3.5 d**).

Interestingly, tissue cytoarchitecture differed between areas that are proximal (hypothalamic-like) and distal (telencephalic-like) to the SHH source. At day

20, PAX6+ distal neuroepithelia acquired circular, rosette-like morphologies (**Figure 3.7 a**), consistent with the self-organizing radial structures described for forebrain organoids (Kadoshima et al., 2013; Lancaster et al., 2013). In contrast, the NKX2.1+ proximal region contained many thin and highly extended neuroepithelia (**Figure 3.7 a**) as observed previously for 3D structures of early hypothalamic lineages (Merkle et al., 2015; Suga et al., 2011). This appears to be restricted to diencephalic NKX2.1+ regions, as presumptive MGE domains (FOXG1+/NKX2.1+) typically acquired circular, rosette-like morphologies (**Figure 3.7 b**). The organizer tissue itself expression of NKX2.1 in at least a subset of the organizer cells, suggesting hypothalamic identity (Blaess et al., 2014) (**Figure 3.7 c**). Our data indicate that an asymmetric SHH cue enables the ordered patterning and topographical organization of brain regions without interfering in the self-organization process, which leads to the establishment of region-specific tissue microarchitectures. We term these 3D cultures SHH-organoids.

To explore the degree to which topography is maintained over time, SHH-organoids were cultured for up to 70 days. Typically the OTX2+ domain was located proximal to the organizer, while the FOXG1+ domain was located distally (**Figure 3.8 a**). Radially organized PAX6, TBR2, TBR1 cerebral cortex-like tissue and DARPP32+ striatum-like tissue emerged in distal domains (**Figure 3.8 b-d**), while hypothalamic-like tissue, expressing LHX6, OTP, POMC, and TH, was found in the immediate vicinity of the organizer cells (**Figure 3.8 d-f**). In some instances organoid topography was more difficult to discern because the organizer seemed to have dispersed throughout the tissue. In addition, tissue necrosis in the center of those larger organoids may degrade organization over time.

Figure 3.5. *In vivo*-like topographic organization of major forebrain subdivisions in SHH-organoids.
(a) In situ hybridization images depicting forebrain topography. The neocortex co-expresses FOXG1 (blue) and PAX6 (green), and is located distal to the SHH source (red). The lateral ganglionic eminence (LGE) co-expresses FOXG1 and GSH2 (orange), and is just ventral to the neocortex. The medial ganglionic eminence (MGE) co-expresses FOXG1 and NKX2.1 (light purple), and is just ventral to the LGE. NKX2.1 is also expressed in the ventro-posterior hypothalamus (dark purple), and is distinguished from the MGE by exclusion of FOXG1 expression. The anterior hypothalamus expresses NKX2.2 (cyan) and resides in the gap between MGE and Hypothalamic NKX2.1 expression. SHH (red) is strongly expressed in the ventro-posterior hypothalamus, and is also weakly expressed in the ventral MGE. Images are from © 2008 Allen Institute for Brain Science. Allen Developing Mouse Brain Atlas. Available from developingmouse.brain-map.org. **(b)** Quantification of position of regional domains in SHH organoids (400 ng/ml doxycycline). Immunofluorescent signals of each regional marker is plotted as a function of distance from the SHH organizer (see methods for details of quantification). Mean ± S.E.M. N=30 organoids from 6 batches. **(c)** Sections from a single organoid showing that SHH-organoids develop with *in vivo*-like topography. A neocortex-like domain that co-expresses FOXG1 and PAX6 is distal to the SHH organizer (red). A GSH2 domain is observed partially overlapping with and adjacent to the PAX6 domain. A FOXG1/NKX2.1 co-expression domain occupies an intermediate position between the neocortex-like domain and the SHH organizer. Finally, NKX2.2 and NKX2.1 are expressed in the FOXG1 negative territory, with NKX2.1 most proximal to the SHH organizer. Colors of arrowheads correspond to regional identities in panel 2a and delineate presumptive forebrain domains. **(d)** SHH-organoids grown without doxycycline treatment mainly express FOXG1 (N=6 organoids) and PAX6 (N=6 organoids), with some GSH2 expression (N=4 organoids). No NKX2.1 or NKX2.2 is observed (N=3 organoids). All experiments and analysis by Gustav Cederquist and Lorenz Studer.


a
Cortex
LGE/MGE
Hypothalmus
FOXG1
PAX6
GSH2
NKX2.1
NKX2.2
SHH
b
SHH-Organoid Topography
Frequency
Distance (Microns)
FOXG1
PAX6
GSH2
NKX2.2
NKX2.1
c
FOXG1 PAX6
GSH2 PAX6
FOXG1 NKX2.1
NKX2.2 NKX2.1
+ Dox (400 ng/mL)
d
FOXG1
PAX6
GSH2
NKX2.1 NKX2.2
- Dox (0 ng/mL)
Organizer (iSHH) DAPI

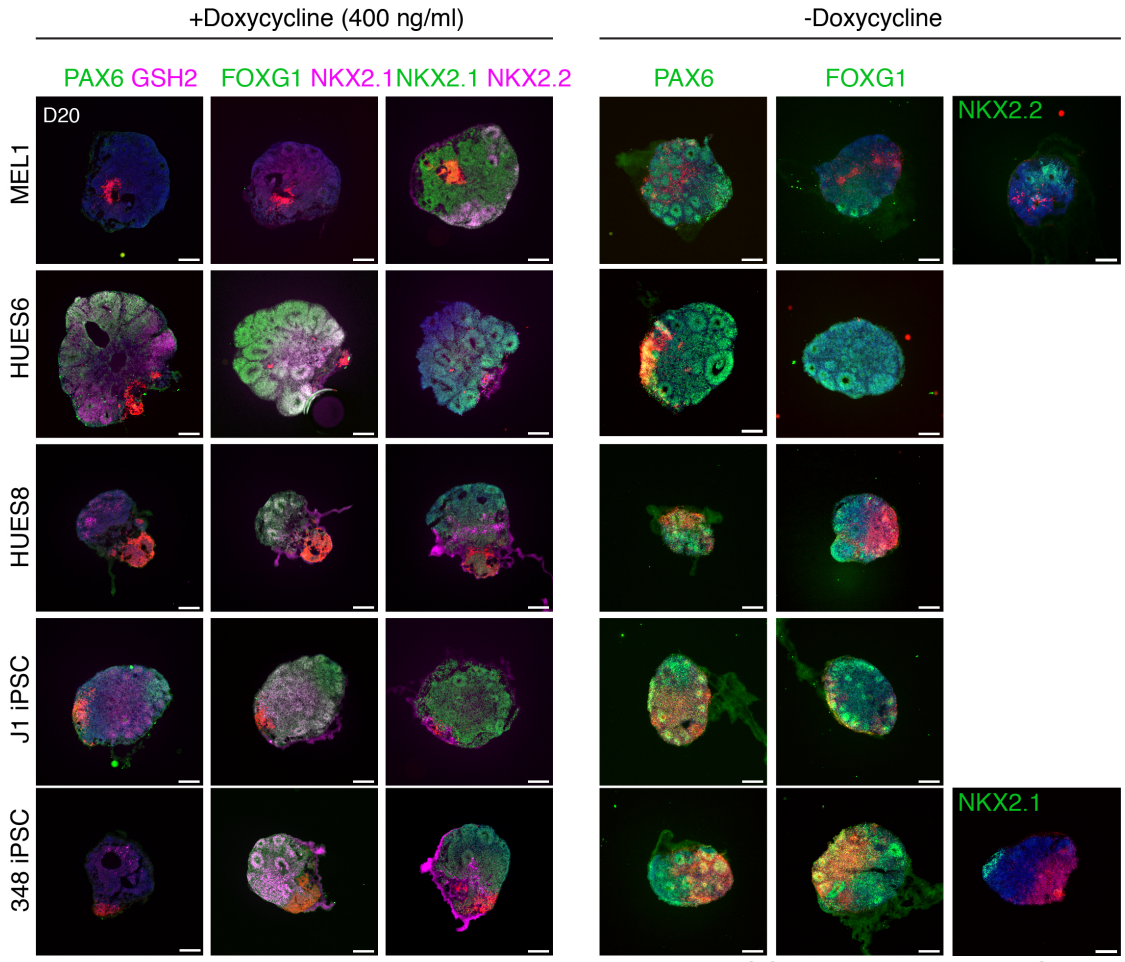


Figure 3.6. Topographic patterning in 3 additional hPSC lines (MEL1, HUES6, HUES8) and 2 iPSC lines (J1 and 348).

The iSHH organizer can induce distinct regional domains that emerge in the anatomically correct topographic order. However, the size of domains and overall growth rate of organoids may differ between lines. Without doxycycline all lines are predominantly PAX6 and FOXG1 positive, though sparse induction of NKX2.2 and NKX2.1 was observed in MEL1 and 348 lines. +Doxycycline, N=8 organoids, 2 batches for all lines; -Doxycycline, N=4 organoids, 1 batch. Scale bars: 200μm. All experiments and analysis by Gustav Cederquist and Lorenz Studer.

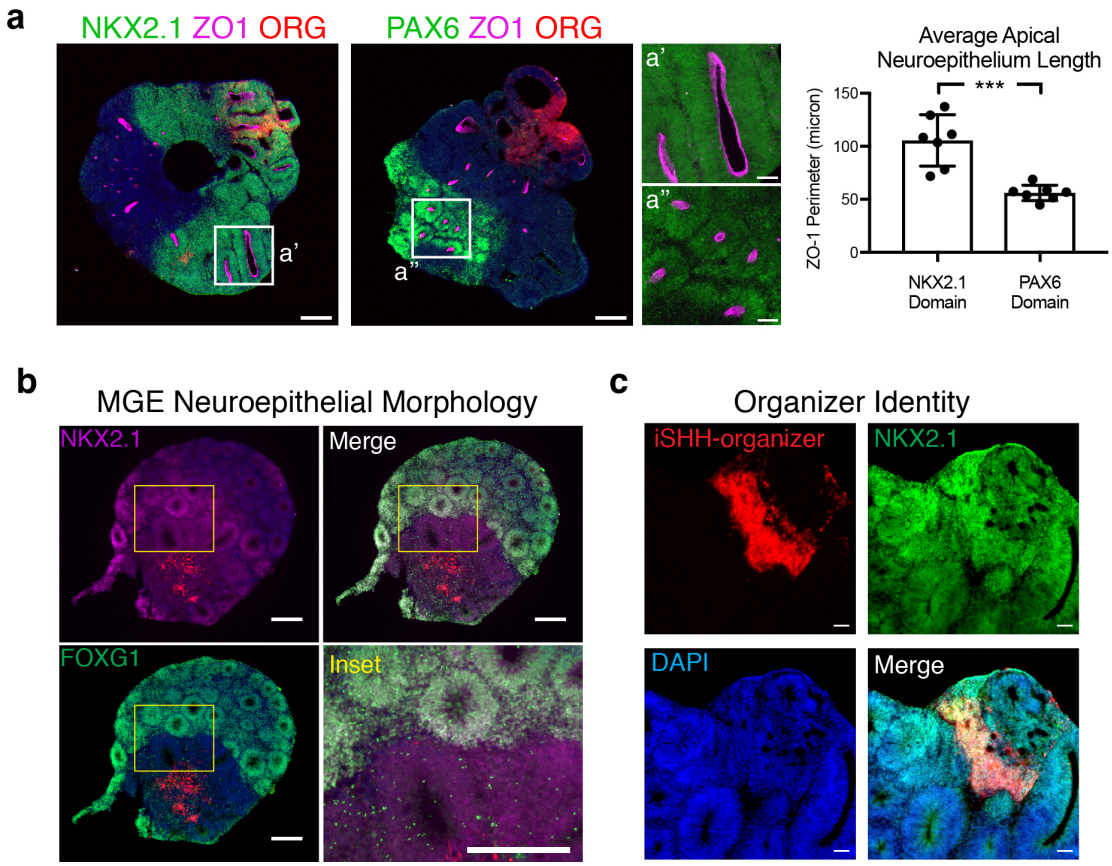


Figure 3.7. Differential local self-organization in SHH-organoids.
(a) The length of ZO-1 apical neuroepithelia differs between PAX6+ and NKX2.1+ regions. PAX6+ regions contain circular, rosette-like neuroepithelia, while NKX2.1+ regions contain highly-extended neuroepithelia. High-magnification images correspond to insets. Quantification shows average neuroepithelium length in PAX6 versus NKX2.1 domains. Dots represent individual organoids. Student two-tailed t-test. n = 7 organoids, 2 batches for each condition. **(b)** MGE-like neuroepithelium (FOXG1+/NKX2.1+) acquires a circular, rosette-like structure suggestion a radial organization. **(c)** iSHH organizer cells (red) at least partially express NKX2.1 (green) and are negative for FOXG1, suggesting hypothalamic identity. Scale bars: 200 mm (Low magnification), 50 mm (High magnification). All experiments and analysis by Gustav Cederquist and Lorenz Studer.

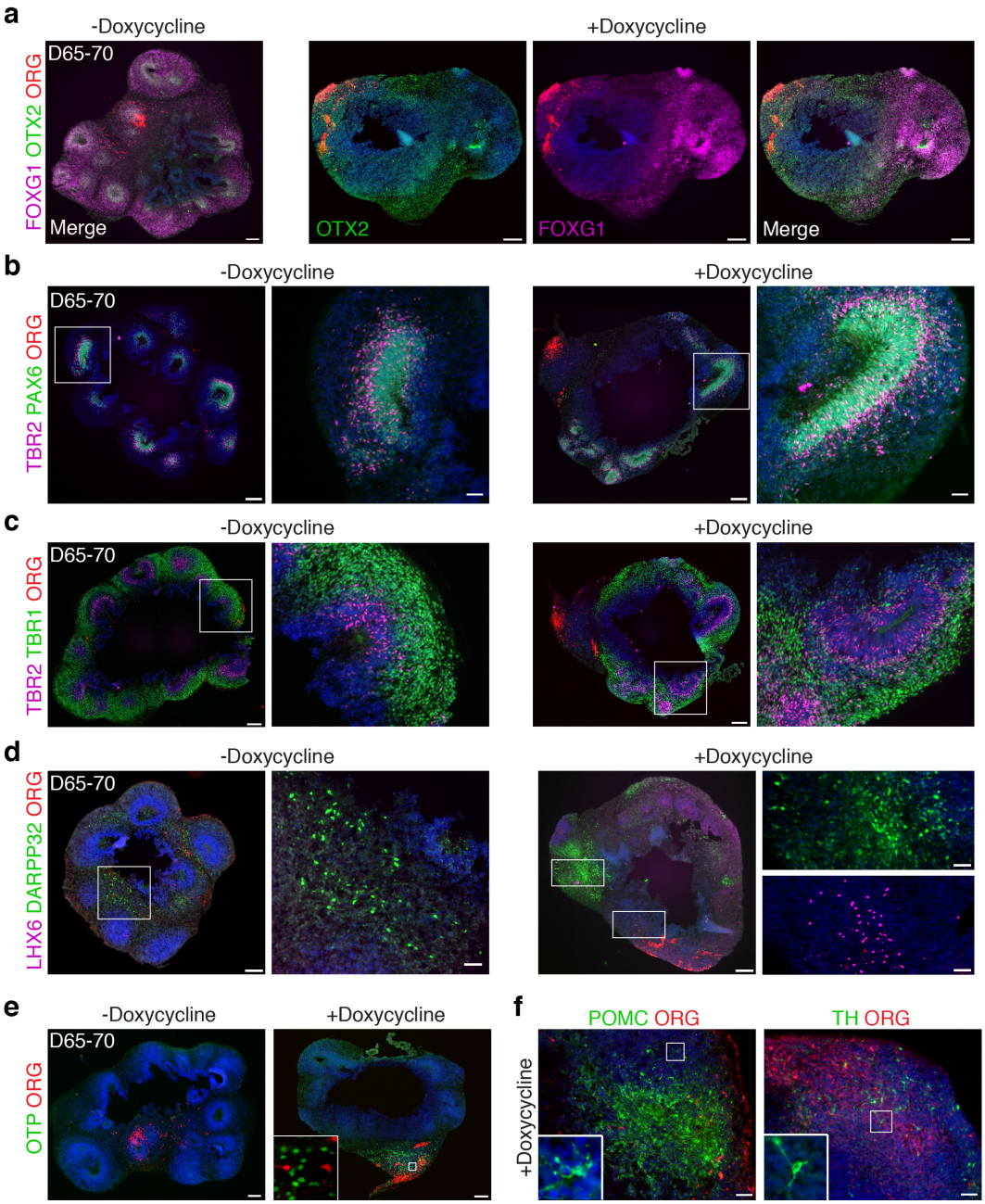


Figure 3.8. Characterization of maintenance of SHH-organoid topography. **(a)** OTX2 and FOXG1 regions remain largely distinct over time, and appear to retain their orientation with respect to the organizer. In some instance, the organizer tissue disperses throughout the organoid, making it difficult to determine orientation. **(b,c)** Cerebral cortex-like tissue of radially organized bands of PAX6, TBR2, and TBR1 emerge in distal regions. **(d)** Striatum-like tissue expressing DARPP32 emerges in distal regions. LHX6+ cells are found proximal to the organizer. **(e,f)** Hypothalamic-like cells, which express OTP, POMC, and TH, are found in the immediate vicinity of the organizer. Scale bars: 50μm (high magnification), 200μm (low magnification). All experiments and analysis by Gustav Cederquist and Lorenz Studer.

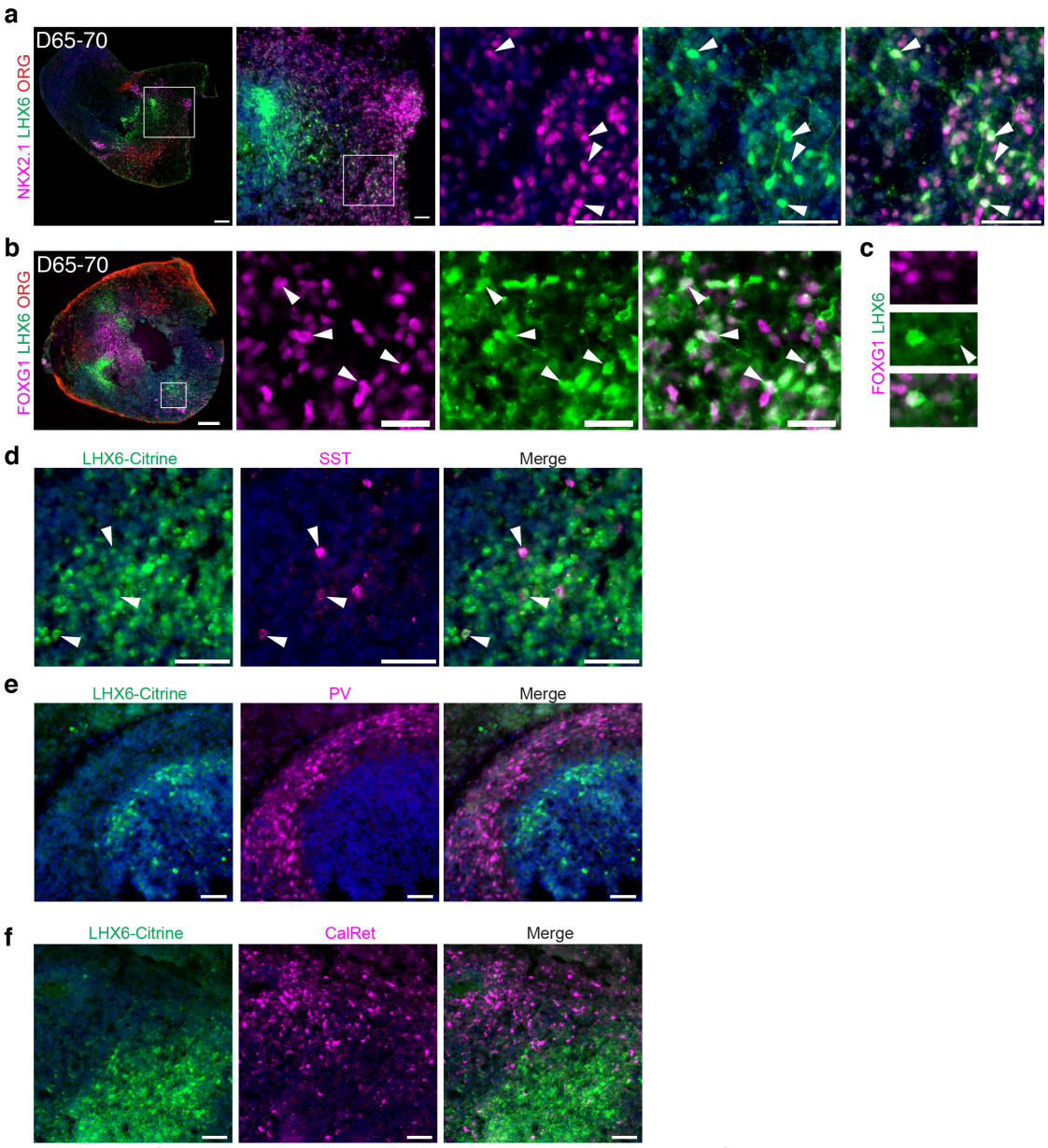


Figure 3.9. Characterization of interneuron diversity in SHH-organoids using LHX6-citrine line.

(a) LHX6+ cells emerge in regions proximal to the organizer, and co-express NKX2.1. **(b)** A subset of LHX6+ cells expresses FOXG1, consistent with striatal or cortical interneuron identity. **(c)** Some FOXG1+/LHX6+ cells have a leading process morphology, characteristic of migrating cortical interneurons. (d,e,f) Diverse interneuron populations expressing somatostatin (d), parvalbumin (e), and calretinin (f) are observed. Parvalbumin+ cells do not express LHX6, suggesting a non-MGE source of these cells. Scale bars: 50μm (high magnification), 100 μM (intermediate magnification), 200 μm (low magnification). All experiments and analysis by Gustav Cederquist and Lorenz Studer.

The identity of LHX6+ cells including the presence of cortical interneurons was further defined by differentiating SHH-organoids using an LHX6-citrine hPSC line. As expected, LHX6+ putative hypothalamic precursors emerge near the organizer, with some cells co-expressing NKX2.1 (**Figure 3.9 a**). A subset of LHX6+ cells co-express FOXG1+ (**Figure 3.9 b**), while a subset are negative for FOXG1, suggesting the presence of both MGE and hypothalamic LHX6 lineages. Some FOXG1+/LHX6+ cells exhibit a leading-process morphology characteristic of migrating interneurons (**Figure 3.9 c**). Cells expressing somatostatin, parvalbumin, or calretinin were observed in SHH-organoids (**Figure 3.9 d-f**), suggesting the emergence of diverse interneuron populations. However, parvalbumin+ cells did not co-express LHX6, suggesting those cells may not represent cortical interneurons.

We next sought insight into the mechanisms by which a SHH signaling gradient establishes forebrain topography. Using the suppression of PAX6 as a read-out of SHH signaling activity, we compared the radial extent of PAX6 suppression in 2D versus 3D cultures. High-dox conditions were used in assessing SHH activity. In 3D cultures, we observe the suppression of PAX6 expression at an average distance of 335 ± 54 μm from the organizer tissue (**Figure 3.10 a**). Interestingly, this range is similar to the distances reported for SHH signaling *in vivo* based on studies across multiple tissues of the developing mouse or chick embryos (Briscoe et al., 2001; Fan and Tessier-Lavigne, 1994; Honig, 1981). However, in 2D cultures in which a similar 1:10 ratio of iSHH to wildtype cells is maintained, we saw suppression of PAX6 at much shorter average distances (75 ± 31 μm) from the SHH source tissue (**Figure 3.10 a**).

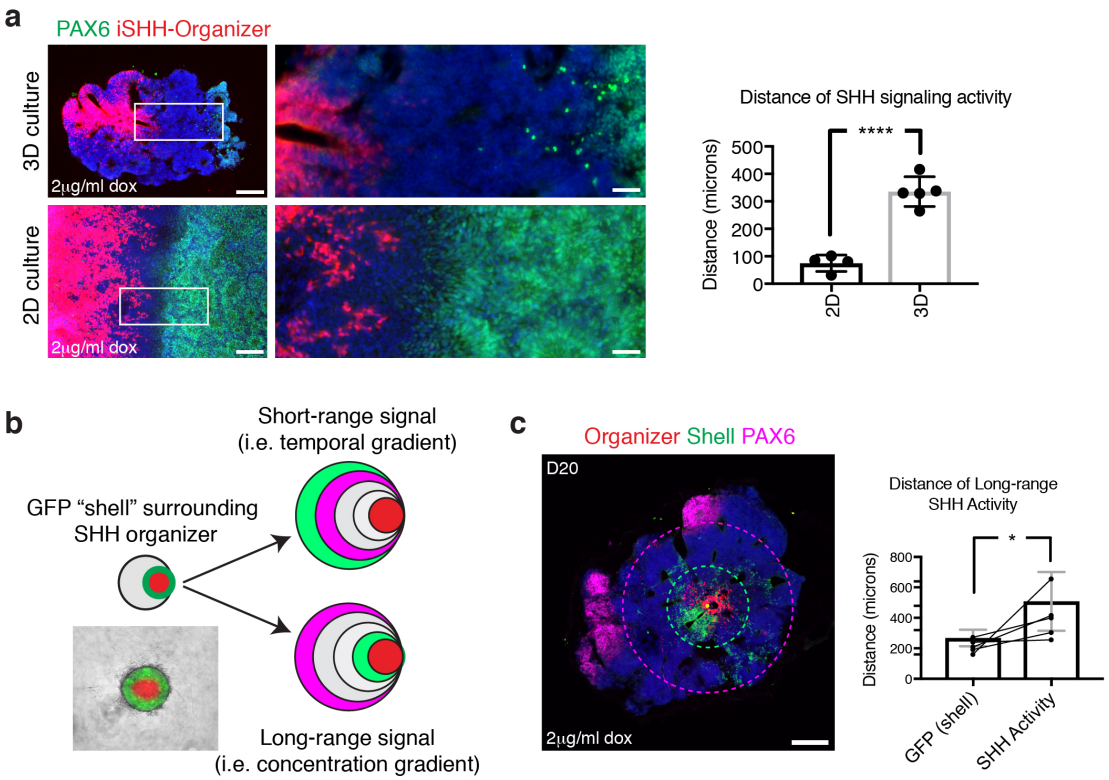


Figure 3.10. Characterization of long-range SHH signal in 3D culture.
(a) Comparison of distance of SHH-signaling activity after 12 days of differentiation in 3D and 2D culture using 2μg/ml doxycycline. In 3D culture, SHH acts over 335 ± 54 μm from the iSHH-organizer, assessed by suppression of PAX6 expression. In 2D culture, SHH acts 75 ± 31 μm from the iSHH-organizer cells. Insets show high-magnification images. Quantification shows mean ± S.D. Individual replicates are plotted on graphs. Student two-tailed t-test. **** $P < 0.0001$. N=5 organoids, 2 batches (3D), N=4 differentiations (2D). **(b)** Depiction of experiment to distinguish between two theoretical mechanisms for long-range SHH activity in 3D culture. A shell of GFP-expressing hPSCs is embedded around the iSHH organizer (see methods). A temporal model is supported if the all of SHH activity remains encapsulated within the GFP shell. A concentration-dependent model is supported if SHH activity extends past the GFP shell. **(c)** After 20 days of differentiation in 2μg/ml doxycycline, SHH activity is observed to extend past the GFP shell. The green dotted circle indicates the boundary of the GFP shell and the purple dotted circle indicates the boundary of the SHH activity, assessed by suppression of PAX6 activity. Quantifications show that the GFP shell extends 268 ± 24 μm from the iSHH-organizer, while SHH activity extends 508 ± 86 μm from the iSHH-organizer. Quantification shows mean ± S.D. Student two-tailed t-test. * $P = 0.028$. Individual replicates are plotted on graphs. N=5 organoids, 2 batches. Scale bars: 200 μm (low magnification), 50 μm (high magnification). All experiments and analysis by Gustav Cederquist and Lorenz Studer.

The observation that SHH has a limited range of activity in 2D culture raises the question how SHH establishes long-range gradients in 3D. We generated 2 theoretical models based on known mechanisms of SHH signaling *in vivo* (**Figure 3.10 b**). The first model is a concentration-dependent model. In this model, SHH signaling activity is determined by the local concentration of SHH morphogen, and is dependent on mechanisms that can transport SHH protein over long distances, such as diffusion, facilitated transport, or cell-to-cell relay (Briscoe et al., 2001; Chen et al., 2004; Fan and Tessier-Lavigne, 1994; Lewis et al., 2001; Sanders et al., 2013; Zeng et al., 2001). The second model is a temporal model. In this model, SHH signaling is restricted to tissue immediately adjacent to the source, which is relieved of SHH signaling as tissue grows away from the source (Harfe et al., 2004). Tissue that grows away from a SHH source will record a brief duration of signaling activity, while tissue that remains near the SHH source will record a long duration of signaling activity. This temporal model does not rely on transport of SHH over long distances.

To experimentally test these models, the iSHH organizer was embedded within a circumscribed ring of GFP-expressing hPSCs. This was achieved by sequential plating of 1,000 iSHH cells, followed by 1,000 GFP cells 8 hours later, followed by 10,000 H9 cells the next day. We reasoned that if the activity of SHH extends past the GFP boundary, then this would support a long-range concentration dependent mechanism. On the other hand, if the GFP boundary expands coincident with the extent of SHH activity, this would support a temporal model, in which tissue growth rather than protein transport establishes long-range SHH activity (**Figure 3.10 b**). After growing SHH-GFP- organoids for 20 days in doxycycline (2mg/ml), we observe that the GFP

expressing cells stay within 285 ± 44 μm to the SHH organizer cells. The activity of SHH, determined as suppression of PAX6 expression, extends to 430 ± 96 μm, thus extending past the green boundary (**Figure 3.10 c**). These data provide evidence for a concentration-dependent signaling mechanism in which SHH protein is transported long distances away from producing cells. Still, because the long-range SHH signal degrades in 2D culture, but a short-range signal persists, our data do not exclude the possibility of a concurrent short-range signaling mechanism, in which SHH protein is retained in the immediate vicinity of producing cells. In fact, during limb patterning, SHH is known to utilize both short- and long-range signaling mechanisms (McGlinn and Tabin, 2006) and such an interplay of short- and long-range signaling may also be critical during forebrain development (Blaess et al., 2014; Ericson et al., 1995; Zhang and Alvarez-Bolado, 2016).

The reproducible topography of SHH-organoids might make them suitable to discern phenotypes related to the development and maldevelopment of forebrain morphogenesis. As proof-of-principle we sought to use SHH-organoids to study the potential adverse effects of statins during fetal forebrain development. Statins are among the most widely prescribed drugs and are known to lower cholesterol levels. Cholesterol is an integral component of the SHH pathway as both a post-translational ligand modification (Creanga et al., 2012; Lewis et al., 2001; Tian et al., 2005; Tukachinsky et al., 2012) and as an agonist of the smoothened receptor (Byrne et al., 2016; Huang et al., 2016). Interestingly, a retrospective case series examining birth defects associated with statin use during first trimester pregnancy found a high percentage of infants born with a range of adverse birth outcomes ranging from growth delay to limb, heart and CNS midline

defects, which may be associated with reduced SHH signaling (Edison and Muenke, 2004a). Subsequent follow up studies did not observe the same effects in alternative cohorts (Kazmin et al., 2007; Taguchi et al., 2008; Winterfeld et al., 2013). F1 progeny of pregnant mice exposed to high-doses of statins typically exhibited fetal loss, neurologic, or skeletal abnormalities (Edison and Muenke, 2004b). Thus, it remains an important, unresolved question whether statins can interrupt SHH-related signaling events during human embryonic development.

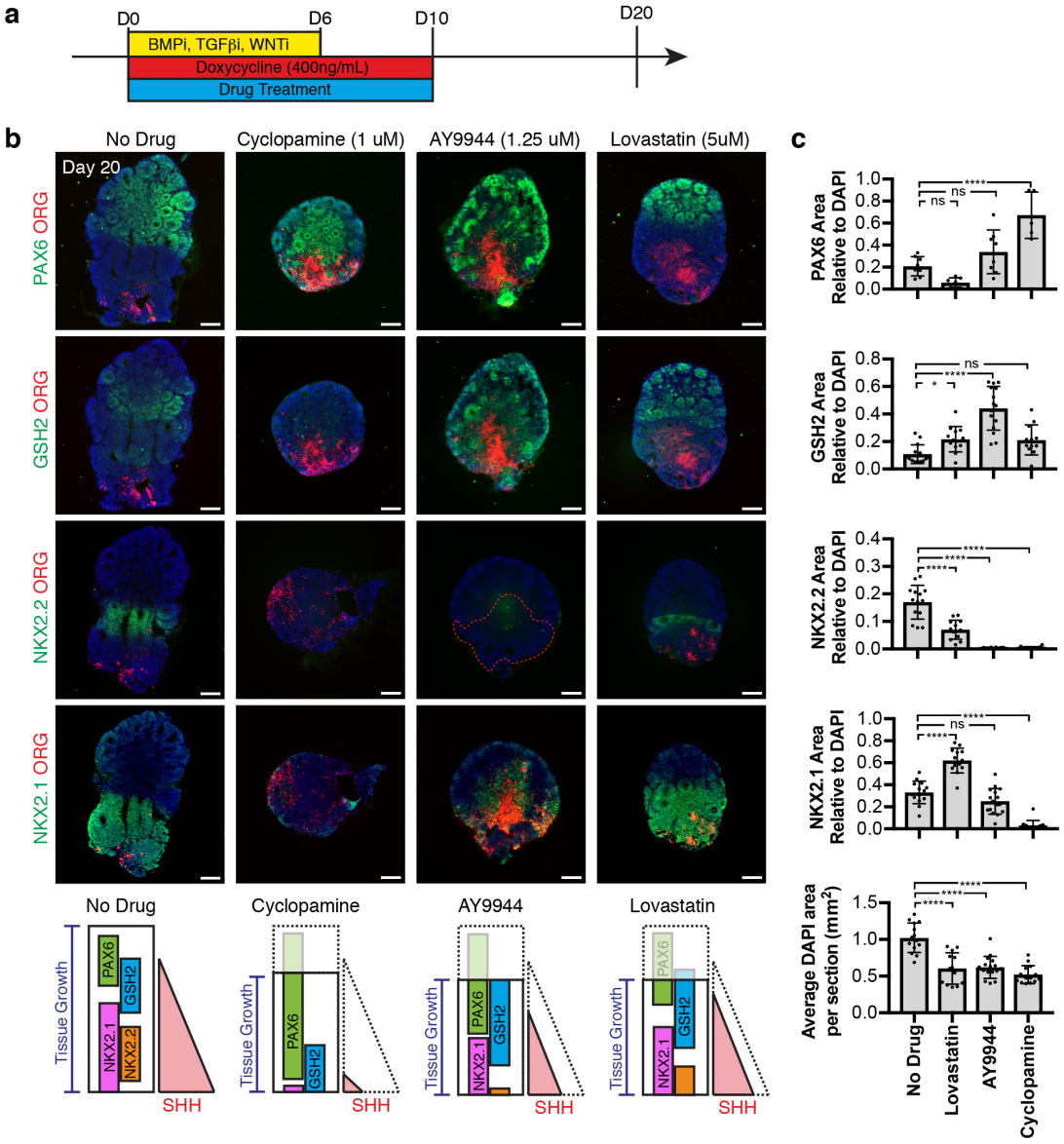
To test whether statins perturb fetal brain growth or patterning, SHH-organoids were treated with doxycycline (400 ng/ml) plus lovastatin (5 μM (Blassberg et al., 2016)) from day 0-10 of differentiation (**Figure 3.11 a**), and analyzed at day 20. The effects of AY9944 (1.25 μM (Blassberg et al., 2016)), a specific inhibitor of the 7-dehydrocholesterol reductase enzyme (DHCR7) that facilitates the terminal step of cholesterol synthesis, or cyclopamine (1 μM (Gaspard et al., 2008)) a Smoothened antagonist that blocks all SHH activity were also tested. In the no drug condition, SHH-organoids established distinct positional domains marked by the expression of PAX6, GSH2, NKX2.2, and NKX2.1 at discrete distances from the SHH-organizer (**Figure 3.11 b**). As expected, SHH-organoids grown in the presence of cyclopamine were nearly uniform for PAX6 expression, with some GSH2 expression, consistent with a near complete inhibition of SHH signaling activity (**Figure 3.11 b,c**). AY9944 treatment reduced the efficacy and range of SHH signaling activity, shifting the organoid to a more dorsal identity. NKX2.2 induction was largely blocked, and NKX2.1 expression was only retained near the organizer. Cells just adjacent to the organizer expressed GSH2 and PAX6 (**Figure 3.11 b,c**). Lovastatin treatment caused similar, but milder effects, than AY9944. There was a partial

block in NKX2.2 induction, with concomitant increase in GSH2 expression (**Figure 3.11 b,c**), consistent with a dorsal shift and reduction in the efficacy and range of SHH signaling activity.

Interestingly, the organizer retained diencephalic expression of OTX2 and NKX2.1 when treated with lovastatin or AY9944 (**Figure 3.12**), suggesting that cholesterol inhibition has minimal effects on the local levels of SHH signaling activity. This observation raises the possibility that distinct signaling mechanisms are used to effect short-range versus long-range SHH activity, as occurs during limb patterning (McGlinn and Tabin, 2006).

In addition to the effects on patterning, all drugs reduced organoid size (**Figure 3.11 b,c**), resulting in a complex phenotype in which both growth and patterning are affected. To dissociate growth and patterning phenotypes, we analyzed organoids at day 6, during the early stages of SHH patterning, and prior to the major phase of tissue growth (which occurs after matrigel embedding). We used NKX2.2, a direct target of GLI, as a read-out of SHH signaling activity. By measuring the distance at which NKX2.2 is expressed from the SHH producing cells, we could infer the distance of SHH signaling activity. Lovastatin reduced the distance at which NKX2.2 is induced from the organizer in a dose-dependent manner (**Figure 3.13**). These data support a model in which inhibition of cholesterol synthesis by Lovastatin or AY9944 result in a complex developmental phenotype characterized by impaired growth and a reduction in the range of SHH signaling activity.

Figure 3.11. Cholesterol synthesis inhibition perturbs tissue growth and SHH signaling.
(a) Experimental design to test effect of statins on SHH-organoid development. **(b)** Four distinct positional domains are identified in day 20 SHH-organoids (400 ng/ml doxycycline). SHH-organoids treated with cyclopamine (1μm) largely express PAX6, with some GSH2 immunoreactivity, suggesting a near complete inhibition of SHH activity. SHH-organoids treated by AY9944 (1.25μm) fail to induce NKX2.2 and only retain NKX2.1 expression in the immediate vicinity of the organizer, suggesting a strong reduction the range of SHH-signaling activity. SHH organoids treated with Lovastatin (5μm) exhibit a moderate reduction in NKX2.2 induction, with a concomitant increase in the relative area of the GSH2 expression domain. Summary of drug phenotypes and potential impact on SHH signaling activity is depicted at bottom of panel. **(c)** Quantification of area of each positional domain, relative to total area. Dots represent individual organoids. **(d)** Organizer tissue retains expression of OTX2 and NKX2.1 in presence of Lovastatin and AY9944, but not cyclopamine. Error bars are S.D. No drug N=8-14; Cyclopamine N=5-14; AY9944 N=8-16; Lovastatin N=7-13. * $P < 0.05$; *** $P < 0.001$, **** $P < 0.0001$. Scale bars: 200 μm (b), 100 μm (d). All experiments and analysis by Gustav Cederquist and Lorenz Studer.


a
D0
D6
D10
D20
BMPi, TGFβi, WNTi
Doxycycline (400ng/mL)
Drug Treatment
b
No Drug
Cyclopamine (1 uM)
AY9944 (1.25 uM)
Lovastatin (5uM)
Day 20
PAX6 ORG
GSH2 ORG
NKX2.2 ORG
NKX2.1 ORG
No Drug
Cyclopamine
AY9944
Lovastatin
Tissue Growth
PAX6
GSH2
NKX2.1
NKX2.2
SHH
c
PAX6 Area Relative to DAPI
GSH2 Area Relative to DAPI
NKX2.2 Area Relative to DAPI
NKX2.1 Area Relative to DAPI
Average DAPI area per section (mm²)
ns
No Drug
Lovastatin
AY9944
Cyclopamine

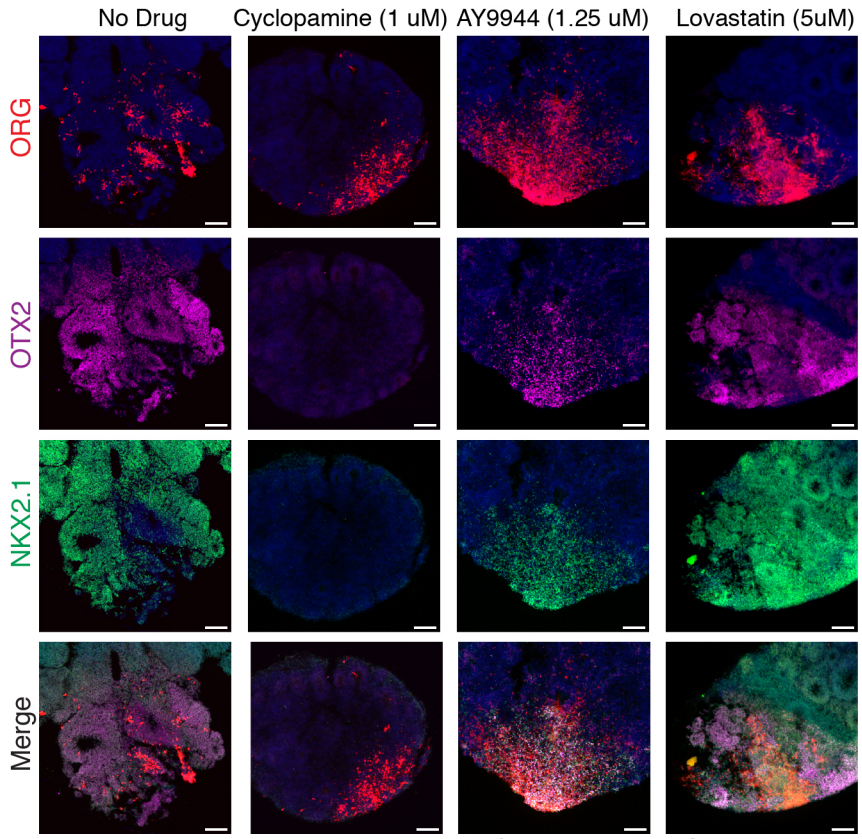


Figure 3.12. Minimal perturbation of short-range SHH activity by cholesterol synthesis inhibition.

Organizer tissue retains expression of OTX2 and NKX2.1 in presence of Lovastatin and AY9944, but not cyclopamine. No drug N=14; Cyclopamine N=15; AY9944 N=15; Lovastatin N=14. Scale bars: 100 μm. All experiments and analysis by Gustav Cederquist and Lorenz Studer.

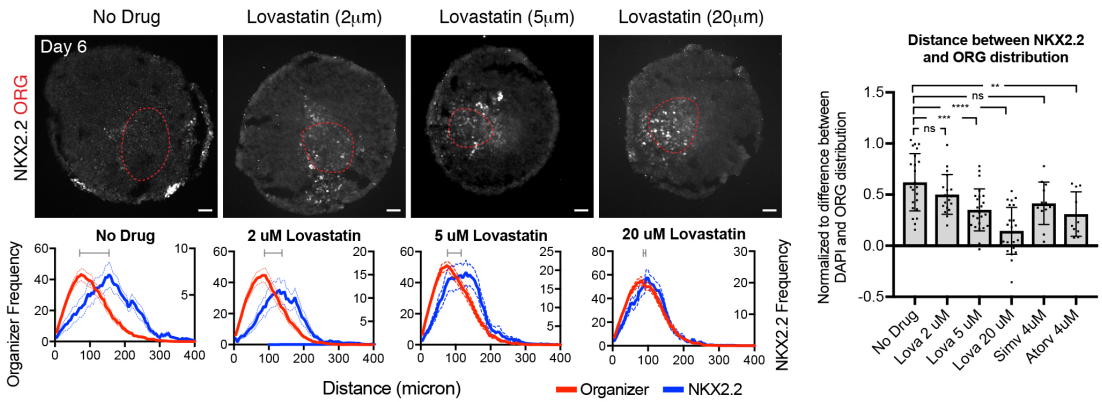


Figure 3.13. Dose-dependent reduction in range of SHH signal by lovastatin. Lovastatin produces a dose-dependent reduction in the distance at which NKX2.2 is induced from the organizer (red dotted boundary). Frequency histograms plot the relative distribution of organizer and NKX2.2 positive cells as a function of distance from the center of the organizer. Dotted lines represent S.E.M. Bar graphs quantify the average distance between the center of NKX2.2 and organizer distributions, illustrated by grey bars in frequency histograms. Dots represent individual organoids. Error bars are S.D. No drug, N=23; 2μm N=18; 5μm N=24; 20 μm N=23; Simvastatin N=12; Atorvastatin N=10. Scale bars: 50 μm. All experiments and analysis by Gustav Cederquist and Lorenz Studer.

Discussion

This study provides the first evidence that an asymmetric morphogenetic cue provides positional information from which *in vivo*-like topography can emerge in brain organoids. Furthermore, we present evidence that SHH uses a long-range signaling mechanism during forebrain organoid development, recapitulating its *in vivo* behavior (McGlinn and Tabin, 2006). The emergence of long-range activity in 3D that is not present in 2D cultures highlights the importance of tissue geometry on the spatial dynamics of signaling activity.

The reproducibility and robustness of our strategy allowed us to use SHH organoids as a developmental toxicology platform. Statins have been reported to be associated with fetal maldevelopment in a small number of cases (Edison and Muenke, 2004a). Our data suggest that statins can impair forebrain development by degrading long-range SHH signaling and reducing

tissue growth. These effects were observed in our organoid system in the low micromolar range, which is significantly higher than the typical plasma concentration of statins (15-20 nM) (Bellosta et al., 2004). This discrepancy could suggest that the normal plasma concentration of statins would have minimal effects of fetal development. However, statins may affect a subset of patients who harbor a genetic predisposition to holoprosencephaly, or who present with additional co-morbidities that allow lipophilic statins to accumulate in developing fetal tissue (Edison and Muenke, 2004b). Mechanistically the precise role of cholesterol in SHH signaling has been controversial (Briscoe and Therond, 2013; Guerrero and Chiang, 2007). Our data are consistent with evidence that cholesterol supports long-range SHH signaling (Lewis et al., 2001), which may occur through the formation of soluble multimeric complexes (Zeng et al., 2001) and by facilitating interactions with SHH-transport proteins essential for long-range activity (Creanga et al., 2012; Tukachinsky et al., 2012).

A key feature of the patterning strategy presented here is the establishment of well-ordered positional domains during early organoid development. In future studies it will be important to determine to what extent the topography of these domains can be retained during long-term *in vitro* culture for the study of region-specific neuronal diversity as well as connectivity and function. It is possible that additional tissue engineering approaches will be required to facilitate the maintenance of topography over time and to prevent increasing cell mixing or inappropriate migration, processes which could lead to a blurring of the boundaries between positional domains over time. For example, manipulating soluble ECM components has

been developed as a strategy to improve the maintenance of radially organized structures within organoids over time (Lancaster et al., 2017).

For disease modeling, current organoid technologies have been successful at modeling cortical progenitor dynamics that ultimately result in changes in brain size (Di Lullo and Kriegstein, 2017). Pre-patterning and fusion of pallial- and subpallial-specific spheroids have been used to study specific phenotypes related to cortical interneuron migration and connectivity (Birey et al., 2017). In our current study, the establishment of forebrain topography in organoids opens the possibility of studying a wide range of phenotypes in a single organoid system. In particular, it would be interesting to study complex neurodevelopmental diseases that have been hypothesized to alter or derive from regional specification during forebrain patterning, including autism (Mariani et al., 2015), epilepsy (Cobos et al., 2005), bi-polar disorder (Chen et al., 2014), or suprasellar pediatric gliomas (Fontebasso et al., 2014). Finally, it will be exciting to test whether our strategy could be used as a general approach to establish topographies across all regions of the CNS beyond the forebrain. Establishing properly organized forebrain organoids, with the correct regional topography is an important step on the road of realizing the full potential of human “mini-brain” technology.

CHAPTER FOUR: A HUMAN PLURIPOTENT STEM CELL BASED MULTIPLEX ANALYSIS PLATFORM IDENTIIES MOLECULAR AND FUNCTIONAL SUBTYPES OF AUTISM

Abstract

Autism is a clinically heterogeneous neurodevelopmental disorder characterized by impaired social interactions, restricted interests and repetitive behaviors. Despite significant advances in uncovering its immense genetic diversity (De Rubeis et al., 2014; Durand et al., 2007; Glessner et al., 2009; Iossifov et al., 2015; Iossifov et al., 2014; Jamain et al., 2003; Krumm et al., 2015; Ronemus et al., 2014; Sanders et al., 2015; Sebat et al., 2007), a systematic understanding of how autism mutations perturb brain development and ultimately affect clinical outcome has remained elusive. This challenge reflects a broader limitation in studying human disorders, as most

experimental models fail to capture the genetic heterogeneity and cell type specific vulnerability characteristic of complex disease (McClellan and King, 2010). Human pluripotent stem cells (hPSCs) offer a promising alternative for modeling complex disorders such as autism. However, the laborious nature of studying individual mutations in hPSCs, concerns about line-to-line variability and marked cellular heterogeneity remain major stumbling blocks.

Here, we present a novel platform to study 30 isogenic hPSC lines in parallel, including 27 lines representing high-confidence *de novo* autism mutations. All hPSC lines are pooled in a single dish and differentiated into disease-relevant cell types of prefrontal cortex (PFC) identity. Cell line specific genetic markers are used to test early-developmental hypotheses of autism (Courchesne et al., 2007; De Ferrari and Moon, 2006; Ernst, 2016; Kalkman, 2012; Krumm et al., 2014; Packer, 2016a, b) for each individual mutation across all hPSC lines. We demonstrate that 59% of the mutations (16/27) perturb prefrontal cortex (PFC) neurogenesis through dysregulation of SOX2+ stem cell behavior, a phenotype further correlated to abnormal WNT/bcatenin responses. Mutations fall into two distinct classes. Class 1 mutations (8/27) inhibit, while Class 2 mutations (8/27) enhance PFC neurogenesis. Remarkably, analysis of clinical patient data reveals that individuals with Class 1 versus Class 2 mutations exhibit distinctive autism profiles based on their trajectory of language acquisition. These results point to a surprising level of structure across autism mutations and reveal brain endophenotypes to define novel, clinically relevant patient subpopulations. Finally, our multiplex hPSC technology should be suitable to disentangle genetic heterogeneity across other complex human disorders and facilitate evolving efforts in precision medicine[19].

Introduction

Neuroimaging and neuropathology studies show frequent alterations in PFC growth and neurogenesis in autism patients (Courchesne et al., 2011; Courchesne et al., 2007; Hazlett et al., 2017; Stoner et al., 2014). In addition, bioinformatic approaches indicate that autism-associated genes interact with transcriptional networks of the frontal cortex and cerebellum (Willsey et al., 2013), and segregate into two temporal categories with peak expression at post-conception week (PCW) 8-20 or shortly after birth (Parikshak et al., 2013). The early category of genes is associated with transcription and chromatin remodeling while the later category of genes is associated with synapse development and function.

The question of whether a given autism mutation directly perturbs PFC growth and neurogenesis can be studied using traditional animal models. For multi-gene disorders however, the functional characterization of dozens of genes in animal or cell-based models remains challenging and typically restricted to resource intensive settings such as large-scale consortia (Sweet, 2017). hPSCs have the potential to solve this problem in three ways. First, hPSCs provide access to disease-relevant human tissue through high-quality differentiation protocols (Sterneckert et al., 2014). Second, CRISPR/Cas9 allows rapid engineering of disease lines (Hsu et al., 2014). Third, cell lines can be pooled into a single dish to increase throughput and reduce assay variability as pioneered in cancer cell lines (Birsoy et al., 2014; Yu et al., 2016).

Here, we designed an hPSC-based multiplex platform in which multiple disease lines are pooled and differentiated into disease-relevant cell types (**Figure 4.1**). Phenotypes for each line are then discerned by measuring changes in the relative allele frequency over time (e.g. growth phenotype (Birsoy et al., 2014)) or across physically separated phenotypic cell fractions (e.g. cell-state or drug-response (Yu et al., 2016)).

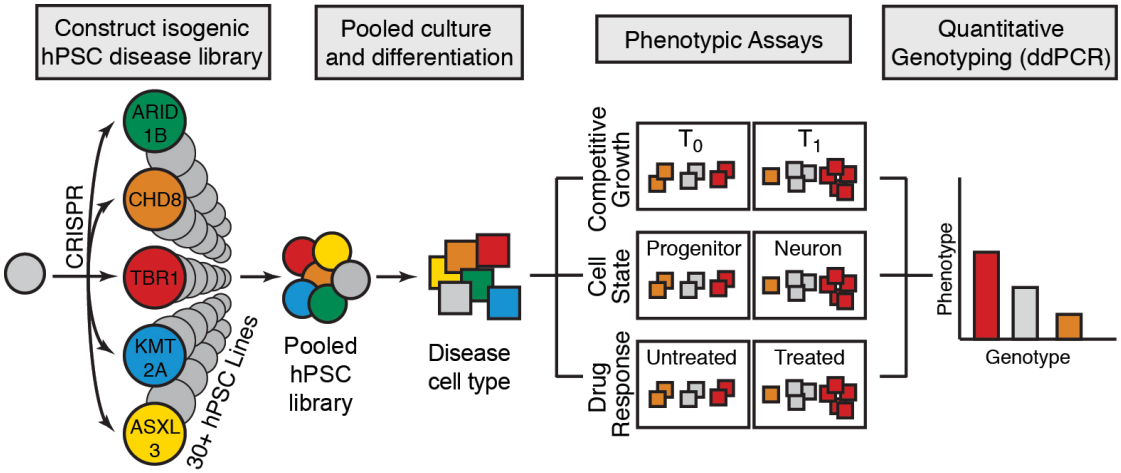


Figure 4.1. Design of hPSC-based multiplex analysis platform.
Individual disease-associated hPSC lines are generated using CRISPR/Cas9, pooled, and differentiated into a disease-relevant tissue. The pooled differentiation can be assayed for growth, cell-state, or drug-response phenotypes by determining relative allele frequencies for each line in comparison to an internal standard (negative control). For example, growth phenotypes are determined by measuring changes in allele frequency over time (T_1 vs T_0). Cell-state phenotypes are determined by measuring differences in allele frequency between physically separated populations (e.g. neurons versus progenitors separated via fluorescence activated cell sorting (FACS)). Drug response phenotypes are determined by measuring differences in allele frequency between treated and untreated pools. Allele frequencies are measured using ddPCR. Experimental Design by Gustav Cederquist, Mark Tomishima, Lorenz Studer.

Allele frequencies are measured using droplet digital PCR (ddPCR) due to its high sensitivity and reproducibility (Hindson et al., 2013) (**Figure 4.2a,** limit of detection > 1:7000 genomes). The multiplex platform was next validated for a physical

separation-based assay measuring differential CTNNB1 protein expression, which accurately identified a CTNNB1-deficient line from a pooled hPSC culture of 8 lines (**Figure 4.2b**).

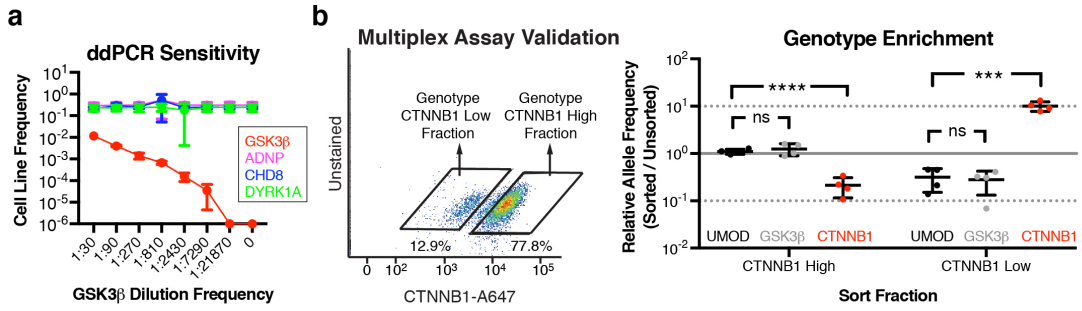


Figure 4.2. Validation of hPSC-based multiplex analysis platform.
(a) ddPCR is a sensitive and accurate method to measure allele frequency, as determined by a serial dilution assay. ddPCR could detect the GSK3β line within a 5-line mixture until it reached a frequency between 1:7290 and 1:21870, using a read depth of ~15,000. The number of ddPCR reads is directly proportional to the number of GSK3β cells. n = 3 dilution series, mean ± s.d. **(b)** Validation of multiplex assay. A pool of 8 hPSC lines, including CTNNB1, UMOD, and GSK3β mutant lines, was separated into CTNNB1-low and CTNNB1-high expressing fractions using intracellular FACS with a CTNNB1 antibody. Each fraction was genotyped with ddPCR to calculate relative allele frequencies (allele frequency in sorted fraction / allele frequency in unsorted fraction). The CTNNB1 mutant line was depleted in the CTNNB1-high fraction (corrected p = 0.0007), and enriched in the CTNNB1-low fraction (corrected p = 0.0001), relative to the negative control line UMOD. P values are corrected using an ANOVA followed by Dunnett's test. n = 4 independent trials, mean ± s.d. Experiments performed and analyzed by Gustav Cederquist and Lorenz Studer.

The first key feature of the multiplex platform is its ability to capture the genetic heterogeneity of complex disease. Toward this end, we used CRISPR/Cas9 to construct an isogenic disease library of high-confidence autism mutations from a 46XY founder hPSC line (**Figure 4.3a, Table 4.1**). Frameshift indels were preferentially introduced into the specific exons that are mutated in patients (De Rubeis et al., 2014; Iossifov et al., 2014). Genes were selected based on a high confidence score in the SFARI Gene Database and further filtered for genes with early expression during cortical development *in vivo* (BrainSpan) and *in vitro*

(Cortecon). The resulting library was comprised of 27 autism lines and was enriched for genes related to transcription/chromatin-remodeling (Parikshak et al., 2013). As a negative control line, we targeted the intron of a gene that is not expressed in neural tissue (UMOD). Two WNT/βcatenin-related positive controls (CTNNB1 and GSK3β) were also included.

We made three independent 30-line mixtures by pooling all lines at the pluripotent stage (MIX30A, B, C) (**Figure 4.3b**). MIX30 pools retained expression of pluripotency genes (**Figure 4.3c**, $OCT4^{+}/SSEA^{+}$ 89.7% ± 1.08%) and all individual lines were well represented at the pluripotent stage (**Figure 4.3d**). Spontaneous TP53 mutations (Merkle et al., 2017) were not detected at levels above background in any of the 12 lines examined (**Figure 4.4, Figure 4.5**).

Allele frequencies were stable at the pluripotent stage, with less than 3-fold average change through seven passages (**Figure 4.6 a,b**), and allele frequencies were not drastically affected by freeze-thaw cycles (**Figure 4.6c**), allowing for expansion, quality-control, and long-term storage of each of the three pools.

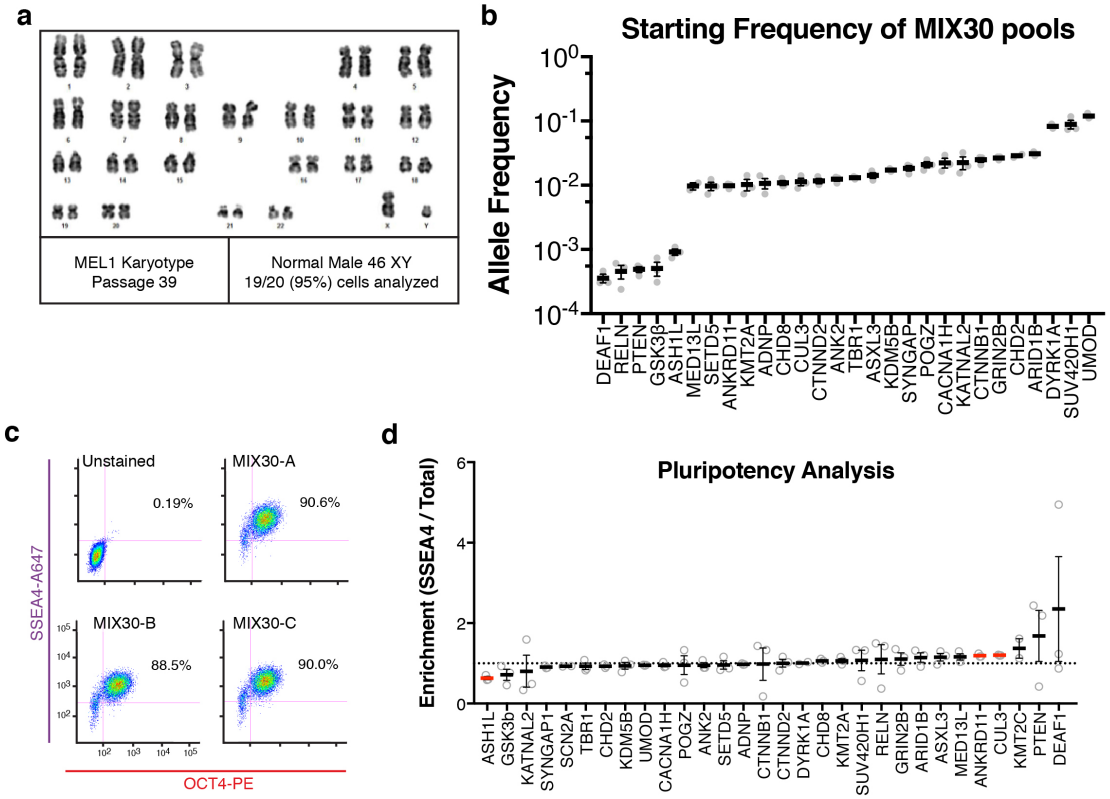


Figure 4.3. Additional characterization of MIX30 pooled autism library.
(a) Karyotype analysis of 46XY MEL1 founder line demonstrates normal karyotype in 19/20 (95%) of cells. **(b)** Average starting frequency of each line in MIX30 pools. Some lines that showed increased fitness in preliminary studies were mixed at low frequencies while the negative control, UMOD, and some lines that showed decreased fitness were mixed at high frequencies. n = 3 MIX30 pools. **(c)** Flow cytometry analysis for expression of pluripotency markers OCT4 and SSEA4 in MIX30 pools. **(d)** Average cell line enrichment in MIX30 SSEA4 sorted fractions relative to unsorted MIX30 fractions. Red bars indicate cell lines with significant increases or decreases in enrichment score compared to UMOD (FDR < 0.05). n = 3 MIX30 pools. Dots represent independent MIX30 pools. Error bars are s.e.m. Experiments performed and analyzed by Gustav Cederquist and Lorenz Studer.

Table 1. Isogenic autism library

Gene	Function	Zygosity	Indel	Indel	Location (GRCh38/hg38)	Cortecon									Brainspan
						0	7	12	19	26	33	49	63	77	PCW8
ADNP	Transcription/Chromatin	Het	1 bp deletion	GTCCCCTTAG[-]CTCGAATGCC	Chr20/-/50894043	7690	12854	10048	7557	6846	7799	8044	6277	6656	4.9948
ANK2	Cytoskeleton	Homo	1 bp insertion	TCACCTGCCG[G]ACTGGTCAAG	Chr4/+/113330366	2187	5750	2992	4457	2917	3059	3757	5763	9000	3.4381
ANKRD11	Transcription/Chromatin	Het	1 bp insertion	GTTCGAAAAA[A]GGCCTCAAGC	Chr16/-/89285534	9570	5530	4485	4145	4608	4716	5297	4408	5253	2.60271429
ARID1B	Transcription/Chromatin	Homo	1 bp insertion	CTGATGAAACA[A]CGCAGGCGCC	Chr6/+/157167135	4584	5168	4608	4395	3808	3407	3601	3528	4572	2.50658571
ASH1L	Transcription/Chromatin	Homo	1 bp insertion	ATCTACTCTT[T]CTTGGTAGCC	Chr1/-/155482239	1484	2428	2345	1825	2070	2652	2973	3078	3960	2.56501429
ASXL3	Transcription/Chromatin	Het	14 bp deletion	AACCAGTCTC[--------------]GAG	Chr18/+/33742955	41	483	594	737	714	630	633	632	803	1.69975714
CACNA1H	Membrane Channel	Homo	1 bp insertion	AACGTGTGCC[C]GCTCGGGTGA	Chr16/+/1200495	944	3535	2453	3741	2322	2653	3696	3739	3683	2.72537143
CHD2	Transcription/Chromatin	Homo	23 bp deletion	ATATGGATGC[----------]	Chr15/+/93020177	3573	4436	5558	4747	5835	7231	6192	5243	5104	3.73221429
CHD8	Transcription/Chromatin	Het	1 bp insertion	GGCAACCTTC[C]GACAGGCTGC	Chr14/-/21403068	8081	7617	6548	4805	5220	5825	5900	4360	5064	3.55768571
CTNND2	Cell Adhesion	Het	7 bp deletion	GCTTAAACAC[-------]GGGGATGGCT	Chr5/-/11732123	322	3651	4823	4906	5962	5646	5830	3598	4232	4.68905714
CUL3	Ubiquitination	Homo	1 bp insertion	TTTCCAAGCA[A]CATGAAGACT	Chr2/-/224511348	6053	7025	7124	5949	4929	5460	5576	5153	5167	4.14752857
DEAF1	Transcription/Chromatin	Homo	5 dp deletion	AGCTGCCCGT[-----]GCCGGAACAT	Chr11/-/687911	1644	2546	2397	2786	3155	2805	2923	3174	2914	3.92462857
DYRK1A	Intracellular signaling	Het	1 bp insertion	GCCAGCCAAA[A]CATAAGTGAC	Chr21/+/37472814	3015	4542	3831	2592	2829	3427	3109	2985	2672	4.1017
GRIN2B	Membrane Channel	Het	8 bp deletion	tcgtcagAGC[--------]GCTGACGTA	Chr12/-/13611738	26	115	84	58	78	85	158	104	101	0.6305
KATNAL2	Cytoskeleton	Het	1 bp insertion	GAATTGGCAG[G]CCGTGGTGAG	Chr18/+/47063348	322	713	529	761	534	508	568	986	642	0.53881429
KDM5B	Transcription/Chromatin	Homo	1 bp deletion	TTGCAAAGTA[-]TGGGAGTTA	Chr1/-/202774629	9749	18472	16540	13981	16316	16850	16979	11740	10387	4.68171429
KMT2A	Transcription/Chromatin	Homo	1 bp deletion	GTCCTCGAAA[-]CCTCGTGGGA	Chr11/+/118471682	4644	5559	4168	3211	3858	5074	3909	3016	3414	2.64244286
KMT2C	Transcription/Chromatin	Het	4 bp deletion	CAAAAATTGG[----]GGTAAACCAG	Chr7/-/152247959	2007	2989	2868	1840	2508	2566	2418	2237	2719	1.80098571
MED13L	Transcription/Chromatin	Het	1 bp insertion	CAAAAGAAGC[C]GCACGGGCAG	Chr12/-/115987136	3964	8413	6497	5219	4984	5400	5157	4189	4989	3.08994286
POGZ	Transcription/Chromatin	Het	1 bp insertion	GGTCAACTCGA[A]GTGTGGCAGA	Chr1/-/151405301	3127	5466	5359	3692	3758	3990	3998	2797	3185	3.51881429
PTEN	Intracellular signaling	Homo	1 bp deletion	CAGGACCCACA[-]GACGGGAAGA	Chr10/+/87957916	1551	1527	1789	1797	1610	1982	1919	2370	2773	3.28431429
RELN	Extracellular signaling	Homo	1 bp insertion	ATTTGAGAGCCAGCCT[T]ACAGG	Chr7/-/103682147	9	583	1957	2478	1388	2510	1917	3348	8144	4.04584286
SCN2A	Membrane Channel	Homo	1 bp deletion	CTCACCTTGG[-]ATTGGCCATC	Chr2/+/165344653	90	9	8	18	28	32	65	67	136	1.39938571
SETD5	Transcription/Chromatin	Het	5 bp deletion	ATAATAGAGT[-----]GGGAAAGTCA	Chr3/+/9441689	10144	14485	12065	8089	8008	9608	9039	6221	7140	3.45837143
SUV420H1	Transcription/Chromatin	Homo	8 bp deletion	GTTAAGCAAA[--------]CAACGATAA	Chr11/-/68158023	2091	4393	4275	3290	4350	3215	3945	3567	3034	3.58494286
SYNGAP1	Intracellular signaling	Homo	1 bp insertion	GGACATCGCA[A]GACAGGCTTA	Chr6/+/33440808	1717	2243	2172	2280	2326	2062	2863	2940	3329	3.0595
TBR1	Transcription/Chromatin	Homo	1 bp insertion	GCAGCCCTAG[T]CCGCTACATG	Chr2/+/161416839	2	1	1	0	7	75	198	430	659	5.21248571
Internal Standards															
UMOD	Wildtype Control	Homo	1 bp insertion	TACCAGGGTT[T]ACAACAATGAA	Chr16/-/20352392	1		0	0	0		0	0	0.3	0
CTNNB1	WNT Low	Homo	23 bp deletion	AGAAAAGCGG[----------]	Chr3/+/41224597	33046	26345	32980	30942	34353	38004	38416	27681	26968	6.88938571
GSK3b	WNT High	Homo	1 bp insertion	CACCTCCTTT[T]GCGGAGAGCT	Chr3/-/120093347	3724	5556	6239	4507	4966	5290	5109	3927	3843	4.26818571

Information regarding isogenic hPSC lines in MIX30 library. Gene names, annotated gene function, zygosity and sequence of CRISPR indel, and gene expression data from hPSC-derived cortical cultures (Cortecon database) and post-conception week 8 human fetal tissue.

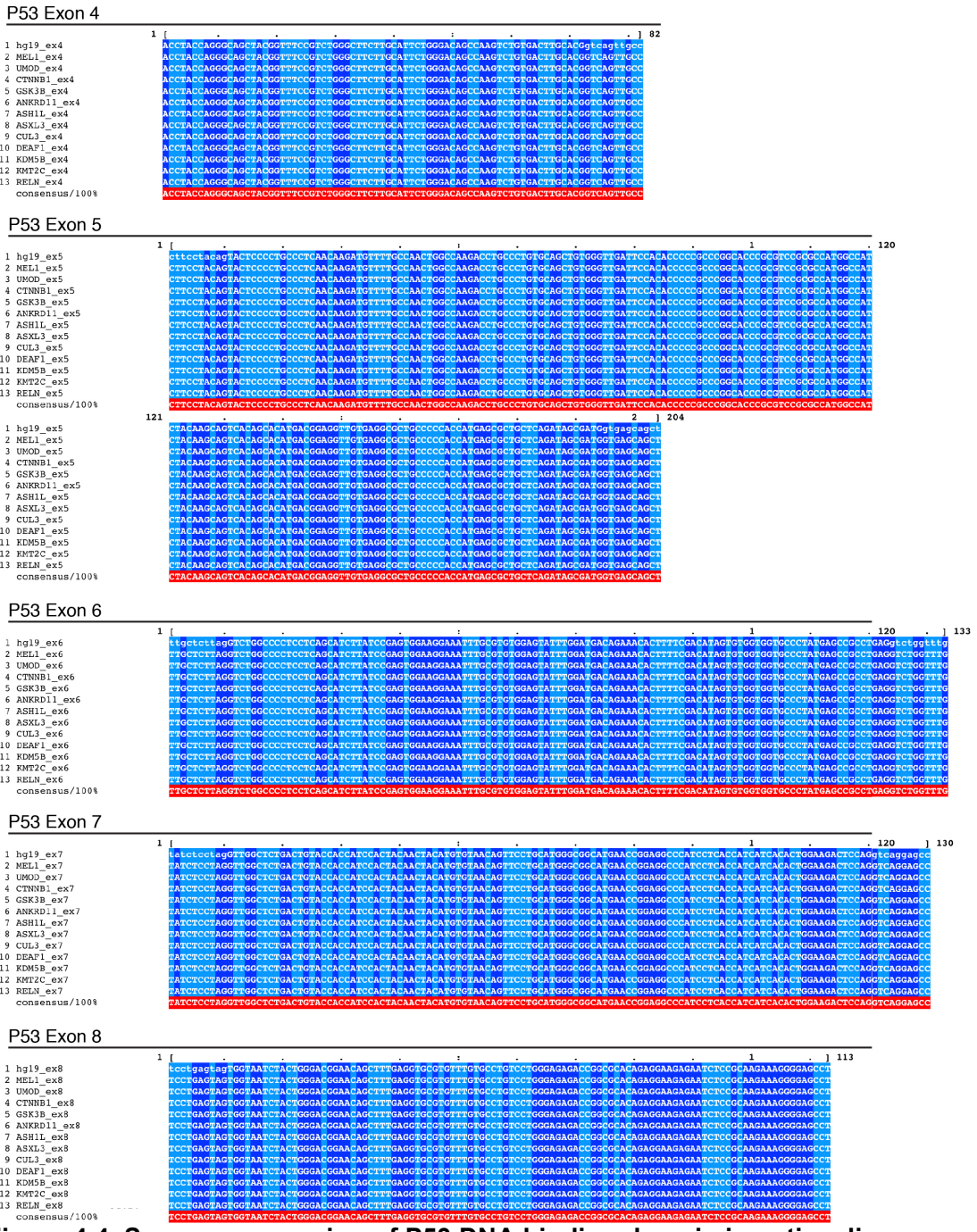


Figure 4.4. Sanger sequencing of P53 DNA binding domain in autism lines. Sequencing of the P53 DNA binding domain (amino acids 102-292) and intervening exon-intron boundaries did not identify mutations in control lines (MEL1 founder, UMOD, CTNNB1, and GSK3β) or Class 1 lines (ANKRD11, ASH1L, ASXL3, CUL3, DEAF1, KDM5B, KMT2C, and RELN), which showed competitive growth advantages. Reference genome GRCh37/hg19 is shown in top row, and consensus from sequencing results is shown in the bottom row (red). Ex, exon. Experiments performed and analyzed by Gustav Cederquist and Lorenz Studer.

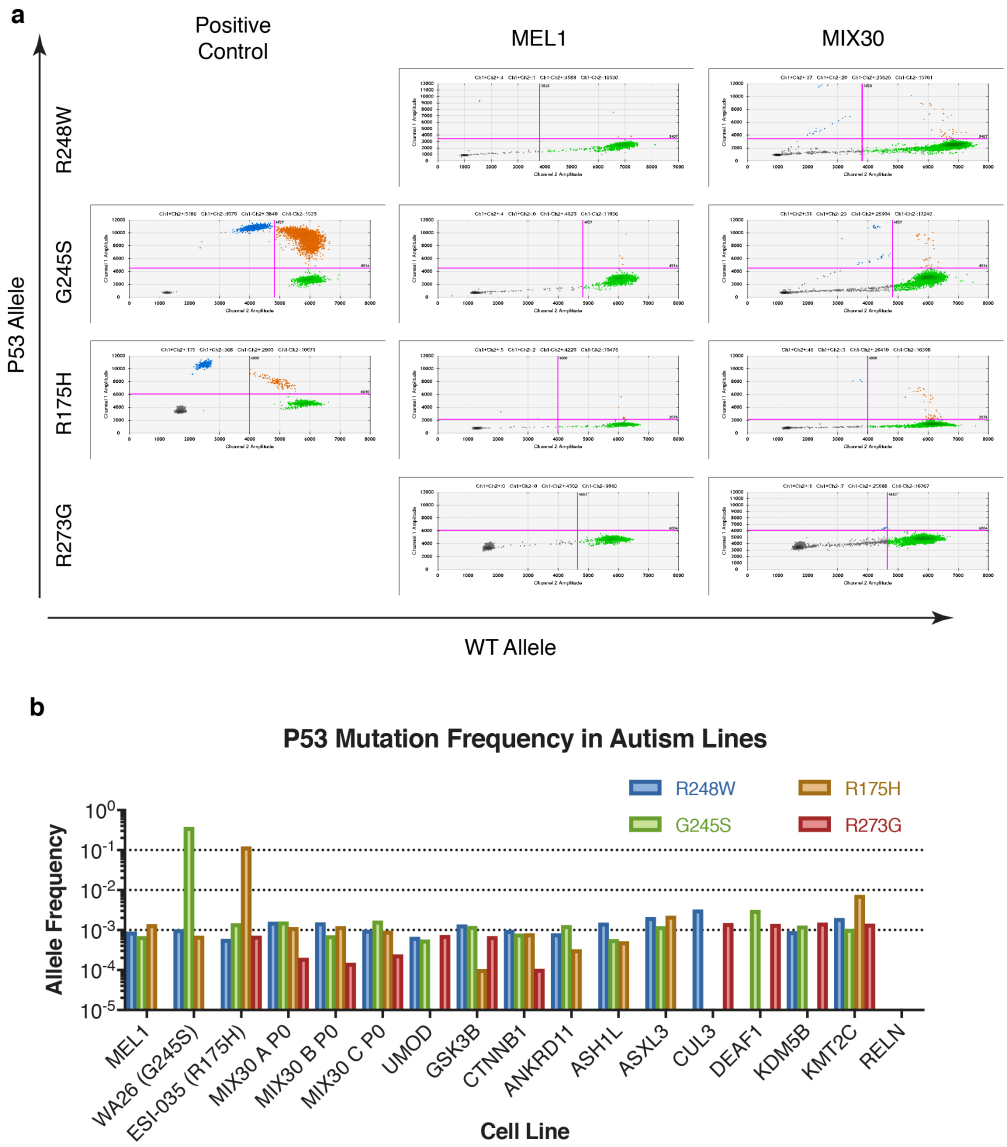


Figure 4.5. ddPCR for common P53 DNA binding domain mutations in autism lines.

(a) ddPCR analysis was used to assess minor allele frequency of P53 DNA binding domain mutations that have been reported to arise spontaneously in pluripotent stem cells(Merkle et al., 2017). Positive controls were available for two of four reported mutations (WA-026 (G245S), ESI-035 (R175H)). (**b)** Bar plots summarizing results of ddPCR analysis on control lines (MEL1 founder, UMOD, CTNNB1, and GSK3β) and Class 1 lines (ANKRD11, ASH1L, ASXL3, CUL3, DEAF1, KDM5B, KMT2C, and RELN), which showed competitive growth advantages. Typically, minor allele frequencies were estimated to be ~0.1% or less, though the R175H mutation may approach 1% frequency in the KMT2C line. ddPCR, droplet digital PCR. Experiments performed and analyzed by Gustav Cederquist and Lorenz Studer.

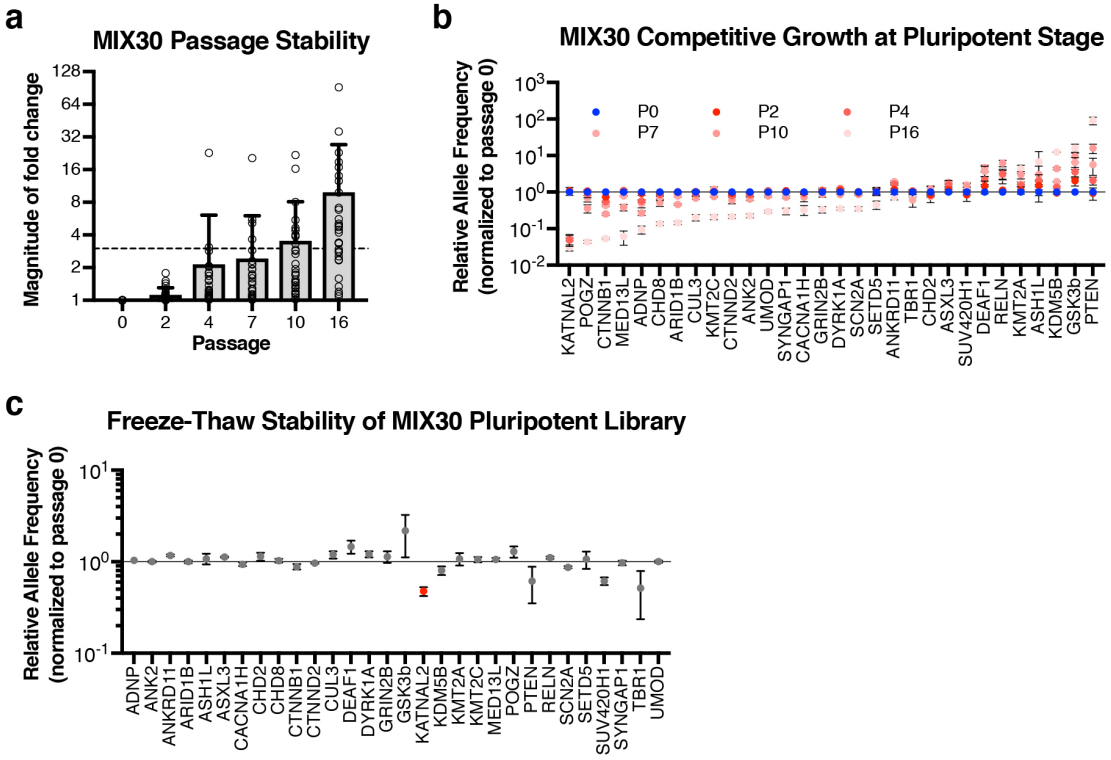


Figure 4.6. Stability of MIX30 pools.
(a) Average magnitude of allele frequency fold change in MIX30 library, normalized to passage 0, remained < 3x for 7 passages. Individual data points represent average fold change per line across n = 3 MIX30 pools. mean ± s.d. **(b)** Competitive growth dynamics of all lines in MIX30 library at pluripotent stage. Lines with selective growth advantage (e.g. PTEN and GSK3β) appear to suppress growth of most other lines by passage 16. n = 3 MIX30 pools, mean ± s.e.m. **(c)** Allele frequencies measured one day after thawing are largely unaffected by freeze-thaw cycle. Red indicates line with significantly reduced allele frequency (FDR < 0.05). n = 3 MIX30 pools, mean ± s.e.m. Experiments performed and analyzed by Gustav Cederquist and Lorenz Studer.

A second key feature of the proposed, hPSC-based multiplex platform is that it can provide access to disease-relevant cell-types through directed differentiation of hPSCs. Since the PFC is a major locus of autism pathology (Willsey et al., 2013), we devised a strategy utilizing FGF8b, a classic organizer of anterior cortical development *in vivo* (Fukuchi-Shimogori and Grove, 2001), to pattern cortical

progenitors to a PFC-like identity (**Figure 4.7a**). By day 18, PFC cultures are composed of near homogenous neuroepithelial rosettes of dorsal telencephalic identity (**Figure 4.7b, Figure 4.8 a,b**) and express high levels of the frontal cortex marker SP8 but low levels of the occipital cortex (OCC) marker COUPTF1 (O'Leary et al., 2007) (**Figure 4.7b, Figure 4.8c**). The BrainSpan transcriptional atlas was used to define 14 gene transcripts that are differentially expressed between human fetal PFC and OCC at PCW8 (**Figure 4.7c**). qRT-PCR analysis revealed a high correlation of PFC versus OCC markers between *in vivo* and *in vitro* derived tissue (**Figure 4.7c, Figure 4.8d**, $R^2 = 0.6191$, $p = 0.0008$).

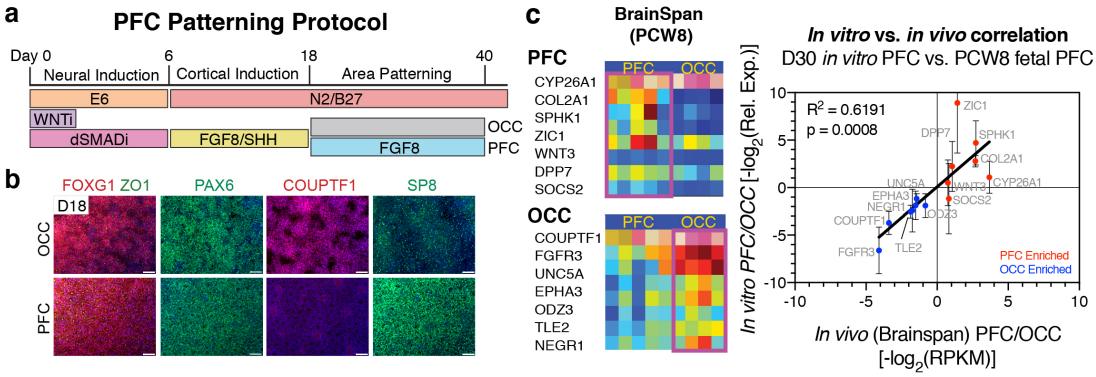
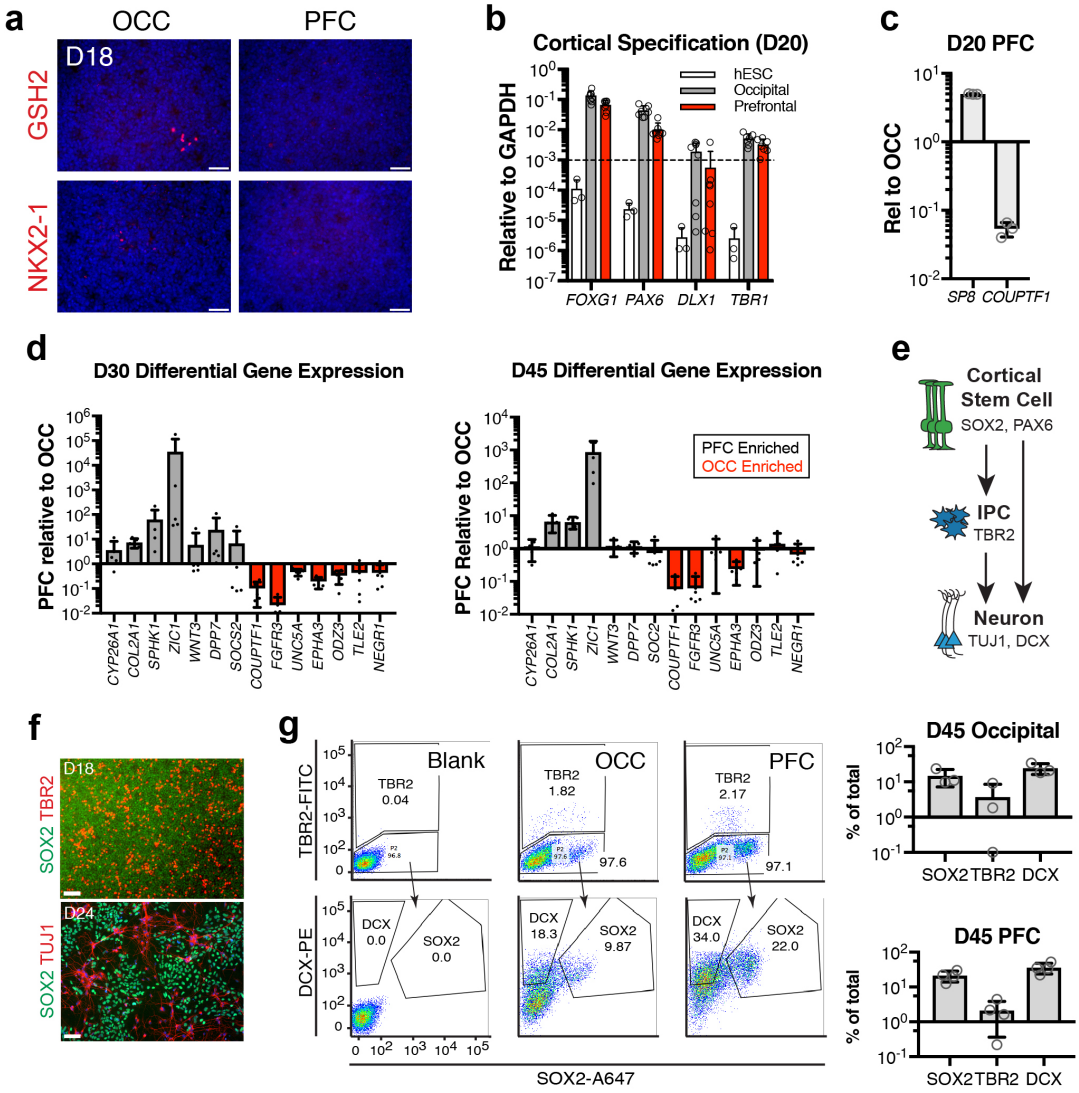


Figure 4.7. Protocol and generation of hPSC-derived prefrontal cortex tissue. **(a)** Schematic illustration for hPSC-derived PFC cultures. **(b)** PFC and OCC cultures express appropriate regional markers. FOXG1 and PAX6 indicate a general cortical identity, while SP8 and COUPTF1 are PFC and OCC markers, respectively. n = 3 differentiations. **(c)** Differential transcript expression of 14 genes between human fetal PFC and OCC, defined using BrainSpan transcriptional atlas (Brainspan.org), were highly correlated *in vitro* and *in vivo* ($R^2 = 0.6191$, $p = 0.0008$). n = 5 differentiations, mean ± s.d. ddPCR, droplet digital PCR. dSMADi, dual-smad inhibitors SB431542 and LDN193189; E6, Essential 6 base media; N2/B27, base media; PCW, post-conception week; PFC, prefrontal cortex; OCC, occipital cortex; WNTi, tankyrase inhibitor XAV939. Scale bars = 100 μm. Experiments performed and analyzed by Gustav Cederquist, Jason Tchieu and Lorenz Studer.

Figure 4.8. Additional characterization of hPSC-derived PFC.
(a) Immunocytochemistry for NKX2-1 and GSH2 shows that PFC and OCC differentiations contain few ventral telencephalic progenitors at day 18. n = 3 differentiations. (**b)** qRT-PCR analysis of cortical identity genes at day 20 shows high expression of *FOXG1* and *PAX6*. Contamination of ventral telencephalic cells was assessed by *DLX1* expression (dotted line, quality control cut-off > 0.098% of *GAPDH* expression). n = 8 differentiations. (**c)** qRT-PCR analysis of differential gene expression between day 20 PFC and OCC cultures for *SP8* (4.972 ± 0.045, $p < 0.0001$) and *COUPTF1* (0.05339 ± 0.013, $p < 0.0001$). n = 3 differentiations, p values from two-tailed student t-test. (**d)** Day 30 and day 45 qRT-PCR analysis for differential gene expression between PFC and OCC of regionally enriched transcripts defined in Figure 4.7. n ≥ 5 differentiations. (**e)** Simplified schematic illustration of cortical neurogenesis. (**f)** Representative immunocytochemistry for SOX2, TBR2, and TUJ1 demonstrates appropriate neurogenic capacity during PFC differentiation. (**g)** Day 45 fixed intracellular flow cytometry analysis and percent of total quantification for SOX2, TBR2, and DCX populations in PFC (SOX2: 21.18 ± 7.36, TBR2: 2.12 ± 1.76, DCX; 35.88 ± 12.34) and OCC (SOX2: 14.86 ± 7.68, TBR2: 3.71 ± 4.84, DCX: 24.37 ± 8.28) differentiations. n = 4 differentiations. PFC, prefrontal cortex; OCC, occipital cortex; QC, quality control. Scale bars = 50 μm. Error bars are s.d. Experiments performed and analyzed by Gustav Cederquist and Lorenz Studer.


a
OCC
PFC
D18
GSH2
NKX2-1
b
Cortical Specification (D20)
hESC
Occipital
Prefrontal
Relative to GAPDH
FOXG1 PAX6 DLX1 TBR1
c
D20 PFC
Rel to OCC
SP8 COUPTF1
d
D30 Differential Gene Expression
PFC relative to OCC
CYP26A1 COL2A1 SPHK1 ZIC1 WNT3 DPP7 SOCS2 COUPTF1 FGFR3 UNC5A EPHA3 ODZ3 TLE2 NEGR1
D45 Differential Gene Expression
PFC Relative to OCC
PFC Enriched
OCC Enriched
CYP26A1 COL2A1 SPHK1 ZIC1 WNT3 DPP7 SOC2 COUPTF1 FGFR3 UNC5A EPHA3 ODZ3 TLE2 NEGR1
e
Cortical Stem Cell
SOX2, PAX6
IPC
TBR2
Neuron
TUJ1, DCX
f
D18
SOX2 TBR2
D24
SOX2 TUJ1
g
Blank
OCC
PFC
TBR2-FITC
TBR2 0.04
TBR2 1.82
97.6
TBR2 2.17
97.1
DCX-PE
DCX 0.0
SOX2 0.0
DCX 18.3
SOX2 9.87
DCX 34.0
SOX2 22.0
SOX2-A647
D45 Occipital
% of total
SOX2 TBR2 DCX
D45 PFC
% of total
SOX2 TBR2 DCX

After establishing regional identity, we next sought to identify specific neurogenic cell-types within PFC cultures relevant to autism (Courchesne et al., 2011; Courchesne et al., 2007; Stoner et al., 2014). Neurons (DCX^+) are born from multipotent cortical neural stem cells ($SOX2^+$) or from proneural intermediate progenitor cells (IPCs, $TBR2^+$) (**Figure 4.8 e**). Immunocytochemistry and fluorescence activated cell sorting (FACS)-analysis revealed the presence of all three cell-types in our PFC culture system (**Figure 4.8 f,g**).

To test the impact of autism mutations on PFC neurogenesis, we differentiated MIX30 pools into day 45 PFC (**Figure 4.9**) and used FACS to isolate bulk (All), neural stem cell ($SOX2^+$), IPC ($TBR2^+$), and neuronal (DCX^+) fractions (**Figure 4.10, Figure 4.11 a-c**). While all autism lines showed a comparable efficiency for neural induction (**Figure 4.9 b**), 59% of autism lines (16/27, FDR < 0.05 (14 lines) and FDR < 0.1 (2 lines)) showed abnormal PFC neurogenesis as assessed by neuronal production (DCX/SOX2 ratio) and neural stem cell enrichment (SOX2/All ratio) (**Figure 4.10 b, Table S2**). The GSK3β control line showed a strong neurogenic phenotype in agreement with studies in mouse (Ahn et al., 2014; Marcus et al., 1998). The substantial enrichment in number of hits during PFC neurogenesis compared with hits during neural induction (**Figure 4.10 c**) demonstrates specificity and suggests a low false positive rate for PFC phenotypes.

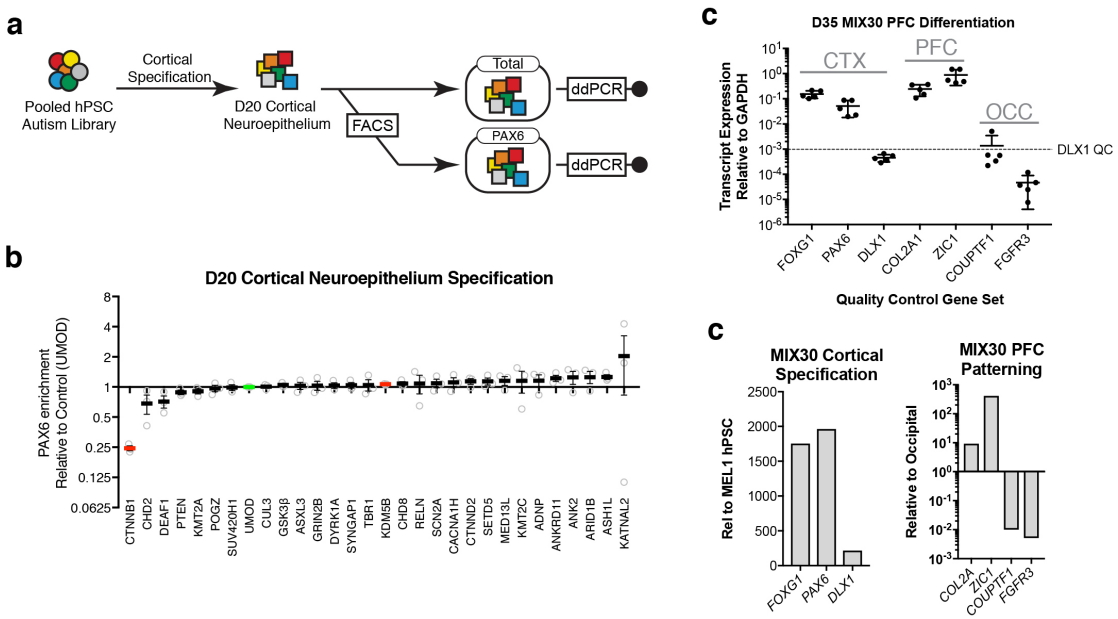
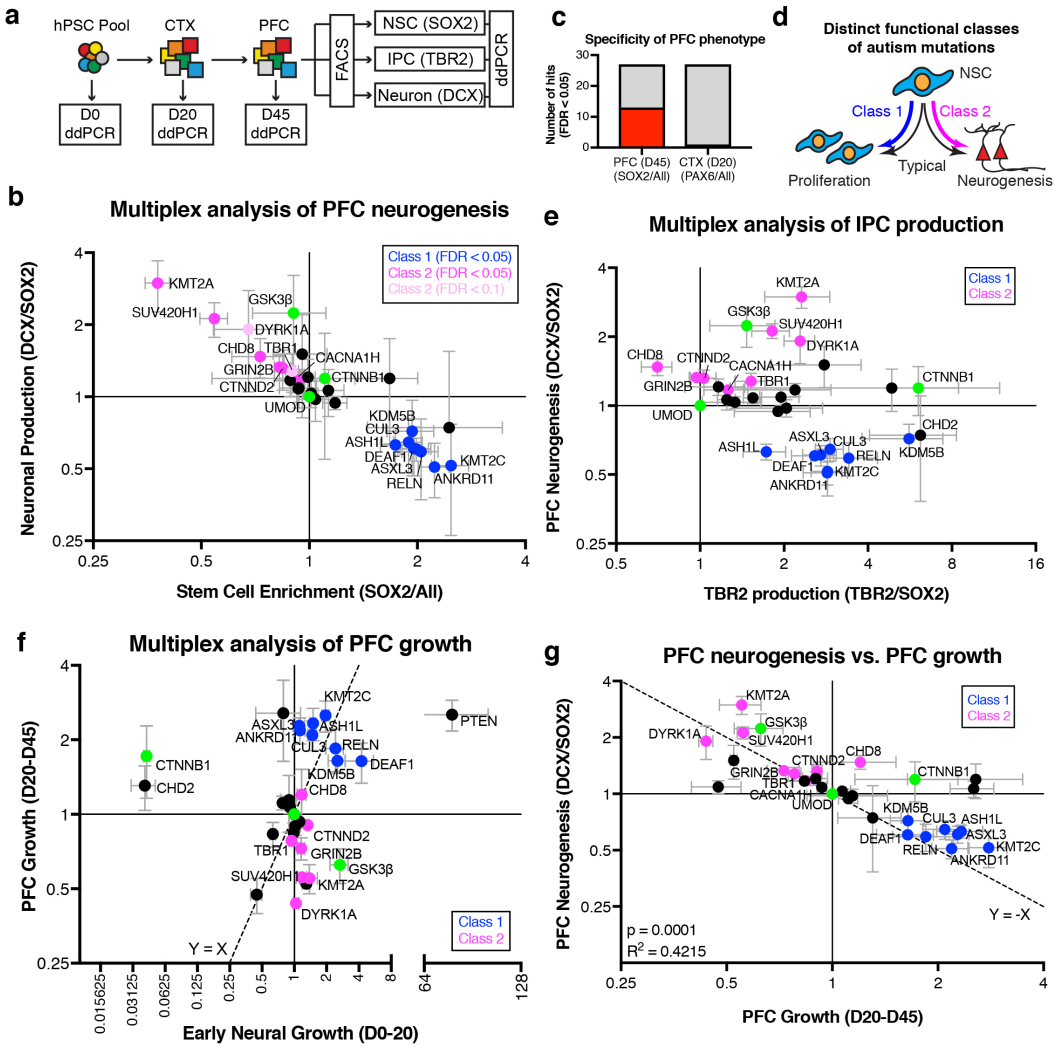
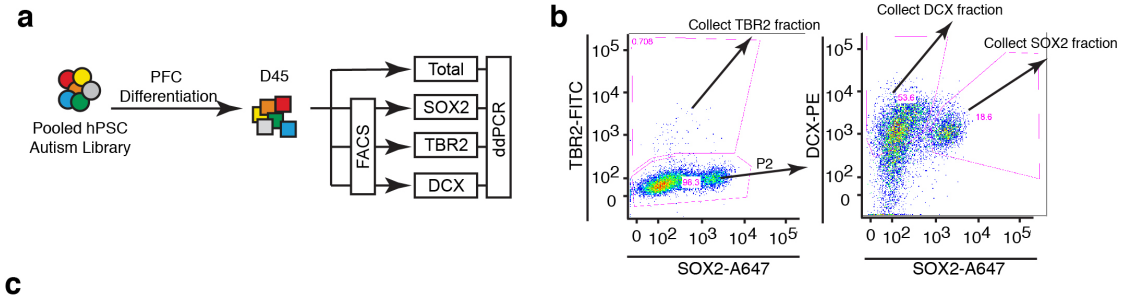


Figure 4.9. Characterization of MIX30 pooled PFC differentiation.
(a) Schematic illustration of multiplex strategy to test autism lines for appropriate cortical patterning during PFC differentiation. At day 20, the PAX6^{+} population is isolated using FACS. (**b)** Average relative cell line enrichment in PAX6^{+} fraction, relative to an unsorted day 20 MIX30 fraction. Red bars indicate cell lines with significant increases or decreases in enrichment score compared to UMOD (FDR < 0.05). n = 3 differentiations, error bars are s.e.m. (**c)** *GAPDH* normalized qRT-PCR analysis of D35 MIX30 PFC differentiation shows appropriate PFC patterning. n = 5 differentiations from three MIX30 pools, error bars are s.d. (**d)** Mean values from **c** plotted relative to controls from figure 4.8. PFC, prefrontal cortex; FACS, fluorescent activated cell sorting. Experiments performed and analyzed by Gustav Cederquist and Lorenz Studer.

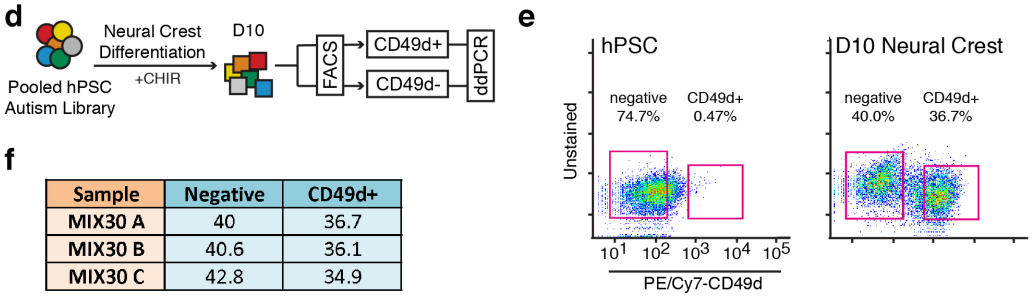
Figure 4.10. Multiplex analysis reveals that autism mutations alter PFC neurogenesis in a class-specific manner.
(a) Schematic illustration of multiplex strategy to test autism mutations for alterations in PFC growth and neurogenesis. Growth phenotypes are determined by measuring changes in allele frequency from D0 to D20 (early neural growth), and D20 and D45 (PFC growth). PFC neurogenesis phenotypes are determined by measuring changes in allele frequency between NSC (SOX2), IPC (TBR2), and Neuron (DCX) sorted fractions. **(b)** Scatter plot of multiplex neurogenesis assay showing changes in neuronal production (DCX/SOX2 ratio) and stem cell enrichment (SOX2/All ratio), normalized to a negative control (UMOD). Class 1 mutations (8/27, blue: FDR < 0.05) exhibit decreased neuronal production and increased stem cell enrichment. Class 2 mutations (8/27, magenta: FDR <0.05, light pink: FDR <0.1) exhibit increased neuronal production and decreased stem cell enrichment. n = 5 differentiations from three MIX30 pools. **(c)** 14/27 cell lines exhibited a stem cell enrichment phenotype (SOX2/All) during PFC development at day 45, while only 1/27 lines exhibited a neural induction phenotype (PAX6/All) during an earlier cortical development phase at day 20. Red and grey bars indicate number of cell lines with positive and negative phenotypes respective (FDR < 0.05). **(d)** Summary of class-specific PFC development phenotypes. Neural stem cell behavior is characterized by a balance between proliferation and neurogenesis (black arrows). Autism mutations skew this balance toward proliferation (Class 1, blue arrow) or neurogenesis (Class 2, magenta arrow). **(e)** Scatter plot of multiplex neurogenesis assay showing changes in IPC production (TBR2/SOX2 ratio) correlated with PFC neurogenesis phenotypes (DCX/SOX2 ratio), normalized to a negative control (UMOD). Class 1 mutations uniformly exhibit increased IPC production (FDR<0.05), while Class 2 mutations are variable in IPC production phenotypes. n ≥ 3 differentiations from at least two MIX30 pools. **(f)** Scatter plot of multiplex assay showing competitive growth phenotypes during early neural (D20/D0 ratio) and PFC growth phases (D45/D20 ratio), normalized to a negative control (UMOD). Class 1 mutations show increased PFC growth (6/8, FDR<0.05) while Class 2 mutations exhibit decreased PFC growth (6/8, FDR<0.05). Dotted line plots y = x values. n = 5 differentiations from three MIX30 pools. **(g)** Significant negative correlation between PFC neurogenesis and PFC growth parameters (R^2=0.4215, p=0.0001). Dotted line plots y = x values. All error bars are mean ± s.e.m. Internal standards are colored green. CTX, cerebral cortex; D, day; ddPCR, droplet digital PCR; FACS, fluorescent activated cell sorting; hPSC; human pluripotent stem cell; NSC, neural stem cell; PFC, prefrontal cortex; IPC, intermediate progenitor cell. Experiments performed and analyzed by Gustav Cederquist and Lorenz Studer.


a
hPSC Pool
CTX
PFC
FACS
NSC (SOX2)
IPC (TBR2)
Neuron (DCX)
ddPCR
D0 ddPCR
D20 ddPCR
D45 ddPCR
c
Specificity of PFC phenotype
Number of hits (FDR < 0.05)
PFC (D45) (SOX2/All)
CTX (D20) (PAX6/All)
d
Distinct functional classes of autism mutations
NSC
Class 1
Class 2
Proliferation
Typical
Neurogenesis
b
Multiplex analysis of PFC neurogenesis
Class 1 (FDR < 0.05)
Class 2 (FDR < 0.05)
Class 2 (FDR < 0.1)
Neuronal Production (DCX/SOX2)
Stem Cell Enrichment (SOX2/All)
e
Multiplex analysis of IPC production
PFC Neurogenesis (DCX/SOX2)
TBR2 production (TBR2/SOX2)
f
Multiplex analysis of PFC growth
PFC Growth (D20-D45)
Early Neural Growth (D0-20)
Y = X
g
PFC neurogenesis vs. PFC growth
p = 0.0001
R2 = 0.4215
Y = -X



c

	D45 PFC			+CHIR		
Sample	**SOX2**	**TBR2**	**DCX**	**SOX2**	**TBR2**	**DCX**
P4R1 A	18.6	0.708	53.6	35	7.76	49.6
P4R1 B	28	2.57	43.8	33.9	6.72	45.9
P2R2 A	20.2	1.28	32.5	23	2.79	34.8
P2R2 B	27.9	0.508	23	Not Treated		
P2R2 C	22.9	0.312	37.9	36.2	0.796	40.9



f

Sample	Negative	CD49d+
MIX30 A	40	36.7
MIX30 B	40.6	36.1
MIX30 C	42.8	34.9

Figure 4.11. FACS-based strategy for multiplex assays.
(**a**) Schematic illustration of multiplex strategy to test autism lines for PFC neurogenesis phenotypes. At day 45, MIX30 pools are separated into bulk, stem cell ($SOX2^+$), IPC ($TBR2^+$), and neuronal (DCX^+) fractions using FACS. Allele frequencies are calculated for each fraction using ddPCR. (**b**) Representative FACS plots showing isolation of $TBR2^+$ fraction, followed by isolation of $SOX2^+$ and DCX^+ fractions. CHIR experiments were performed using the same sorting strategy. **(c),** Percent of total values for each sorted fraction for all five replicates of multiplex neurogenesis assay (see Figure 4.10) and all four replicates of multiplex PFC WNT response assay (see Figure 4.16). **(d)** Schematic illustration of multiplex strategy to test autism lines for cranial neural crest phenotypes. At day 10, MIX30 pools are separated into CNC ($CD49^+$) and negative ($CD49d^-$) fractions using FACS. Allele frequencies are calculated for each fraction using ddPCR. **(e)** Representative FACS plots showing isolation of $CD49^+$ and $CD49d^-$ fractions during CNC differentiation. **(f)** Percent of total values for sorted fractions for all three replicates of multiplex neural crest assay (See Figure 4.16). CNC, cranial neural crest; ddPCR, droplet digital PCR; FACS, fluorescent activated cell sorting; IPC, intermediate progenitor cell; PFC, prefrontal cortex. FACS: Gustav Cederquist and Kirin Ramnarine. Data analysis Gustav Cederquist and Lorenz Studer.

Abnormal patterns of neurogenesis fell into two distinct classes (**Figure 4.10 d**). Class 1 mutations (ANKRD11, ASH1L, ASXL3, CUL3, DEAF1, KDM5B, KMT2C, and RELN) showed neural stem cell enrichment and decreased neuronal output. Class 1 mutations uniformly exhibited increased IPC production (**Figure 4.10 e**), suggesting a block in cell-cycle exit. Class 2 mutations (CACNA1H, CTNND2, CHD8, DYRK1A, GRIN2B, KMT2A, TBR1, and SUV420H1) showed neural stem cell depletion and increased neuronal output (**Figure 4.10 b**). In general, Class 1 mutations exhibited increased PFC growth and Class 2 mutations exhibited decreased growth (**Figure 4.10 f**). The significant negative correlation between PFC growth and neurogenesis (**Figure 4.10 g,** $R^2 = 0.4215$, $p = 0.0001$) illustrates that autism mutations coordinately dysregulate proliferation and differentiation, highlighting neural stem cells of the PFC as a convergence point in autism (**Figure 4.10 d**).

PFC neurogenesis phenotypes were validated using single genotype differentiations for six Class 1 lines (**Figure 4.12**). In addition, studies of neurogenesis in animal models confirm the overall findings for ANKRD11 (Gallagher et al., 2015), ARID1B (Jung et al., 2017), CHD2 (Shen et al., 2015), CHD8 (Durak et al., 2016; Gompers et al., 2017), DYRK1A (Kurabayashi and Sanada, 2013), KMT2A (Huang et al., 2015), and RELN (Deguchi et al., 2003; Hammond et al., 2010; Johnson et al., 2015; Lakoma et al., 2011; Nowakowski et al., 2017), but not PTEN (Chen et al., 2015b). In addition, correlations with data from pilot studies that used 8-line mixtures (MIX8) demonstrated reproducibility and stability of PFC neurogenesis phenotypes to changes in pool number (**Figure 4.13**). PFC neurogenesis phenotypes did not correlate with biologically unrelated assays including hPSC growth (**Figure 4.14 a**), SSEA4 expression (**Figure 4.14 b**), early neural growth (**Figure 4.14 c**), cortical patterning (**Figure 4.14 d**), or cell line zygosity (**Figure 4.14 e**). The coefficient of variation was typically between 0.2 and 0.3 for each assay (**Figure 4.15**)

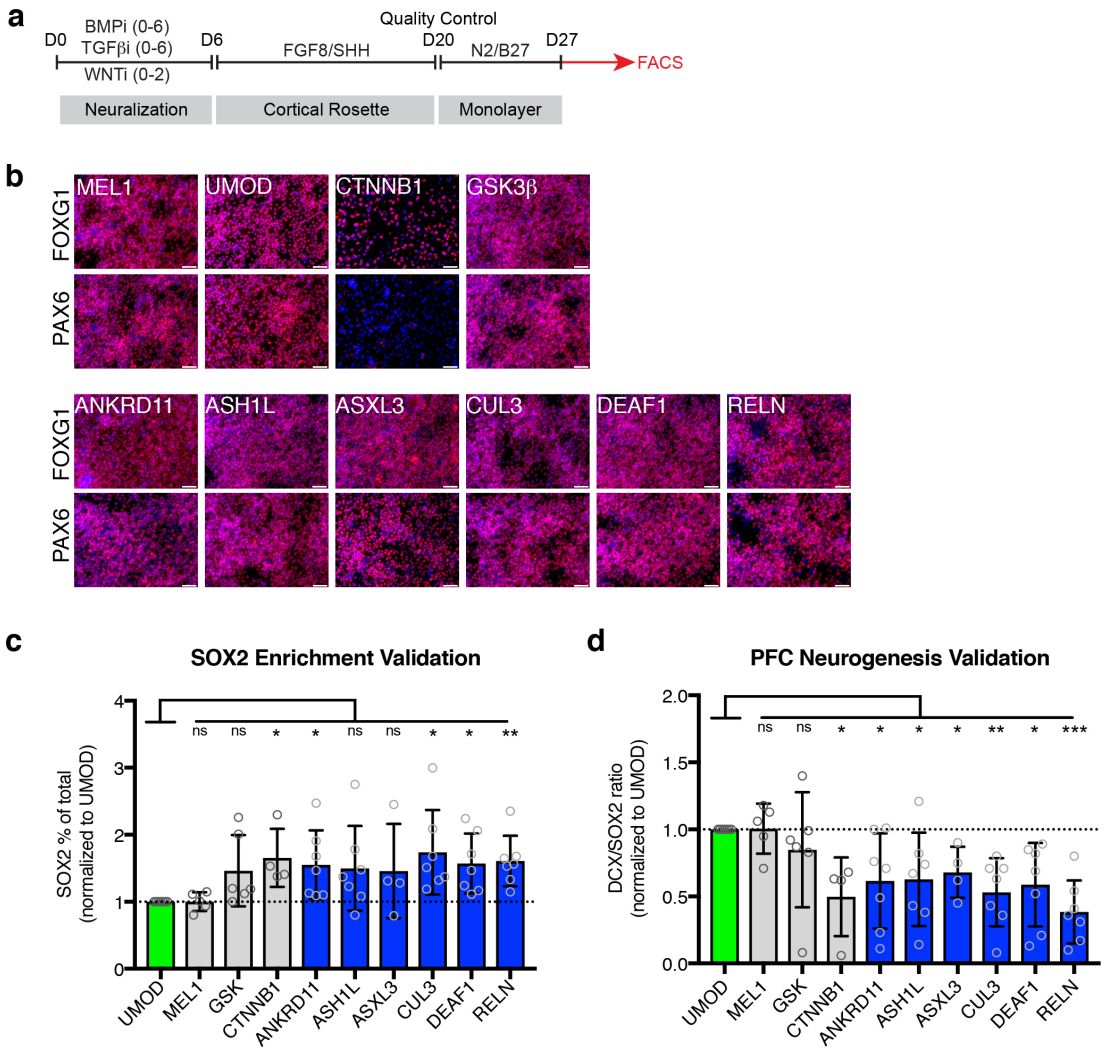
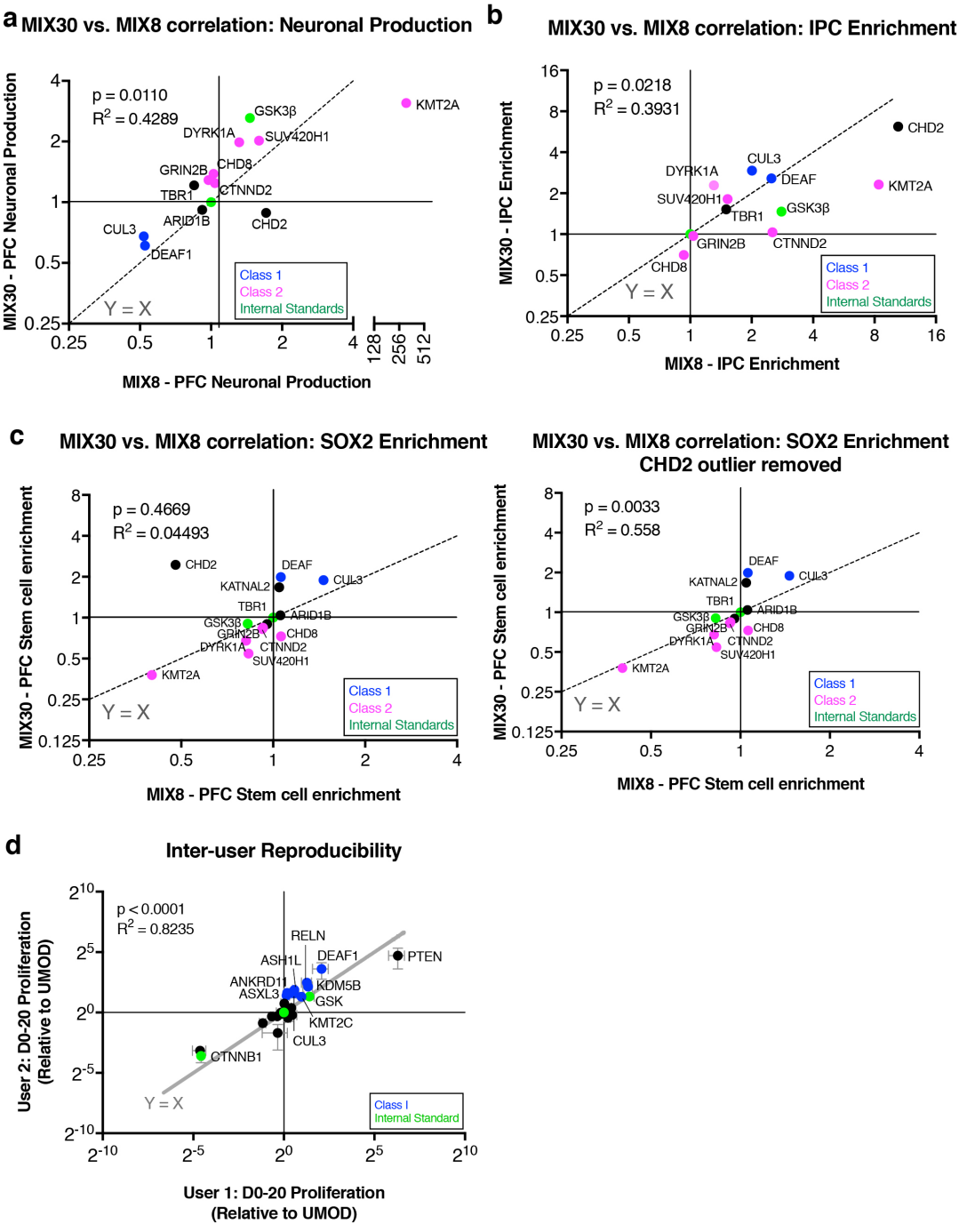


Figure 4.12. Validation of multiplex assay using single genotype differentiations.

(**a**) Schematic illustration of single genotype neurogenesis assay. Quality control was performed at day 20 to ensure appropriate patterning for each line. Lines were passaged at day 20 to a saturating density of 10^6 cells / cm^2 and allowed to undergo neurogenesis for 7 days. A negative control line, UMOD, was included in each differentiation batch for normalization. **(b)** Day 20 immunocytochemistry for FOXG1 and PAX6 to confirm appropriate cortical patterning. **(c)** Percent of SOX2 cells per differentiation, normalized to UMOD. 4/6 class I genes (blue) showed the expected increase in SOX2 percentage (FDR < 0.05), while MEL1 was similar to UMOD. n ≥ 5 differentiations. **(d)** DCX/SOX2 ratio for each line, normalized to UMOD. 6/6 class I genes (blue) showed the expected increase in DCX/SOX2 ratio (FDR < 0.05), while MEL1 was similar to UMOD. n ≥ 5, FACS, fluorescent activated cell sorting. Error bars are s.d. PFC Differentiations: Jason Tchieu; FACS: Gustav Cederquist and Kiran Ramnarine; Data analysis: Gustav Cederquist and Lorenz Studer.

Figure 4.13. Reproducibility of multiplex assay with different pool numbers and different users.
(a-c) Correlation between MIX30 experiments and pilot studies that used mini-pools of only 8 lines (MIX8) demonstrates that multiplex assay phenotypes are stable to changes in pool number and reproducible over time. Positive correlations were found between neuronal production (**a**, R^2 = 0.4289, p = 0.011) and IPC production (**b**, R^2 = 0.3931, p = 0.022). SOX2 enrichment did not show a positive correlation (**c** left, R^2 = 0.04493, p = 0.4669), though this was driven by a single CHD2 data point. In MIX8 cultures, CHD2 absolute allele frequency was near the limit of detection (< 1:13,000 in SOX2 fraction), precluding reliable analysis. When this data point was removed, the SOX2 correlation became significant (**c** right, R^2 = 0.558, p = 0.0033). Time point for MIX30 data is day 45, and time point for MIX8 data is day 55. MIX30 data is from **Fig. 2**. n = 1 differentiation for MIX8 data. (**d)** High reproducibility for day 0-20 MIX30 growth phenotypes when assays are performed by two different users (R^2 = 0.8235, p < 0.0001). User 1 data is from Figure 4.10. n = 3 differentiations from 3 MIX30 pools from user 2. Gray line is plotted at Y = X. Error bars are s.e.m. Experimental Design and Analysis: Gustav Cederquist and Lorenz Studer.


a MIX30 vs. MIX8 correlation: Neuronal Production
p = 0.0110
R2 = 0.4289
MIX30 - PFC Neuronal Production
MIX8 - PFC Neuronal Production
Y = X
Class 1
Class 2
Internal Standards
b MIX30 vs. MIX8 correlation: IPC Enrichment
p = 0.0218
R2 = 0.3931
MIX30 - IPC Enrichment
MIX8 - IPC Enrichment
c MIX30 vs. MIX8 correlation: SOX2 Enrichment
p = 0.4669
R2 = 0.04493
MIX30 - PFC Stem cell enrichment
MIX8 - PFC Stem cell enrichment
MIX30 vs. MIX8 correlation: SOX2 Enrichment
CHD2 outlier removed
p = 0.0033
R2 = 0.558
d Inter-user Reproducibility
p < 0.0001
R2 = 0.8235
User 2: D0-20 Proliferation (Relative to UMOD)
User 1: D0-20 Proliferation (Relative to UMOD)
Class I
Internal Standard

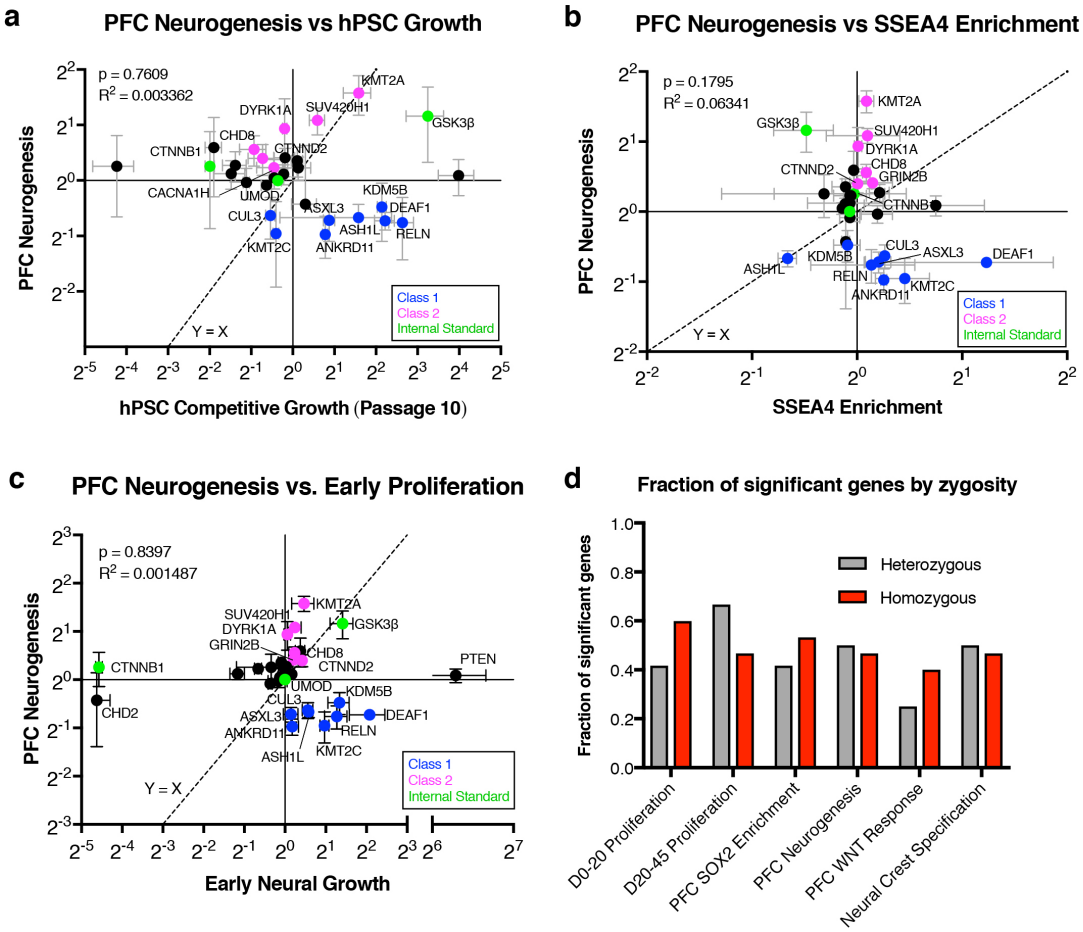
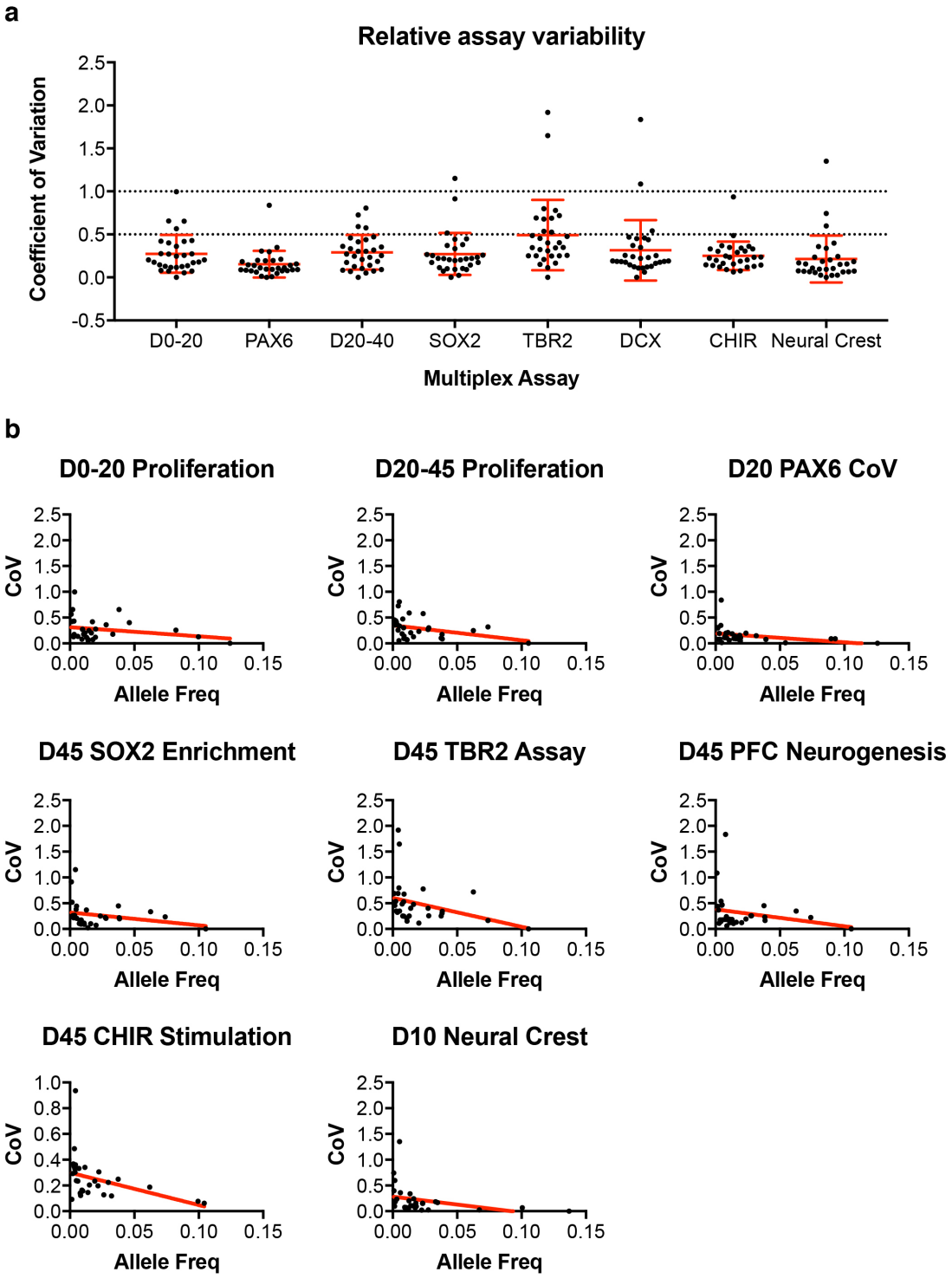


Figure 4.14. PFC neurogenesis phenotypes do not correlate with hPSC characteristics, early neural growth, or zygosity.
(a) No correlation between cell line growth at the pluripotent stage (Figure 4.3) and PFC neurogenesis phenotype (Figure 4.10) (R^2 = 0.00336, p = 0.7809). **(b)** No correlation between SSEA4 enrichment (Figure 4.4) and PFC neurogenesis phenotype (R^2 = 0.06341, p = 0.1795). **(c)** No correlation between early neural growth (Figure 4.10) and PFC neurogenesis (R^2 = 0.001487, p = 0.8397). **(d)** No correlation between cortical induction (Figure 4.9) and PFC neurogenesis (R^2 = 0.003348, p = 0.7614). (e) No apparent bias for homozygous or heterozygous mutations for any examined phenotypes. Gray line is plotted at Y = X. Error bars are s.e.m. Experimental Design and Analysis: Gustav Cederquist and Lorenz Studer.

Figure 4.15. Variability of multiplex assay.
(a) Coefficient of variation for 8 multiplex assays (D0-20: early neural growth, Fig. 2e; PAX6: cortical specification, Fig. S5; D20-45: PFC growth, Fig. 2e; SOX2: PFC stem cell enrichment, Fig. 2b; TBR2: IPC production, Fig. 2c; DCX: PFC Neuronal production, Fig. 2b; CHIR: PFC WNT response, Fig. 3b; Neural Crest: neural crest WNT response, Fig. 3d). Each dot represents the mean value for an individual cell line. Red lines show mean ± s.d. for all cell lines. **(b)** Coefficient of variation plotted as a function of total allele frequency. Relative variation increases as allele frequency decreases. IPC, intermediate progenitor cell; PFC, prefrontal cortex. Experimental Design and Analysis: Gustav Cederquist and Lorenz Studer.


a
Relative assay variability
Coefficient of Variation
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
D0-20
PAX6
D20-40
SOX2
TBR2
DCX
CHIR
Neural Crest
Multiplex Assay
b
D0-20 Proliferation
D20-45 Proliferation
D20 PAX6 CoV
D45 SOX2 Enrichment
D45 TBR2 Assay
D45 PFC Neurogenesis
D45 CHIR Stimulation
D10 Neural Crest
CoV
Allele Freq
0.00
0.05
0.10
0.15

In addition to probing developmental phenotypes related to cell fate specification and proliferation, our multiplex platform also allows us to evaluate the cell-type specific activity of key molecular pathways. The WNT/βcatenin pathway is a critical regulator of stem cell proliferation and neurogenesis during cortical development (Chenn, 2008; Hirabayashi et al., 2004; Munji et al., 2011), and is a central node among a network of autism-related genes (Gilman et al., 2011; Krumm et al., 2014; Packer, 2016a). We therefore tested autism lines for the ability to respond to WNT/βcatenin signaling by treating day 35 MIX30 PFC cultures with the GSK3α/β inhibitor CHIR99021 (3μM) for 10 days, using stem cell proliferation as an initial readout of WNT activity (Kim et al., 2009) (**Figure 4.16 a**). Strikingly, Class 1 mutations were uniformly hyporesponsive to WNT-induced stem cell proliferation (**Figure 4.16 b**). Lineage specificity of the WNT-response phenotype was assessed by differentiating the MIX30 library to $CD49d^{+}$ cranial neural crest (CNC) precursors (Fattahi et al., 2016), whose specification is dependent on WNT/βcatenin activity (Dorsky et al., 1998) (**Figure 4.16 c**). Class 1 mutations were inefficient in CNC specification (**Figure 4.16 d**), further supporting class-specific WNT dysregulation.

The observed WNT-dependent defects in CNC development could explain the high rate of facial dysmorphism in some autism patients (Cordero et al., 2011; Miles et al., 2008). In fact, facial dysmorphism has been reported in patients for 7 out of 8 Class 1 genes (Balasubramanian et al., 2017; Faundes et al., 2018; Koemans et al., 2017; Ockeloen et al., 2015; Okamoto et al., 2017; Redin et al., 2017; Vulto-van Silfhout et al., 2014). To explore these clinical observations and to further validate our in vitro multiplex data, we generated mosaic F0 loss-of-function zebrafish of Class 1 genes and assessed lower jaw development, a parameter known to critically rely on WNT-dependent CNC function (Curtin et al., 2011; Dougherty et al., 2013; Kamel et al., 2013; Rochard et al., 2016). ANKRD11, CUL3, and KMT2C mutants significantly

increased the fraction of jaw hypomorphs, while ASH1L, DEAF1, and KDM5B mutants showed statistically non-significant increases (**Figure 4.17 a**). Together these data indicate that Class 1 genes are hyporesponsive to WNT signaling across multiple developmental lineages (**Figure 4.17 b**).

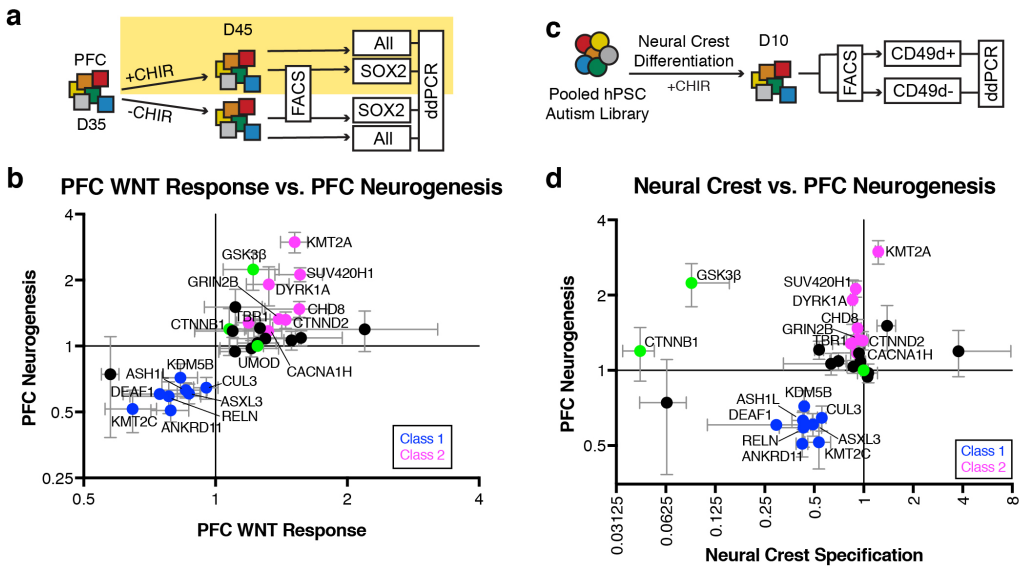


Figure 4.16. Class-specific dysregulation of WNT signaling.
(a) Schematic illustration of multiplex strategy to test autism mutations for WNT/βcatenin response during PFC growth. D35 Pooled PFC cultures are treated with the CHIR99021 (GSK3 inhibitor, 3μM) for 10 days and compared to untreated cultures. Stem cell enrichment is used as a read-out of WNT activity. **(b)** Correlation of PFC WNT response with PFC neurogenesis phenotype. 7/8 Class 1 mutations (blue) are hyporesponsive to WNT signaling (FDR < 0.05), relative to the negative control UMOD. n = 4 differentiations from three MIX30 pools. **(c)** Schematic illustration of multiplex strategy to test autism mutations for WNT/βcatenin response during neural crest development. MIX30 hPSC pools are differentiated toward neural crest for 10 days using an established WNT-dependent protocol. Allele frequencies are then compared between cranial neural crest positive (CD49d+) and negative (CD49d-) sorted fractions. **(d)** Correlation of neural crest specification with PFC neurogenesis phenotypes reveals that 8/8 Class 1 mutations are inefficient at cranial neural crest specification. Experiments performed and analyzed by Gustav Cederquist and Lorenz Studer.

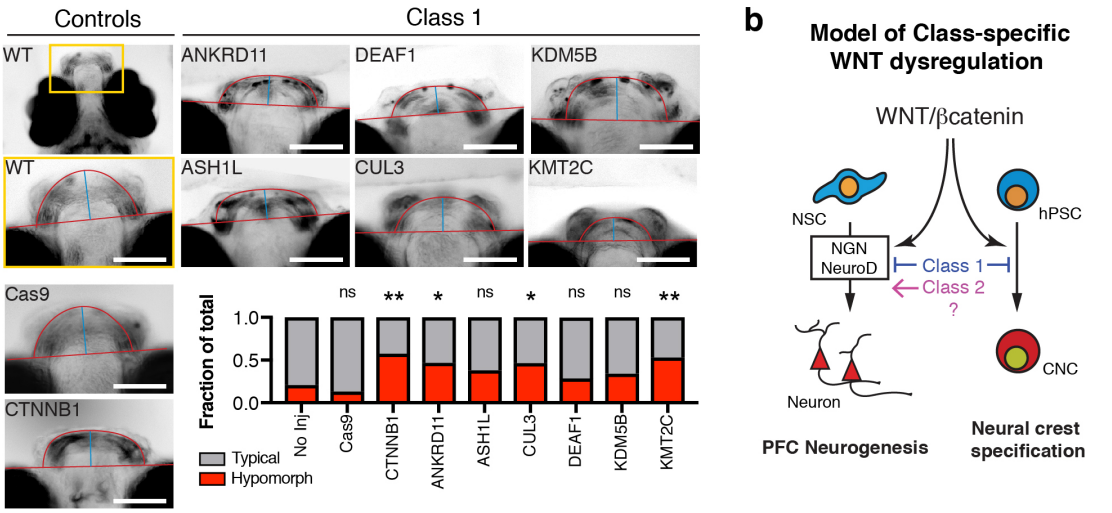


Figure 4.17. Validation of WNT phenotypes in Zebrafish.
(a) Analysis of zebrafish jaw development in F0 mosaic loss-of-function animals imaged ventrally at 7 dpf. Top left image depicts area of high magnification. Cas9 alone and CTNNB1 gRNA injections serve as controls. Class 1 mutations exhibit hypomorphic jaw phenotypes that resemble those of CTNNB1 mutants (CTNNB1 p = 0.0064, ANKRD11 p = 0.0488, CUL3 = 0.0488, KMT2C = 0.0088). p values calculated using Fisher's exact test corrected for multiple comparisons using Benjamini-Hochberg method. No injection n = 43, Cas9 alone n = 23, CTNNB1 n = 40, ANKRD11 n = 32, ASH1L n = 68, CUL3 n = 43, DEAF1 n = 28, KDM5B n = 41, KMT2C n = 47 fish. All injections performed on at least 2 clutches. Experimenter was blinded to genotype during analysis. **(b)** Potential model of WNT-dysregulation for class 1 and 2 autism mutations. WNT/βCatenin promotes neural crest specification and regulates cortical neurogenesis through transcriptional activation of proneural factors NGN and NeuroD. Class 1 mutations block WNT-dependent responses during neural crest specification and neurogenesis. All error bars in figure 3 are mean ± s.e.m. Internal standards are colored green. Scale bars, 100μm. Dpf, days post fertilization. Experimental design: Scott Callahan, Richard White; Zebrafish husbandry, injections, imaging: Scott Callahan. Image and data analysis: Gustav Cederquist, Scott Callahan, Richard White, Lorenz Studer.

Table 2. Summary of phenotypes from multiplex experiments

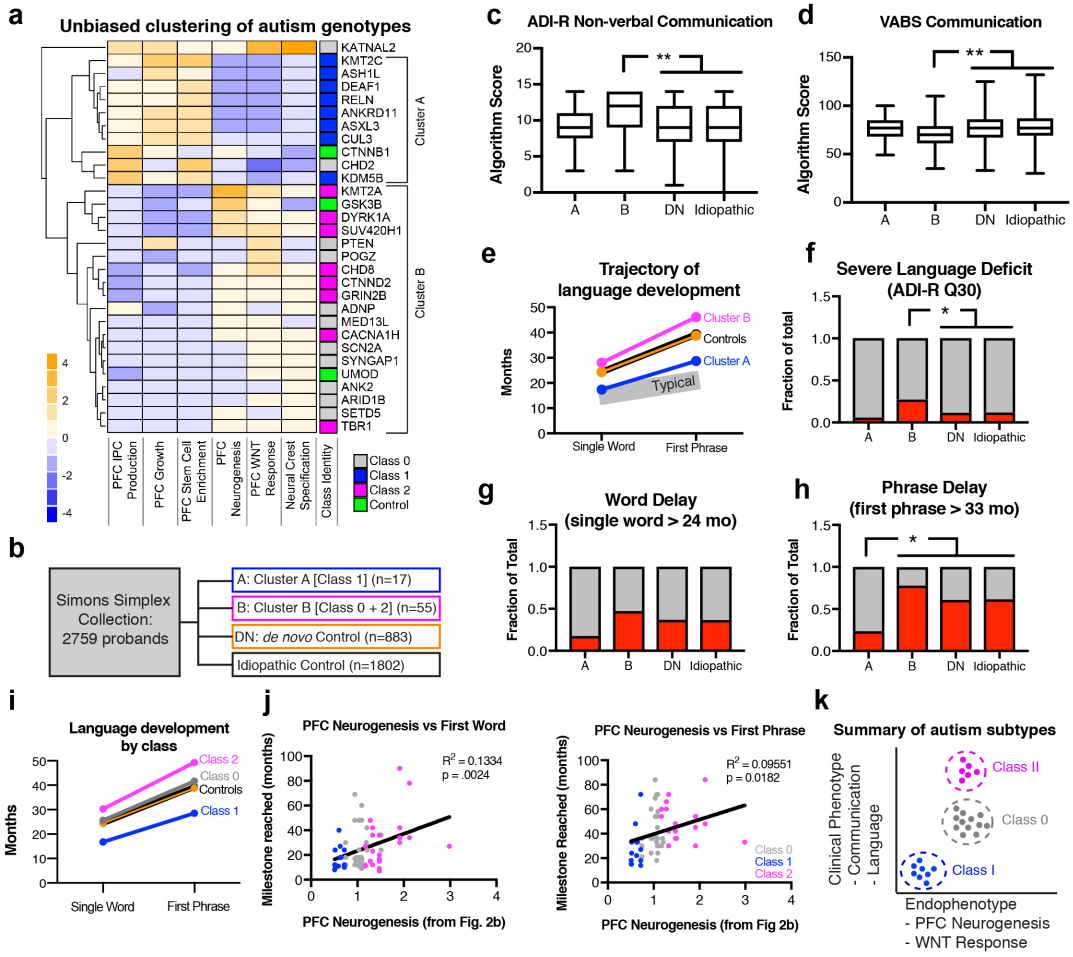
Table S2: Summary of Multiplex Phenotype						Significance		
						FDR < 0.05	FDR < 0.10	
Gene Mutation	Neuroepithelial Growth	PFC Growth	PFC SOX2 Enrichment	PFC Neurogenesis	IPC Production	PFC WNT Response	Neural Crest Specification	
Class 1								
ANKRD11	4.055E-01	7.375E-03	2.518E-03	4.380E-04	6.952E-03	1.881E-02	6.891E-05	
ASH1L	3.565E-03	1.304E-02	4.891E-03	6.884E-04	2.843E-02	4.200E-02	7.236E-04	
ASXL3	5.412E-01	6.494E-04	3.485E-03	1.314E-03	1.809E-03	1.731E-02	1.431E-06	
CUL3	2.417E-02	1.304E-02	3.485E-03	5.475E-03	4.342E-04	3.236E-02	1.432E-04	
DEAF1	5.321E-02	1.044E-01	3.485E-03	2.753E-05	5.066E-03	4.200E-02	1.125E-02	
KDM5B	2.610E-02	1.529E-02	2.650E-02	6.466E-02	3.375E-02	1.881E-02	1.431E-06	
KMT2C	8.972E-04	2.888E-02	2.953E-04	9.938E-03	5.965E-04	1.881E-02	8.044E-03	
RELN	3.011E-02	7.799E-02	6.017E-03	9.628E-03	1.988E-02	6.114E-02	1.098E-02	
Class 2								
CACNA1H	8.831E-01	1.175E-01	1.095E-03	3.798E-02	2.859E-01	6.642E-01	4.025E-01	
CHD8	2.134E-01	5.618E-01	1.009E-01	1.285E-02	6.528E-03	5.882E-01	1.662E-01	
CTNND2	4.988E-05	2.586E-02	1.496E-02	4.658E-02	7.400E-01	3.727E-01	7.014E-01	
DYRK1A	5.795E-01	4.445E-08	8.524E-02	7.794E-02	1.420E-03	7.947E-01	2.468E-02	
GRIN2B	1.994E-01	1.529E-02	1.095E-03	6.766E-03	5.529E-01	6.642E-01	4.898E-01	
KMT2A	2.613E-01	1.382E-03	7.395E-07	1.314E-03	4.343E-02	1.590E-01	6.894E-02	
TBR1	3.390E-01	6.789E-04	5.155E-02	5.620E-02	1.646E-02	6.642E-01	2.980E-01	
SUV420H1	3.014E-01	2.481E-07	1.752E-04	6.884E-04	1.960E-02	3.815E-01	4.765E-03	
Class 0								
ADNP	3.961E-02	4.501E-09	8.596E-01	1.925E-01	1.101E-01	6.642E-01	1.315E-01	
ANK2	2.409E-04	1.891E-01	1.344E-01	9.233E-02	1.564E-01	4.536E-01	4.898E-01	
ARID1B	5.620E-01	6.418E-01	8.595E-01	7.874E-01	1.815E-01	9.448E-01	5.855E-01	
CHD2	2.722E-12	3.206E-01	2.681E-01	5.754E-01	2.759E-02	2.263E-04	1.667E-06	
KATNAL2	6.108E-01	1.703E-01	1.320E-01	5.569E-01	3.123E-01	6.642E-01	4.629E-01	
MED13L	8.831E-01	2.445E-01	8.596E-01	1.072E-01	5.313E-01	9.448E-01	1.432E-04	
POGZ	5.437E-05	6.789E-04	7.303E-01	4.151E-01	1.088E-02	6.642E-01	1.686E-02	
PTEN	2.610E-02	9.028E-03	3.973E-01	6.102E-01	2.818E-02	4.287E-02	2.316E-01	
SCN2A	2.694E-02	7.799E-02	1.688E-01	2.757E-01	1.056E-03	9.448E-01	2.980E-01	
SETD5	6.716E-05	1.600E-01	2.998E-01	8.813E-02	1.168E-01	6.642E-01	2.278E-01	
SYNGAP1	3.961E-02	3.219E-01	8.596E-01	6.120E-01	2.686E-01	1.000E+00	2.316E-01	
Internal Standards								
GSK3B	2.694E-02	1.304E-02	7.971E-01	4.658E-02	1.468E-01	9.448E-01	6.442E-06	
CTNNB1	5.572E-17	2.713E-01	8.596E-01	5.800E-01	3.547E-01	6.642E-01	1.111E-06	
UMOD	1.000E+00	1.000E+00	1.000E+00	1.000E+00	1.000E+00	1.000E+00	1.000E+00	

Phenotype Heatmap				Scale				
				0	0.5	1	1.5	2
Gene Mutation	Neuroepithelial Growth	PFC Growth	PFC SOX2 Enrichment	PFC Neurogenesis	IPC Production	PFC WNT Response	Neural Crest Specification	
WNT-Low								
ANKRD11	1.13	2.19	2.23	0.51	2.86	0.79	0.42	
ASH1L	1.49	2.33	1.74	0.63	1.73	0.86	0.43	
ASXL3	1.11	2.28	1.96	0.61	2.71	0.87	0.49	
CUL3	1.47	2.09	1.89	0.65	2.93	0.95	0.56	
DEAF1	4.23	1.64	1.99	0.60	2.57	0.75	0.29	
KDM5B	2.52	1.64	1.94	0.72	5.62	0.83	0.44	
KMT2C	1.96	2.79	2.48	0.22	2.86	0.73	0.53	
RELN	2.42	1.85	2.05	0.59	3.41	0.78	0.43	
Neurogenic								
CACNA1H	0.99	0.84	0.94	1.17	1.26	1.32	0.91	
CHD8	1.17	1.20	0.73	1.47	0.70	1.55	0.92	
CTNND2	1.34	0.91	0.84	1.32	1.03	1.45	0.99	
DYRK1A	1.04	0.44	0.68	1.91	2.29	1.32	0.86	
GRIN2B	1.16	0.73	0.83	1.33	0.97	1.40	0.92	
KMT2A	1.38	0.55	0.38	2.98	2.31	1.52	1.22	
TBR1	0.94	0.78	0.90	1.28	1.52	1.19	0.84	
SUV420H1	1.18	0.56	0.54	2.12	1.81	1.56	0.90	
Unclassified								
ADNP	1.30	0.52	0.96	1.51	2.79	1.11	1.39	
ANK2	0.78	1.11	1.18	0.94	1.90	1.11	1.05	
ARID1B	0.89	1.14	1.04	0.98	2.04	1.21	1.07	
CHD2	0.04	1.30	2.45	0.74	6.18	0.58	0.06	
KATNAL2	0.79	2.56	1.67	1.19	4.87	2.19	3.78	
MED13L	1.02	0.90	0.99	1.21	1.16	1.26	0.54	
POGZ	0.45	0.47	0.94	1.09	1.95	1.57	0.70	
PTEN	78.21	2.53	1.13	1.06	1.25	1.49	0.63	
SCN2A	1.12	0.93	0.93	1.08	1.55	1.30	0.96	
SETD5	0.63	0.83	0.88	1.17	2.19	1.09	0.94	
SYNGAP1	0.91	1.07	1.01	1.03	1.34	1.25	0.87	
Internal Standards								
GSK3B	2.67	0.63	0.90	2.24	1.47	1.22	0.13	
CTNNB1	0.04	1.72	1.10	1.19	6.06	1.07	0.04	
UMOD	1.00	1.00	1.00	1.00	1.00	1.25	1.00	

We next asked whether functional autism classes defined by our multiplex platform could define clinically distinct subgroups of autism patients, using proband data from the Simons Simplex Collection (SSC). To define cohorts in an unbiased manner that most accurately represented overall multiplex data, we performed hierarchical clustering of all lines across six phenotypic assays related to PFC development and WNT signaling. This analysis revealed two major functional groups (**Figure 4.18 a**). Cluster A contained all Class 1 mutations and the CTNNB1 control, further supporting a WNT-hyporesponsive phenotype. Cluster B included all Class 2 mutations and the GSK3β control, perhaps suggesting a contrasting relationship to WNT signaling. Based on this analysis, SSC patients were divided into 4 cohorts: (1) Cluster A (A, n = 17), (2) Cluster B (B, n = 55), (3) *de novo* control (DN, n = 883), (4) idiopathic control (Idiopathic, n = 1802) (**Figure 4.18 b**). Cohorts were similar in their demographic profiles (**Figure 4.19 a-d**), autism severity (**Figure 4.19 e**), and average head circumference (**Figure 4.19 f**). Cluster B exhibited a reduced IQ when compared to control cohorts (**Figure 4.19 g**).

Autism Diagnostic Interview-Revised (ADI-R) scores were used to assess major autism behavioral domains and revealed that Cluster B exhibited an increased severity in communication deficits (**Figure 4.18 c, Figure 4.19 h-j**), and was corroborated by the Vineland adaptive behavioral scale (**Figure 4.18 d**). We next assessed language development, a major dimension of communication behavior. Interestingly, Cluster A patients on average reached language milestones earlier than control patients, while Cluster B patients were further delayed than controls (**Figure 4.18 e-h**). These findings are corroborated by assessments of large CHD8 (Bernier et al., 2014) and DYRK1A (Earl et al., 2017) cohorts (Cluster B, Class 2), which demonstrated frequent language abnormalities.

Figure 4.18. Multiplex analysis reveals clinical subclasses of autism with distinct trajectories of language development.
(a) Unbiased hierarchical clustering of phenotypic data from six multiplex assays reveals two major functional groups of autism mutations from the MIX30 library. Class 1 mutations largely fall into cluster A, while Class 2 mutations fall into cluster B. **(b)** 2759 probands from the Simons Simplex Collection (SSC) were segregated into cohorts based on the presence of a *de novo* mutation in genes from Cluster A (17 patients) or Cluster B (55 patients). Control cohorts were defined as patients with any other *de novo* mutation not in the MIX30 library (*de novo* control, 883 patients) or patients without any known *de novo* mutation (idiopathic control, 1802 patients). **(c)** Cluster B patients exhibit increased severity in ADI-R non-verbal communication scores (ANOVA p = 0.0025; B vs. DN corrected p = 0.0022; B vs. Idiopathic corrected p = 0.0009). **(d)** Cluster B patients are on average reduced in the communication domain of the Vineland adaptive behavior scale (ANOVA p = 0.0034; B vs. DN corrected p = 0.0034; B vs. Idiopathic corrected p = 0.0013). **(e)** Average trajectories of language development. Control cohorts speak single words at ~24 months (DN=24.4 mo; idiopathic=24.4 mo), and speak their first phrases at ~39 months (other=38.8 mo; idiopathic=39.3 mo). Cluster B patients speak words at 28.02 months and first phrases at 46.1 months. Cluster A patients speak single words at 17.4 months and first phrases at 28.7 months. Typical language development is depicted in gray. **(f)** Cluster B cohort contains an increased fraction of patients with severe language deficit when compared to control cohorts (corrected p = 0.0404). **(g)** Cluster A exhibits a trend toward decreased fraction of patients with word delay compared to cluster B, though this did not survive multiple comparison testing (corrected p = 0.0895). **(h)** Cluster A contains a decreased fraction of patients with phrase delay than other cohorts (corrected p = 0.0144). **(i)** Class 0 exhibits an intermediate phenotype of average language development, between Class 1 and 2. **(j)** Correlation of language development with *in vitro* PFC neurogenesis. PFC neurogenesis values (**Figure 4.9**, DCX/SOX2 ratio) were assigned to probands using proband genotypes. Both first word (R^2 = 0.1334, p = 0.0024) and first phrases milestone (R^2 = 0.09551, p = 0.0182) showed significant positive correlations with the extent of PFC neurogenesis. Each dot represents one proband. **(k)** Summary of functional/molecular subtypes of autism defined using hPSC-based multiplex analysis platform. ASD, autism spectrum disorder. ADOS, Autism Diagnostic Observation Schedule. ADI-R, Autism Diagnostic Interview-Revised. Clustering: Jason Tchieu; Clinical analysis: Gustav Cederquist, Chelsea Rittenhouse, Sun-Young Chung, Lorenz Studer.


a
Unbiased clustering of autism genotypes
KATNAL2
KMT2C
ASH1L
DEAF1
RELN
ANKRD11
ASXL3
CUL3
CTNNB1
CHD2
KDM5B
KMT2A
GSK3B
DYRK1A
SUV420H1
PTEN
POGZ
CHD8
CTNND2
GRIN2B
ADNP
MED13L
CACNA1H
SCN2A
SYNGAP1
UMOD
ANK2
ARID1B
SETD5
TBR1
Cluster A
Cluster B
4
2
0
-2
-4
PFC IPC Production
PFC Growth
PFC Stem Cell Enrichment
PFC Neurogenesis
PFC WNT Response
Neural Crest Specification
Class Identity
Class 0
Class 1
Class 2
Control
b
Simons Simplex Collection: 2759 probands
A: Cluster A [Class 1] (n=17)
B: Cluster B [Class 0 + 2] (n=55)
DN: de novo Control (n=883)
Idiopathic Control (n=1802)
c
ADI-R Non-verbal Communication
Algorithm Score
**
A
B
DN
Idiopathic
d
VABS Communication
Algorithm Score
**
A
B
DN
Idiopathic
e
Trajectory of language development
Months
Cluster B
Controls
Cluster A
Typical
Single Word
First Phrase
f
Severe Language Deficit (ADI-R Q30)
Fraction of total
*
A
B
DN
Idiopathic
g
Word Delay (single word > 24 mo)
Fraction of Total
A
B
DN
Idiopathic
h
Phrase Delay (first phrase > 33 mo)
Fraction of Total
*
A
B
DN
Idiopathic
i
Language development by class
Months
Class 2
Class 0
Controls
Class 1
Single Word
First Phrase
j
PFC Neurogenesis vs First Word
Milestone reached (months)
R2 = 0.1334
p = .0024
PFC Neurogenesis (from Fig. 2b)
PFC Neurogenesis vs First Phrase
Milestone Reached (months)
R2 = 0.09551
p = 0.0182
Class 0
Class 1
Class 2
PFC Neurogenesis (from Fig 2b)
k
Summary of autism subtypes
Clinical Phenotype
- Communication
- Language
Class II
Class 0
Class I
Endophenotype
- PFC Neurogenesis
- WNT Response

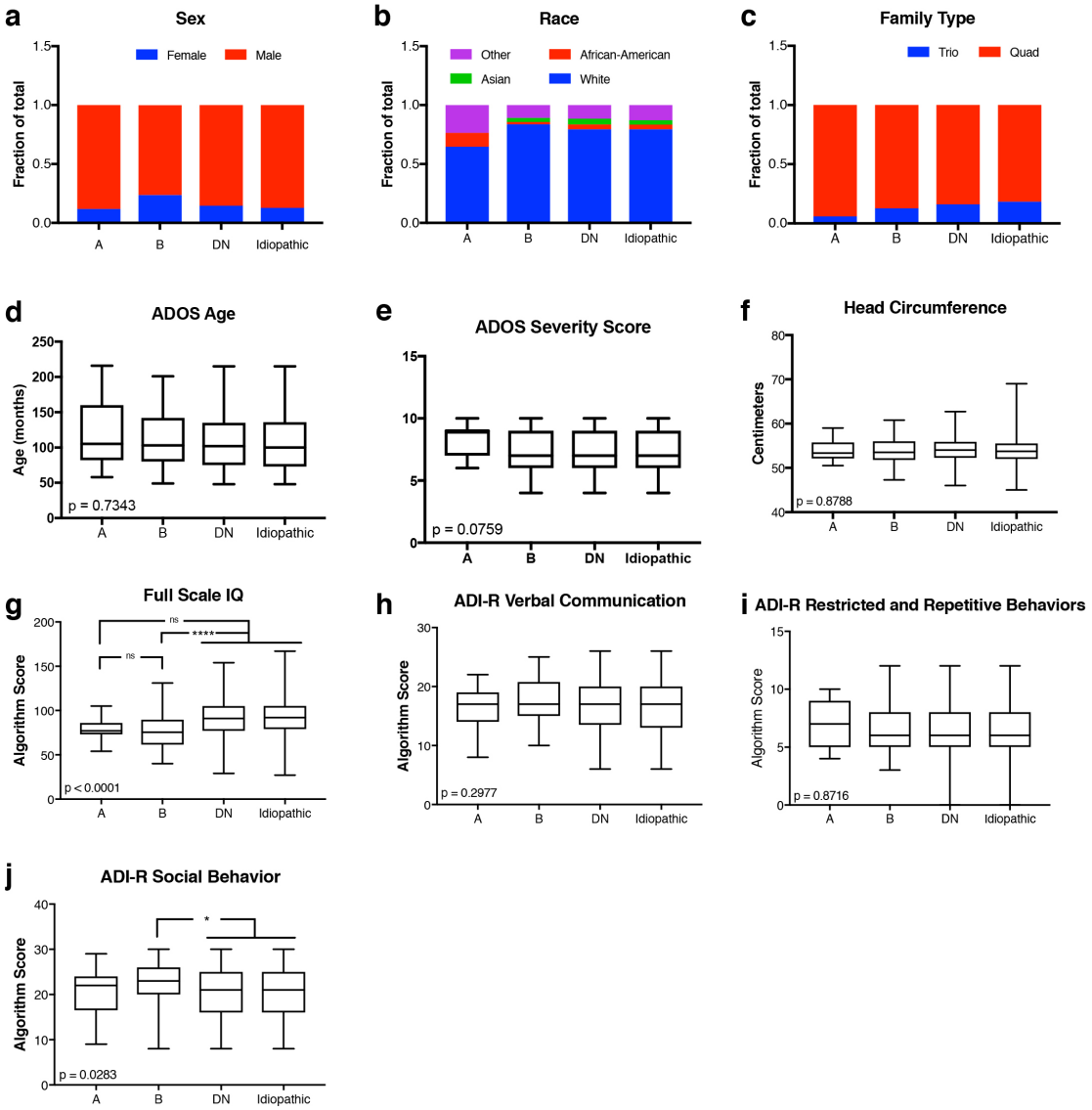


Figure 4.19. Additional Clinical Correlations.
(a-c) Demographic sex, race, and family data for five autism cohorts, plotted as fraction of total. **(d-f)** Autism cohorts were not significantly different in age (ANOVA p = 0.7763), autism severity score (ANOVA p = 0.1766), or head circumference (ANOVA p = 0.8788). **(g)** Cluster A and B were not different in IQ, but cluster B showed a reduction in full scale IQ compared to control cohorts (ANOVA p < 0.0001, B vs. DN: corrected p = 0.0002, B vs. Idiopathic: corrected p < 0.0001). **(h)** No significant between group differences on ADI-R verbal communication scores (ANOVA p = 0.3117). **(i)** No significant between group differences on ADI-R restricted and repetitive behavior scores (ANOVA p = 0.8817). **(j)** Cluster B exhibited increased severity on ADI-R social behavior measure (ANOVA p = 0.0283, B vs. DN corrected p= 0.0142, B vs. Idiopathic corrected p = 0.0172). Group differences were assessed with one-way ANOVA followed by tukey's multiple comparison test. Error bars are s.d. ADI-R, Autism Diagnostic Interview – Revised; ADOS, Autism Diagnostic Observation Schedule; DN, *de novo* control cohort. Data analysis: Gustav Cederquist, and Lorenz Studer.

When we further subdivided cluster B by Class (i.e. by PFC neurogenesis phenotype), we noticed that Class 0 patients tended to have an intermediate language phenotype between that of Class 1 and 2 (**Figure 4.18 i**), mirroring the pattern of PFC neurogenesis phenotypes. Correlating language data with PFC neurogenesis (**Figure 4.9 b**), using patient genotypes, revealed a positive association with the extent of neurogenesis and the severity of language acquisition phenotype (**Figure 4.18 j**). Surprisingly, delayed PFC neurogenesis may have a protective effect among autism patients.

In the future, it will be interesting to determine whether autism genotypes could be used to predict clinical phenotype and guide targeted early intervention. Moreover, exploring the molecular convergence within genotype classes as defined by our novel multiplex human PSC platform could lead to the development of precision therapeutics. At least 5/8 Class 1 genes are known regulators of polycomb activity (**Figure 4.20**), implicating changes in the epigenetic control of cortical neurogenesis. Polycomb regulates WNT-dependent responses to proneural gene transcription (Hirabayashi et al., 2009), which could explain Class 1 neurogenesis defects.

Finally, in addition to studying isogenic hPSCs, the multiplex platform could be adapted to patient-specific autism iPSCs to explore polygenic risk and to assess the impact of genetic background on complex disease. In summary, our platform captures the genetic heterogeneity of autism in a single experimental model, and reveals a well-defined functional architecture to early-developmental autism risk. Autism mutations converge onto discrete biological pathways, associated with PFC development and WNT signaling, that match distinct clinical outcomes in language acquisition (**Figure 4.18 k**). More broadly, our technology bridges a widening gap between the rapid accumulation of genetic information and the limited ability to

assess functional impact in classifying and potentially treating complex human disease.

Relationship to Polycomb

1. ANKRD11 - ?
2. ASH1L - Trithorax Group
3. ASXL3 - Polycomb Repressive Deubiquitinase
4. CUL3 - PcG Ubiquitination
5. DEAF1 - Homeotic Transformation
6. KDM5B - Occupites > 50% of Pcg Sites
7. KMT2C - COMPASS Complex
8. RELN - Suppressed by EZH2

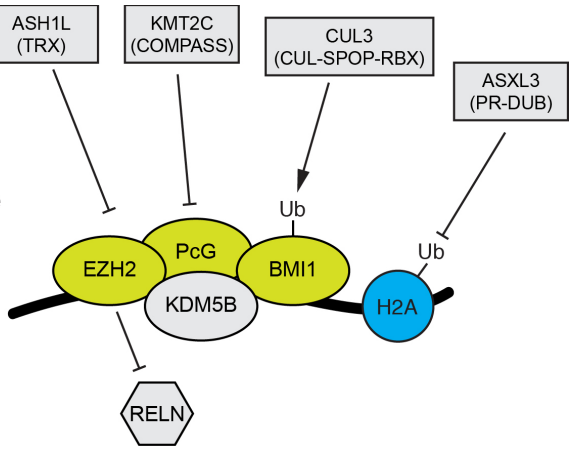


Figure 4.20. Class 1 genes regulate polycomb signaling.
Five of eight class 1 genes are known regulators of polycomb signaling. ASH1L is a trithorax group protein(Gregory et al., 2007), and KMT2C is a member of the COMPASS complex(Piunti and Shilatifard, 2016). ASXL3 is part of the polycomb repressive deubiquitinase complex(Srivastava et al., 2016). CUL3 regulates polycomb through ubiquitination(Hernandez-Munoz et al., 2005). KDM5B occupies over 50% of polycomb sites(Schmitz et al., 2011). In addition, DEAF1 mutant mice have a homeotic transformation phenotype(Hahm et al., 2004).

CHAPTER 5: CONCLUDING REMARKS ON THE PROSPECTS OF USING HPSCS FOR FUNCTIONAL CHARACTERIZATION AND CLASSIFICATION OF COMPLEX DISEASE

In this dissertation I developed novel hPSC-based methods to address current limitations in NDD modeling. Traditional hPSC-based monolayer cultures are an effective means to study pathophysiology at the level of individual cells. However, the study of multicellular developmental processes, such as the establishment of tissue architecture and long-range tissue patterning, remains challenging. I first explored, in collaboration, how hPSC-derived forebrain organoids could be used to model ZIKA-virus induced microcephaly and for identification of small molecule inhibitors of ZIKA virus infection (Chapter 2). These studies led me to conclude that current hPSC-based organoid technologies are useful for modeling disorders of brain growth, as they recapitulate the microarchitecture of progenitor zones in the forebrain, but fall short as models for other disease related processes such as tissue patterning. Therefore, I developed a novel organoid system by introducing a SHH signaling center into forebrain organoids. I demonstrated that the signaling center acts as a

developmental organizer, specifying the appropriate molecular and positional identity for major forebrain subdomains. The utility of this system to model disorders of developmental patterning was demonstrated through a series of experiments that found that cholesterol synthesis inhibiting drugs, including statins, perturb SHH-dependent tissue patterning and growth (Chapter 3).

The genetic diversity of disease risk is another major limitation in NDD modeling. For example, hundreds of distinct gene mutations have been associated with autism, yet attaining a systematic understanding of how these mutations affect brain development remains elusive. I developed an hPSC-based multiplex analysis platform for the simultaneous functional interrogation of at least 30 disease lines. I used this platform to characterize high-confidence, early-developmental, *de novo* autism-associated mutations and identified that mutations could be segregated into distinct functional categories based on PFC development and WNT activity phenotypes. Patients who fell into distinct subgroups exhibited average differences in clinical parameters of communication skills and language development. (Chapter 4).

From these studies, I propose two concepts that could generally be applied to the study of complex disease. The first concept is adequate genetic representation: the inclusion of a large number of disease-associated variants in experimental models of complex disease. Variants should preferably be selected in an unbiased manner such that the model reflects the underlying genetic diversity of the disease. In this way, experimentation can lead to broad insights regarding the molecular nature of disease. For example, one might find that the functional impact of multiple genetic variants converge on a single signaling pathway. Such insights would not be possible in single gene studies. Adequate genetic representation also allows controlled comparisons between disease variants, which could facilitate genotype-based subgrouping of patients to aid efforts in precision medicine.

The multiplex analysis performed in this dissertation was based on a rare variant model of complex disease, in which genetic risk is comprised by a large number of high-effect mutations that are individually rare in the population. In the future, it may be possible to use the multiplex platform to study common variant models of complex disease, in which genetic risk is comprised polygenic variants of low effect. This could be achieved by pooling iPSCs from patients who harbor common variants. Each line would be whole-genome (or exome) sequenced to define unique single nucleotide polymorphisms which act as genetic tags for ddPCR identification. The limitation of this approach is that cell lines are derived from non-isogenic backgrounds, which will make it difficult to discern causative relationships between genotype and phenotype. An alternate approach, which could potentially establish such relationships, is to genome engineer a library of cell lines that contain all permutations of a given set of common variants (e.g. a library containing three common variants A, B, and C would contain the following lines: A, B, C, AB, AC, BC, ABC).

A second general concept that could be applied to experimental modeling of complex disease is phenotypic depth: testing genetic variants against multiple phenotypic parameters. In the multiplex autism study, genetic variants were tested against six WNT and PFC related phenotypes. This strategy revealed consistent relationships between WNT signaling and PFC neurogenesis parameters, and provided a biological rationale for subgrouping of autism patients into cohorts that exhibited differences in clinical phenotype. Thus, phenotypic depth can help to define how genetic, molecular, developmental, and ultimately clinical phenomenology are linked, generating testable hypotheses for the future.

Given the link between PFC neurogenesis and language delay, it would be interesting to track language development in children who have early alterations in

PFC thickness. It would also be interesting to test additional *in vitro* phenotypes for correlations with other clinical parameters. For example, altered striatal neurogenesis may correlate with repetitive motor behaviors. In this way, it would be possible to build a catalogue of associations between endophenotype and clinical phenotype.

Currently, phenotypic depth is limited by the sophistication of hPSC-based experimental models. Organoid technologies are quickly improving our ability to model multicellular developmental processes. Topographically patterned SHH-organoids could be used to study dorso-ventral patterning, which is critical for the balanced generation of excitatory and inhibitory neurons and may underlie some cases of epilepsy. The protracted maturation time of hPSC-derived neurons into functional neural networks is also a limiting factor in the phenotypic characterization of disease models. Ultimately, unbiased phenotypic approaches may prove most fruitful for identification of primary disease pathways in common disease. For example, the combination of single-cell RNA sequencing with multiplex hPSC libraries could search for commonly altered and druggable molecular pathways.

Moving forward, I believe that hPSCs will be an invaluable tool for the study of complex disease, and specifically for initiatives in precision medicine. Precision medicine has the potential to revolutionize clinical care through individualized treatments based on a patient's unique disease profile. The classification of diseases into clinically-actionable subgroups is the first step toward this goal. For example, diabetes, a disease in which blood sugar remains abnormally elevated, is classified as Type I or Type II. Type I diabetes is defined as an inability to produce insulin, which normally shuttles glucose from the blood stream into organs. Type II diabetes is defined as an inability to respond to insulin. This classification scheme is effective because it reflects an underlying biology that directly informs an appropriate medical intervention (insulin replacement versus insulin sensitization).

The current classification of autism is based on a constellation of clinically observable phenomena. The Diagnostic and Statistical Manual 5 (DSM5) defines autism along two behavioral domains: impaired social interactions and restricted and/or repetitive behaviors. Patients are further described by the severity and quantity of symptoms (Grzadzinski et al., 2013). This dissertation illustrates that hPSCs can be an effective tool to systematically identify major disease pathways associated with NDDs. Thus, it may one day be possible to use this information to identify subgroups of autism patients based on molecular-genetic and developmental criteria. It is hoped that such endeavors will ultimately lead to the development of targeted therapies for patients with autism and other debilitating NDDs.

CHAPTER 6: MATERIALS AND METHODS

Chapter 2 Methods

hNPC culture and cortical neuron differentiation

hNPCs were derived and validated as previously described (Topol et al., 2015). Cells were maintained on growth factor reduced Matrigel (BD Biosciences) coated plates in NPC media, including Dulbecco's Modified Eagle Medium/Ham's F12 Nutrient Mixture (ThermoFisher Scientific), 1xN2 (ThermoFisher Scientific), 1xB27 minus RA (ThermoFisher Scientific) and 20 ng/ml recombinant human basic FGF (Peprotech), and split at 1:3 ratio every week with Accutase (Millipore). To induce cortical neuronal differentiation, NPCs were dissociated with Accutase and plated at 2.0 x 10^5 cells per cm^{-2} in NPC media on growth factor reduced Matrigel-coated plates. After overnight culture, media was changed to neural differentiation medium, including DMEM/F12, 1xN2, 1xB27 minus RA, 20 ng/ml BDNF

(Peprotech), 20 ng/ml GDNF (Peprotech), 20 ng/ml NT-3 (Peprotech), 1 mM dibutyryl-cyclic AMP (Sigma), 200 nM ascorbic acid (Sigma) and 1 µg/ml laminin (ThermoFisher Scientific). NPC-derived neurons were differentiated for 1~2 months.

Generation and culture of human forebrain organoids

To induce forebrain organoid differentiation, H9 hESC line (WA-09, WiCell) hPSCs were dissociated to single cells using EDTA, and 9000 cells were reaggregated using low-adhesion V-bottom 96 well plates (Wako) in Essential8 Medium (Fisher Scientific) with 10 µM Y-27632 (Tocris Biosciences). After 24 h (day 0), medium was changed to Essential6 (Fischer Scientific) supplemented with 10 µM SB431542 (Tocris Biosciences), 500 nM LDN193189 (Stem Cell Technologies), and 2 µM XAV939 (Tocris Biosciences) until day 4. From day 4 to day 18, XAV939 was removed. Medium was changed every other day. From Day 18, organoids were maintained in organoid differentiation medium (50% DMEM F-12 (Fisher Scientific), 50% mL Neurobasal (Fisher Scientific), 0.5xN2 supplement (Stem Cell Technologies), 0.025% insulin (Sigma), 5 mM L-Glutamine (1x, Fischer Scientific), 0.7 mM MEM-NEAA (1x, Fischer Scientific), 50 U/mL Penicillin-Streptomycin (1x, Fischer Scientific), 55 μM 2-mercaptoethanol (1x, Fischer Scientific), 1xB27 supplement without Vitamin A (Fischer Scientific).

ZIKV infection of hNPCs and forebrain organoids

hNPCs were plated on 96-well plates at $8.5x10^4/cm^2$. After overnight incubation, hNPCs were infected with ZIKV (MR766 strain, MOI=0.125) for 2 h, and changed to virus-free medium. The cells were maintained in NPC medium with daily medium change. Three days later, the supernatant and cells were collected and used for

assays. hESC-derived forebrain organoids were infected with ZIKV MR766 (5 x 10^5 PFU/ml) for 24 h. After removal of virus containing medium, forebrain organoids were maintained in forebrain organoid medium for an additional 3 or 17 days and then used for assays. To monitor ZIKV dynamics in hNPCs, hNPCs were infected with ZIKV (MR766 strain, MOI=0.125, FSS13025 strain MOI=0.1) for 2 h, and changed to virus-free medium. The cells were maintained in NPC medium with daily medium change. The cells were lysed at the different time points and analyzed using qRT-PCR.

High-content Screen

hNPCs were dissociated and plated on Matrigel-coated 384-well plates for the chemical screen. After overnight incubation, the cells were treated with compounds from an FDA-approved drug library for 1 h with final concentration at 10 μM. DMSO was used as a negative control. ZIKV (MR766 strain) was added at MOI=0.125. After 2 h infection, the supernatant was replaced with fresh medium and compounds from the same chemical library were added to the same wells. After an additional 3 days incubation, cells were fixed, stained with antibodies against ZIKV envelope protein (ZIKV E) and proliferation marker Ki67, and analyzed with the ImageXpress Micro Widefield High-Content Analysis System. A two dimensional analysis was performed to pick the primary hit compounds (Figure S1A). X-axis represents the fold change of total cell number, which was calculated by dividing the total cell number of the chemical treated well by the average of total cell number of DMSO treated wells. Y-axis represents the fold change of the percentage of ZIKV infection, which was calculated by dividing the percentage of ZIKV infection of the chemical treated well by the average of the percentage of ZIKV infection of DMSO

treated wells. The compounds in which the fold change of total cell number >1 and the fold change of the percentage of ZIKV infection <20% were further evaluated.

Generation of inhibitory curves

hNPCs were plated on 96-well plates with Accutase digestion at 8.5x104 /cm2 density and incubated overnight. For generating inhibitory curves of compound for inhibiting ZIKV infection, cells were first treated with DMSO or different doses of hit compounds for 1 h before ZIKV infection. The cells were then infected with ZIKV (MR766 strain) at a MOI=0.125. After 2 h of infection, the virus-containing medium was replaced by fresh virus-free medium and compounds were added again to the same wells. For generating inhibitory curves of compound for eliminating ZIKV, compounds were instead added after ZIKV infection. After 3 days incubation, the culture medium was harvested and applied to the Vero re-infection assay. Vero cells were plated to 96-well plates one night before re-infection. ZIKV containing medium was serial diluted 10, 10^2, 10^3, 10^4, 10^5 and10^6 fold for Vero infection.

RNA-seq analysis

hNPCs at 72 h post-inoculation were collected for RNA-seq analysis, including mock-infected hNPCs treated with DMSO, mock-infected hNPCs treated with 25 μM HH, mock-infected hNPCs treated with 15 μM AQ, ZIKV (MR766 strain, MOI=0.125) infected hNPCs treated with DMSO, ZIKV-infected hNPCs treated with 25 μM HH, and ZIKV-infected hNPCs treated with 15 μM AQ. The RNA from hNPCs in each condition was extracted with Absolutely RNA Nanoprep kit (Agilent Technologies, 400753). The RNA quality was validated with a bioanalyzer (Agilent). The cDNA libraries were synthesized using the TruSeq RNA Sample Preparation kit (Illumina) and sequenced in single-read with the HiSeq4000 sequencer (Illumina) at Weill

Cornell Genomics Resources Core Facility. The reads were aligned to the human hg19 reference genome with Tophat2 (Kim et al., 2013). Gene expression data were analyzed with Cufflinks (Trapnell et al., 2010). To generate heat maps displaying the differential gene expression patterns of different samples, the expression values (RPKM) were normalized per gene over all samples, to be specific, for each gene we calculated the mean and standard deviation (stdev) of expression over all samples, and linearly transformed the expression value using the formula (RPKM-mean)/stdev. The heat maps were then generated by Heatmap.2 in the R plots package. Gene ontology pathway analysis was performed using the DAVID function annotation tool.

Immunofluorescence Staining

Cells were fixed with 4% paraformaldehyde for 20 min at room temperature (RT) and blocked in a solution of Mg^{2+} and Ca^{2+} free PBS containing 5% horse serum and 0.3% Triton-X for 1 h at room temperature and followed by incubation with primary antibody at 4°C overnight. The following primary antibodies and dilutions have been used in this study: mouse anti-Flavivirus group antigen antibody (ZIKV E) (1:2000; Millipore, clone D1-4G2-4-15), mouse anti-Ki67 (1:200, DAKO, MIB-1), rabbit anti-Ki67 (1:500, ThermoFisher, SP-6), rabbit anti-cleaved caspase-3 (1:1000, Cell Signaling Technology, Asp15), goat anti-SOX2 (1:100, Santa Cruz, sc-17320), rabbit anti-SOX2 (1:200, Biolegend, N-term), mouse anti-NESTIN (1:1000, Neuromics, MO15012), rabbit anti-TUJ1 antibody (1:500, Covance, MRB-435P), and chicken anti-MAP2 (1:1000, Abcam, Ab5392). The secondary antibodies include donkey anti-mouse, goat, rabbit or chicken secondary antibodies conjugated with Alexa-Fluor-488, Alexa-Fluor-594 or Alexa-Fluor-647 fluorophore (1:500, Life Technologies). Nuclei were counterstained by DAPI.

The immunohistochemical analysis of mouse tissues was performed on cryosections. In brief, the tissues were fixed overnight with 4% PFA at 4°C, sunk in 20% sucrose, embedded in OCT, and then sectioned at 30 μm for immunostaining. The primary and secondary antibodies were used as described above.

RNA Extraction and qRT-PCR

Total RNA of cultured cells and organs was extracted with RNeasy plus mini kit (Qiagen). vRNA from cell culture supernatant was extracted with QIAampe viral RNA mini kit (Qiagen). Reverse transcription was carried out with strand specific tagged primers of ZIKV plus ACTB gene specific primer of human or mouse using High Capacity cDNA Reverse Transcription kit (Thermo Fisher).

qPCR reactions were performed with PrimeTime Gene Expression 2X Master Mix (IDT DNA) and probes for ZIKV and human or mouse ACTB. The PCR products were analyzed with Sanger sequencing to validate the specificity of "+" and "-"strand primers. ZIKV expression level was then normalized to human ACTB or mouse Actb correspondingly. Primer sets used in RT and qPCR and probes used in qPCR were listed in Table S2.

Adult mouse infection and *in vivo* drug testing

6-8 week old female SCID-Beige mice were infected with ZIKV (2.5 x10^5 PFU in 0.5 ml culture fluid) through intraperitoneal injection. For pre-infection treatment, drug candidates were administrated to mice 12 h before ZIKV inoculation and followed by treatment once per day. HH was administrated subcutaneously at the dose of 100 mg/kg body weight. AQ was administrated with intraperitoneal injection at the dose of 40 mg/kg body weight. PBS was used as the vehicle control. For ZIKV kinetics analysis, 6-8 weeks old female SCID-beige mice were inoculated with (1 x10^6 PFU for

MR766 strain or 1 x10^7 PFU for FSS13025 strain in 0.5 ml culture fluid) through intraperitoneal injection. The mice were euthanized at the indicated time points and analyzed with qRT-PCR to validate the ZIKV level *in vivo*. At 24 hpi or 5 dpi, HH was administrated subcutaneously at the dose of 100 or 200 mg/kg body weight per day for 5 days (MR766 strain). At 7 dpi, HH was administrated subcutaneously at the dose of 100 mg/kg body weight per day for 7 days (FSS13025 strain). After euthanasia, brain, liver, spleen and kidney were collected and analyzed using qRT-PCR. Blood samples from mice treated with HH with dose of 200 mg/kg body weight after 5 days treatment were analyzed with Vero infection assay.

Vero assay to monitor the infectious particular *in vivo*

Whole blood samples were collected from tail vein and serum was isolated with centrifuge. Solid samples (brain, liver, spleen and kidney) were harvested from infected mice after euthanasia. Samples were freshly homogenized with cold culture medium (DMEM + 10% fetal bovine serum) and then centrifuged (2000g for 10 minutes at 4 °C). Supernatant and serum were then diluted serially to infect Vero cells in 96-well plates.

Quantification and Statistical Analysis

Quantification data are presented as mean ± SEM. N=3 independent biological replicates were used if not otherwise specifically indicated. n.s. indicates non-significant difference. For *in vitro* experiments, *p* values were calculated by unpaired two-tailed Student's t-test if not otherwise specifically indicated. For *in vivo* experiments, *p* values were calculated by one-way repeated measures ANOVA or two-way repeated measures ANOVA with a Bonferroni test for multiple comparisons using Prism 6. *$p<0.05$, **$p<0.01$ and ***$p<0.001$.

Data and Software Availability

The accession number for the RNA-seq data reported in this paper is GEO: GSE89334.

Chapter 3 Methods

hPSC maintenance and monolayer differentiation

WA09 and derivatives (EF1α::GFP (GFP), EF1 α::RFP; TRE-SHH (iSHH)) hPSC were maintained with Essential 8 medium or Essential 8 flex (E8, Thermo, A15117001 or A28558501) in feeder-free conditions on vitronectin (VTN-N) substrate (Thermo, A14700). hPSCs were passaged as clumps with EDTA solution (0.5 M EDTA/PBS). For neural monolayer differentiation, hPSCs were dissociated to single cells and plated on matrigel substrate (BD Biosciences, 354234) in E8 at a density of 250,000 cells/cm^2 in the presence of ROCK inhibitor (Y-27632, 10 μM, Tocris 1254) (Day -1). From Day 0 to 7, cells were cultured in Essential 6 medium (E6, Thermo, A1516401) in the presence of TGFβ and BMP inhibitors (LDN193189, 100 nM, Stem Cell Technologies, 72142; SB431542, 10 μM, Tocris, 1614). From day 7 to 12, cells were cultured in E6 alone. Floorplate was specified by addition of SHH (200 ng/ul, R&D Systems 464-SH) and CHIR99021 (0.7 μM, Tocris 4423). Media was changed every day during the differentiation.

SHH inducible line

The iSHH hPSC was generated according to a previously described gene targeting strategy(Gonzalez et al., 2014). Briefly, two donor constructs were targeted to the first intron of the AAVS1 locus using TALEN-facilitated homologous recombination. One

construct contained a constitutively expressed reverse tetracycline transactivator (M2rtTA) and the second construct contained a tetracycline inducible SHH cassette. A constitutive RFP expressing line (WA09 derived, EF1α::RFP) was used, so iSHH cells could be visualized(Fattahi et al., 2016).

SHH-spheroid aggregation

1,000 iSHH cells were dissociated to single cell and aggregated in low-attachment round bottom microwells for 24 hours in the presence of ROCK inhibitor (Y-27632, 10 μM) and WNT inhibitor (XAV939, 5μM). For iSHH-GFP-spheroids, iSHH cells were plated in the morning and allowed to aggregate for 8 hours. 1,000 GFP cells were plated on top and allowed to aggregate overnight. The next day, 10,000 wildtype WAH9 hPSCs were dissociated to single cell and allowed to aggregate on top of the iSHH cells, in E8 medium with ROCK inhibitor (Y-27632, 10 μM) and WNT inhibitor (XAV939, 5μM, Tocris 3748).

SHH organoid differentiation

SHH-spheroids were cultured in E6 with 3-inhibitors (LDN193189, 100 nM; SB431542, 10 μM; XAV939, 5 μM) plus doxycycline (400 ng/ml) until spheroid tissue began to brighten and have smooth edges, typically 6-8 days. SHH-spheroids were then embedded in matrigel drops and cultured in an N2/B27 based organoid media as previously described(Lancaster and Knoblich, 2014b). SHH-spheroids were moved an orbital shaker once formation of translucent, neuroepithelial buds was observed, typically 4 - 7 days after embedding. Organoids that failed to form translucent, neuroepithelial, buds were discarded. Half media was changed every other day.

Histology and immunocytochemistry

Organoids were fixed in 4% PFA overnight, and washed three times with PBS the next day. Day 20 organoid tissue was cryoprotected in 30% sucrose/PBS and sectioned at 30um on a cryostat (Leica 3050 S). Floating Sections were blocked for 30 minutes in 10%FBS, 1%BSA, 0.3%triton PBS, and incubated with primary antibody overnight. Day 6 organoid tissue was cryoprotected in 30% sucrose/PBS and sectioned at 20um directly onto slides. Slides were blocked for 30 minutes in 10%FBS, 1%BSA, 0.3%triton PBS, and incubated with primary antibody overnight. Sections or slides were washed with PBS then incubated with secondary antibody for 3 hours at room temperature.

Quantification of organoid patterning and SHH activity

Microscopy was performed using a standard inverted epifluorescence microscope (Olympus IX71 or Zeiss Axio Observer). Images were acquired using Cell Sens (Olympus) or Zen Pro (Zeiss) software. Min, max and gamma (midtone) adjustments were applied uniformly to images during processing with Adobe Photoshop Creative Cloud.

The position of regional domains in SHH-organoids was quantified by first dividing each image into a grid of 20μm x 20μm squares (4x images) or 5μm x 5μm square (20x images) that form a Cartesian coordinate grid. Each 20μm x 20μm or 5μm x 5μm square is a discrete region of interest (ROI) that is associated with an X and Y coordinate. The origin of the grid was calculated for each image as the "center of mass" of the iSHH-organizer cells: $CM_x = \Sigma m_i x_i / M$ and $CM_y = \Sigma m_i y_i / M$, where m = grey value intensity of iSHH cells for individual ROIs, x or y = coordinate, and M = total intensity for all ROIs, and the sum includes all x or y coordinates in the image. Next, locations of gene expression (e.g. PAX6) were defined by thresholding staining

intensity and determining ROIs with positive immunoreactivity. The linear distance from the origin to each ROI ($sqrt((ROI_x-CM_x)^2 + (ROI_y-CM_y)^2)$) was calculated. In Figures 3-5b and 3-13, ROI frequencies are plotted as a function of distance from the center of mass. Typically 2-5 sections at 180μm intervals were used for quantification of each biological replicate (N).

In Figure 3-10c, the GFP shell distance is defined as the radial distance from the center of mass at which the ratio of GFP ROIs / DAPI ROIs drops below 35%. The SHH-activity distance is the radial distance from the center of mass to the edge of the PAX6-negative territory, defined as the distance at which the ratio of PAX6-negative ROIs / DAPI ROIs dropped below 95% (PAX6-negative ROIs were obtained by subtracting PAX6 thresholded pixels from DAPI thresholded pixels). The 35% and 95% parameters were determined prior to analysis as accurately cut off values by comparing with manual distance measurements. Sections that lacked organizer cells or that had large necrotic centers were excluded from analysis.

The distance of SHH patterning activity in figure 3-10a was obtained by measuring the distance of lines from the edge of PAX6 territory to the nearest organizer cells. At least 5 lines were measured per section. All quantifications were performed using ImageJ.

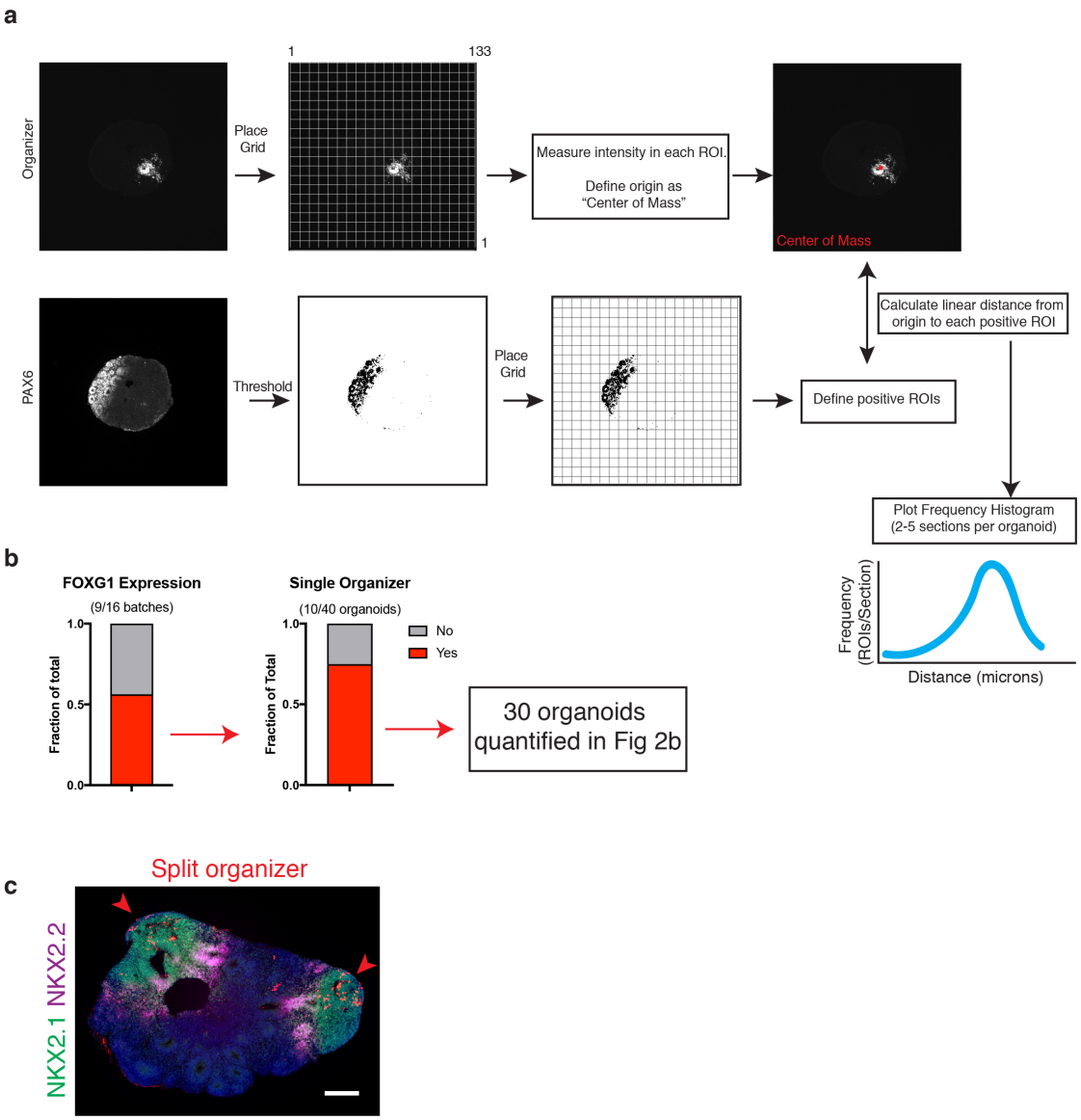


Figure 6.1. Method for quantifying SHH-organoid topography.
(a) SHH-organoids are quantified using a grid of regions of interest (ROIs), and each ROI is associated with an X and Y coordinate. The origin is defined for each section by calculating the "center of mass" of the organizer signal. The grid is then used to define ROIs that are positive for regional markers (e.g. PAX6). The linear distance from all positive ROIs to the origin is calculated, and these data can be plotted as a frequency histogram. (b) Organoids that did not express FOXG1, or that had a split organizer (c) were not included in the quantification. Scale bars: 200μm.

RNA extraction and RT-PCR

RNA was extracted using trizol reagent (Invitrogen, 15596026) followed by chloroform extraction. RNA was precipitated in isopropanol and resuspended in ddH2O. cDNA synthesis was performed using 1 μg of RNA (iScript, Bio-Rad, 1708840). RT-PCR was performed with EvaGreen Supermix (Bio-Rad, 1725202). Three to four individual samples were combined into a single replicate when collecting RNA from spheroids.

Visualization of SHH protein gradient

SHH-spheroids were fixed in 4% PFA for 6 hours at 4C, then washed three times in PBS. Standard immunocytochemistry was performed on wholemount SHH-spheroids using anti-SHH primary antibody (1:100, clone 5E1, DSHB). Secondary antibody was amplified using a tyramide signal amplification kit (Thermo, B40941). Flourescent stained spheroids were visualized using an inverted epifluorescence microscope. Gradients were quantified using line scans that transected the organizer tissue. Quantifications were performed using ImageJ.

Cell-based SHH Palmitoylation Assay

Synthesis of [I^{125}]iodopalmitate was carried out as previously described (Alland et al., 1994; Berthiaume et al., 1995; Buglino and Resh, 2008). iSHH hPSC were dissociated to single cell and seeded as a confluent monolayer on matrigel substrate at a density of 250,000 cells / cm^2 and differentiated for 9 days in E6 with LDN193189, (100 nM) and SB431542 (10 μM) at increasing concentrations of doxycycline (0, 125, 250, 500, 1000, 2000 ng/ml). On day 9, iSHH cells were incubated for 1 h at 37 °C with DMEM containing 2% dialyzed FBS and then labeled for 4 h with 10μCi of [I^{125}]iodopalmitate. The cells were lysed and subjected to immunoprecipitation with anti-Shh antibody, and the immunoprecipitates were diluted

in 100 µl of 2× SDS-PAGE sample buffer containing 10 mM DTT. The samples were analyzed on a 12.5% SDS-PAGE gel, followed by phosphorimaging on a Typhoon FLA-7000 phosphorimager. An aliquot of each immunoprecipitate was analyzed for Shh protein expression by Western blotting.

Statistical Analysis

All reported measurements are from distinct samples. When comparing samples across treatment groups, samples were collected, processed, and analyzed in parallel. Organoid samples are taken from two independent differentiations, except where noted in the figure legend. Statistical tests are listed in the figure legends.

Chapter 4 Methods

Gene selection for autism library

Gene selection for the MIX30 library was performed in the Spring of 2015 using the SFARI gene database. First, all genes with a score of 1 or 2 (high-confidence) were selected. Second, genes were filtered for early developmental expression using the BrainSpan human fetal brain transcriptional atlas (BrainSpan.org, expressed at PCW8) and a hPSC-derived cortical neuron transcriptional atlas (Cortecon.neuralsci.org, expressed on or before day 50).

Generation of multiplex library

CRISPR/Cas9 was used to introduce frameshift mutations into high-confidence autism genes. Guide RNAs (gRNAs) were designed to target exons in which indels or single nucleotide variant (SNV) mutations have been found in patients. If no suitable target sequence was found, then an upstream site was chosen. gRNAs were cloned

into the bicistronic PX458 Cas9-GFP vector (Addgene 48138), and introduced into MEL1 hPSCs (46XY) by nucleofection (Lonza). Nucleofected cells were FACS sorted for GFP, and individual clones were collected on a mouse embryonic fibroblast (MEF, Global Stem) feeder layer in the presence of Rock-inhibitor (Y-27632, 10 mM, Tocris 1254) in knockout serum replacement (KSR; Life Technologies,10828-028) as previously described(Fattahi et al., 2016) for two weeks. Rock-inhibitor was removed after 4 days. Clones were picked onto a vitronectin substrate and further maintained in Essential 8 media (Life Technologies). Clones were screened using sanger sequencing. Compound hets were discarded, and only homozygous or heterozygous clones were kept. Heterozygous clones were inferred bioinformatically (http://yosttools.genetics.utah.edu/PolyPeakParser/). All frozen stocks were sequence validated. Since patient mutations could be gain-of-function or loss-of-function, DNA sequencing rather than protein expression was used for validation.

hPSC maintenance, pooling, and storage.

MEL1 and derivatives were maintained with Essential 8 medium or Essential 8 flex (E8, Thermo, A15117001 or A28558501) in feeder-free conditions on vitronectin (VTN-N) substrate (Thermo, A14700). hPSCs were passaged as clumps with EDTA solution (0.5 mM EDTA/PBS). Pooling was performed by dissociating lines to single cell with EDTA and adding cells at desired frequency. Pools were established in the presence ROCK inhibitor (Y-27632, 10 μM, Tocris 1254) for 1 day. Pooled hPSCs were frozen in E8 with 10% DMSO (Sigma) media and thawed in the presence of ROCK inhibitor (10 μM)

Prefrontal cortex differentiation

hPSCs were dissociated to single cells and plated on matrigel substrate (BD Biosciences, 354234) in E8 at a density of 250,000 cells/cm^2 in the presence of ROCK inhibitor (Y-27632, 10 mM, Tocris 1254) (Day -1). From Day 0 to 6-8, cells were cultured in Essential 6 medium (E6, Thermo, A1516401) in the presence of TGFβ and BMP inhibitors (LDN193189, 100 nM, Stem Cell Technologies, 72142; SB431542, 10 mM, Tocris, 1614). WNT inhibitor (XAV939, 2uM) was also included from D0-2. On day 6-8, monolayer cultures were dissociated with accutase and replated as high-density droplets on laminin/fibronectin, and cultured in N2 media (B27 1:1000, FGF8 50ng/ul, SHH 25 ng/ul) for 4 days, until neuroepithelial rosettes were visible. Droplets were then passaged 1:2 with trypsin onto laminin/fibronectin coated plates and cultured in the same media. At day 20, cultures were passaged using accutase or dispase to a density of 200,000 cells/cm^2 to 400,000 cells/cm^2 and cultured in N2 media (B27 1:50, FGF8 50ng/ul) for up to 20 days. Cells were cultured in N2/B27 (1:50) media after day 40. Cultures in which flat morphology cells arose were discarded.

Neural Crest differentiation

hPSCs were dissociated to single cells and plated on matrigel substrate (BD Biosciences, 354234) in E8 at a density of 200,000 cells/cm^2 in the presence of ROCK inhibitor (Y-27632, 10 mM, Tocris 1254) (Day -1). From day 0-2, cells were cultured in E6 with BMP4 (1ng/mL), CHIR99021 (0.6 μM), and SB431542 (10μM). From day 3-10, cells were cultured in CHIR99021 (1.5 μM) and SB431542 (10μM). Cells were dissociated with accutase for FACS.

Measurement of allele frequencies using droplet digital PCR

We measured allele frequencies in pooled mixtures using ddPCR. ddPCR assays were designed by the manufacturer (BioRad, see reagents table), and incorporated a 5' fluorescently labeled probe (HEX or FAM) with a 3' quencher (Zen) for wild-type and mutant allele respectively. ddPCR was performed according to the manufacturer protocol. All DNA was measured with a Qubit (Thermo) prior to ddPCR. Wild-type and mutant allele concentrations for each genotype were estimated by a Poisson distribution of the fluorescence reads and converted to a fractional abundance estimate (QuantaSoft, BioRad), and used to calculate mutant allele frequencies. Growth and cell-state phenotypes were determined by calculating changes in relative allele frequency across phenotypic fractions and normalizing each line to the internal negative standard, UMOD, for each replicate. WNT response phenotypes from day 45 PFC cultures were determined by comparing changes in relative allele frequency between treated and untreated conditions.

Flow cytometry and genomic DNA extraction from fixed cells.

Cultures were dissociated with accutase and fixed and permeabilized with BD Cytofix/Cytoperm (BD Bioscience, 554722) for 45 minutes on ice. Fixed cells were washed with BD Perm/Wash Buffer (BD Bioscience, 554723). Cells were stained with primary antibody for 1 hour on ice and secondary antibody for 30 minutes on ice, and sorted using a FACSAria III flow cytometer (BD Bioscience), and FlowJo Software (BD) for analysis. Sorted fixed cells were centrifuged for 5 minutes at 20,000 rcf. Pellets were resuspended in 500 μl lysis buffer (10mM Tris-HCL pH 8.0, 100 mM NaCl, 10mM EDTA, 0.5% SDS, 40mg/mL proteinase K) and incubated at 65C, shaking, overnight. The next day, 300 ul NaCl was added to lysis and incubated on ice for 10 minutes. Samples were centrifuged at 20,000 rcf for 10 minutes and

aqueous phase DNA was precipitated in 650 μl of isopropanol, washed with 70% ethanol, and resuspended in ddH2O.

RNA extraction and qRT-PCR

RNA was extracted using trizol reagent (Invitrogen, 15596026) followed by chloroform extraction. RNA was precipitated in isopropanol and resuspended in ddH2O. cDNA synthesis was performed using 1 μg of RNA (iScript, Bio-Rad, 1708840). RT-PCR was performed with EvaGreen Supermix (Bio-Rad, 1725202) and analyzed on a CFx96 Real-Time System (BioRad)

Immunocytochemistry

Cells were fixed in 4% PFA for 15 minutes, and washed three times with PBS. Cells were blocked for 30 minutes in 10%FBS, 1%BSA, 0.3%triton PBS, and incubated with primary antibody overnight. The next day, sections were washed with PBS then incubated with secondary antibody for 1 hour at room temperature. Microscopy was performed using a standard inverted epifluorescence microscope (Olympus IX71). Images were acquired using Cell Sens (Olympus). Min, max and gamma (midtone) adjustments were applied uniformly to images during processing with Adobe Photoshop Creative Cloud.

Zebrafish husbandry

Zebrafish work was approved by the Institutional Animal Care and Use Committee (IACUC) at MSKCC. Zebrafish were bred and maintained in the Zuckerman fish facility, in temperature (28°C), pH (7.4), and salinity-controlled conditions. All fish were maintained on a 14hr on/10hr off light cycle. Zebrafish used were of the ab strain.

Creation of zebrafish CRISPR F0 mosaic mutants

We designed targeting sgRNAs for the genes of interest for homologous exons that were targeted in hPSC lines, in two zebrafish paralogues if applicable. CHOPCHOP (cite: http://chopchop.cbu.uib.no/). gRNA/Cas9/Tracer complexes were then synthesized using the ALT-R system and prepared according to previously published protocols (https://www.idtdna.com/pages/products/crispr-genome-editing/alt-r-crispr-cas9-system) CRISPR activity was confirmed from a random subset of injected embryos using a surveyor assay (IDT), for at least 1 paralogue of each gene for conditions that showed a significant jaw phenotype.

Zebrafish imaging and image processing

Fish were imaged at 7 dpf using an upright Zeiss Discovery V16 equipped with a motorized stage, brightfield, GFP and tdTomato filter sets. To acquire images, fish were lightly anaesthetized with Tricaine 4 mg ml-1 and placed into agarose molds to properly image the head from a ventral vantage point. Images were acquired with the Zeiss Zen software v1, and the post image processing was done using ImageJ. Zebrafish images were quantified by a blinded observer using ImageJ software. Jaw length was measured as an angle between one line from the top of the eyes and a second line from the top of the right eye to middle of the jaw, depicted in **Fig. S10d**. Jaw angle phenotypes were scored on a binary scale, with a cut-off of < 1 s.d. below WT average.

Functional Clustering of Multiplex Data and Clinical Analysis

Functional classes of autism mutations (**Figure 4.10**) were defined based on a positive PFC neurogenesis phenotype. A positive PFC neurogenesis phenotype was

assigned if a genotype exhibited altered PFC neuronal production or altered PFC stem cell enrichment with an FDR < 0.05, or if a genotype exhibited both altered PFC neuronal production and altered PFC stem cell enrichment with an FDR <0.1. Aggregate multiplex data across PFC and WNT-related assays (PFC growth, PFC IPC production, PFC neurogenesis, PFC stem cell enrichment, PFC WNT response, Neural crest induction) were aggregated and clustered in R using the pheatmap library **(Figure 4.18 a)**.

Proband data was ascertained from the Simons Simplex Collection Clinical Database (SFARIBase). Genotypes were assigned using previously published results from sequencing studies(Iossifov et al., 2014; Krumm et al., 2015; Sanders et al., 2015). Patients in Cluster A and Cluster B were assigned genotypes based on the presence of *de novo* coding or splice-site variants. Non-splice site intronic and inherited mutations were not considered. Patients with *de novo* loss-of-function or MIS3 missense mutations that did not fit into Cluster A or Cluster B were included in the *de novo* Control. All other patients were included in the idiopathic control. The ADI-R verbal communication score excludes patients who severe language deficits, and thus the ADI-R non-verbal communication score was used in order to compare all patients regardless of language ability (**Figure 4.18 c**). Correlations in Figure 4.18 j were performed by assigning a PFC neurogenesis value to each patient based on their genotype. The PFC neurogenesis value is the neuronal production value (DCX/SOX2) for each genotype from Figure 4.10 b.

Statistical Methods

All reported measurements are from distinct samples. At least three independent biological replicates were used for each experiment, derived from at least two independent MIX30 pools for multiplex experiments. Specific data on replicates (n) is

given in the figure legends. Data are presented as mean ± s.e.m., except where noted in the figure legends. False discovery rates (FDR) for multiplex assays were calculated using two-sided t-test to compare the means between autism lines and the control UMOD, and correcting p values for multiple comparisons using the Benjamini-Hochberg method. Comparisons of clinical cohorts were performed using ANOVA with Tukey test (for quantitative parameters). Exact p-values for Fisher's test were corrected with Benjamini-Hochberg method. Fisher's tests were two-sided. Statistical analysis was performed using Prism 7 (Graphpad) or Excel (Microsoft) software. Mean and corrected p values from multiplex assays are included in **Table S2**.

REFERENCES

Ahn, J., Jang, J., Choi, J., Lee, J., Oh, S.H., Lee, J., Yoon, K., and Kim, S. (2014). GSK3beta, but not GSK3alpha, inhibits the neuronal differentiation of neural progenitor cells as a downstream target of mammalian target of rapamycin complex1. Stem Cells Dev *23*, 1121-1133.

Alcamo, E.A., Chirivella, L., Dautzenberg, M., Dobreva, G., Farinas, I., Grosschedl, R., and McConnell, S.K. (2008). Satb2 regulates callosal projection neuron identity in the developing cerebral cortex. Neuron *57*, 364-377.

Alland, L., Peseckis, S.M., Atherton, R.E., Berthiaume, L., and Resh, M.D. (1994). Dual myristylation and palmitylation of Src family member p59fyn affects subcellular localization. J Biol Chem *269*, 16701-16705.

Alsio, J.M., Tarchini, B., Cayouette, M., and Livesey, F.J. (2013). Ikaros promotes early-born neuronal fates in the cerebral cortex. Proc Natl Acad Sci U S A *110*, E716-725.

Anderson, C., and Stern, C.D. (2016). Organizers in Development. Curr Top Dev Biol *117*, 435-454.

Anderson, S.A., Eisenstat, D.D., Shi, L., and Rubenstein, J.L. (1997). Interneuron migration from basal forebrain to neocortex: dependence on Dlx genes. Science *278*, 474-476.

Anderson, S.A., Kaznowski, C.E., Horn, C., Rubenstein, J.L., and McConnell, S.K. (2002). Distinct origins of neocortical projection neurons and interneurons in vivo. Cereb Cortex *12*, 702-709.

Angevine, J.B., Jr., and Sidman, R.L. (1961). Autoradiographic study of cell migration during histogenesis of cerebral cortex in the mouse. Nature *192*, 766-768.

Aoto, K., Nishimura, T., Eto, K., and Motoyama, J. (2002). Mouse GLI3 regulates Fgf8 expression and apoptosis in the developing neural tube, face, and limb bud. Dev Biol *251*, 320-332.

Bagley, J.A., Reumann, D., Bian, S., Levi-Strauss, J., and Knoblich, J.A. (2017). Fused cerebral organoids model interactions between brain regions. Nat Methods *14*, 743-751.

Balasubramanian, M., Willoughby, J., Fry, A.E., Weber, A., Firth, H.V., Deshpande, C., Berg, J.N., Chandler, K., Metcalfe, K.A., Lam, W.*, et al.* (2017). Delineating the phenotypic spectrum of Bainbridge-Ropers syndrome: 12 new patients with de novo, heterozygous, loss-of-function mutations in ASXL3 and review of published literature. J Med Genet *54*, 537-543.

Barrows, N.J., Campos, R.K., Powell, S.T., Prasanth, K.R., Schott-Lerner, G., Soto-Acosta, R., Galarza-Munoz, G., McGrath, E.L., Urrabaz-Garza, R., Gao, J.*, et al.* (2016). A Screen of FDA-Approved Drugs for Inhibitors of Zika Virus Infection. Cell host & microbe *20*, 259-270.

Bellosta, S., Paoletti, R., and Corsini, A. (2004). Safety of statins: focus on clinical pharmacokinetics and drug interactions. Circulation *109*, III50-57.

Bernier, R., Golzio, C., Xiong, B., Stessman, H.A., Coe, B.P., Penn, O., Witherspoon, K., Gerdts, J., Baker, C., Vulto-van Silfhout, A.T.*, et al.* (2014). Disruptive CHD8 mutations define a subtype of autism early in development. Cell *158*, 263-276.

Bershteyn, M., and Kriegstein, A.R. (2013). Cerebral organoids in a dish: progress and prospects. Cell *155*, 19-20.

Bershteyn, M., Nowakowski, T.J., Pollen, A.A., Di Lullo, E., Nene, A., Wynshaw-Boris, A., and Kriegstein, A.R. (2017). Human iPSC-Derived Cerebral Organoids Model Cellular Features of Lissencephaly and Reveal Prolonged Mitosis of Outer Radial Glia. Cell Stem Cell *20*, 435-449 e434.

Berthiaume, L., Peseckis, S.M., and Resh, M.D. (1995). Synthesis and use of iodo-fatty acid analogs. Methods Enzymol *250*, 454-466.

Birey, F., Andersen, J., Makinson, C.D., Islam, S., Wei, W., Huber, N., Fan, H.C., Metzler, K.R.C., Panagiotakos, G., Thom, N.*, et al.* (2017). Assembly of functionally integrated human forebrain spheroids. Nature *545*, 54-59.

Birsoy, K., Possemato, R., Lorbeer, F.K., Bayraktar, E.C., Thiru, P., Yucel, B., Wang, T., Chen, W.W., Clish, C.B., and Sabatini, D.M. (2014). Metabolic determinants of cancer cell sensitivity to glucose limitation and biguanides. Nature *508*, 108-112.

Blaess, S., Szabo, N., Haddad-Tovolli, R., Zhou, X., and Alvarez-Bolado, G. (2014). Sonic hedgehog signaling in the development of the mouse hypothalamus. Front Neuroanat *8*, 156.

Blassberg, R., Macrae, J.I., Briscoe, J., and Jacob, J. (2016). Reduced cholesterol levels impair Smoothened activation in Smith-Lemli-Opitz syndrome. Hum Mol Genet *25*, 693-705.

Boonyasuppayakorn, S., Reichert, E.D., Manzano, M., Nagarajan, K., and Padmanabhan, R. (2014). Amodiaquine, an antimalarial drug, inhibits dengue virus type 2 replication and infectivity. Antiviral Res *106*, 125-134.

Brennand, K.J., Simone, A., Jou, J., Gelboin-Burkhart, C., Tran, N., Sangar, S., Li, Y., Mu, Y., Chen, G., Yu, D.*, et al.* (2011). Modelling schizophrenia using human induced pluripotent stem cells. Nature *473*, 221-225.

Briscoe, J., Chen, Y., Jessell, T.M., and Struhl, G. (2001). A hedgehog-insensitive form of patched provides evidence for direct long-range morphogen activity of sonic hedgehog in the neural tube. Mol Cell *7*, 1279-1291.

Briscoe, J., and Therond, P.P. (2013). The mechanisms of Hedgehog signalling and its roles in development and disease. Nat Rev Mol Cell Biol *14*, 416-429.

Buglino, J.A., and Resh, M.D. (2008). Hhat is a palmitoylacyltransferase with specificity for N-palmitoylation of Sonic Hedgehog. J Biol Chem *283*, 22076-22088.

Butti, C., Santos, M., Uppal, N., and Hof, P.R. (2013). Von Economo neurons: clinical and evolutionary perspectives. Cortex *49*, 312-326.

Byrne, E.F.X., Sircar, R., Miller, P.S., Hedger, G., Luchetti, G., Nachtergaele, S., Tully, M.D., Mydock-McGrane, L., Covey, D.F., Rambo, R.P.*, et al.* (2016). Structural basis of Smoothened regulation by its extracellular domains. Nature *535*, 517-522.

Cao-Lormeau, V.M., Blake, A., Mons, S., Lastere, S., Roche, C., Vanhomwegen, J., Dub, T., Baudouin, L., Teissier, A., Larre, P.*, et al.* (2016). Guillain-Barre Syndrome outbreak associated with Zika virus infection in French Polynesia: a case-control study. Lancet *387*, 1531-1539.

Carteaux, G., Maquart, M., Bedet, A., Contou, D., Brugieres, P., Fourati, S., Cleret de Langavant, L., de Broucker, T., Brun-Buisson, C., Leparc-Goffart, I.*, et al.* (2016). Zika Virus Associated with Meningoencephalitis. The New England journal of medicine *374*, 1595-1596.

Cavodeassi, F. (2014). Integration of anterior neural plate patterning and morphogenesis by the Wnt signaling pathway. Dev Neurobiol *74*, 759-771.

Chailangkarn, T., Trujillo, C.A., Freitas, B.C., Hrvoj-Mihic, B., Herai, R.H., Yu, D.X., Brown, T.T., Marchetto, M.C., Bardy, C., McHenry, L.*, et al.* (2016). A human neurodevelopmental model for Williams syndrome. Nature *536*, 338-343.

Chen, D., Cai, J., Cheng, J., Jing, C., Yin, J., Jiang, J., Peng, Z., and Hao, X. (2015a). Design, Synthesis and Structure-Activity Relationship Optimization of Lycorine Derivatives for HCV Inhibition. Scientific reports *5*, 14972.

Chen, H.M., DeLong, C.J., Bame, M., Rajapakse, I., Herron, T.J., McInnis, M.G., and O'Shea, K.S. (2014). Transcripts involved in calcium signaling and telencephalic neuronal fate are altered in induced pluripotent stem cells from bipolar disorder patients. Transl Psychiatry *4*, e375.

Chen, M.H., Li, Y.J., Kawakami, T., Xu, S.M., and Chuang, P.T. (2004). Palmitoylation is required for the production of a soluble multimeric Hedgehog protein complex and long-range signaling in vertebrates. Genes Dev *18*, 641-659.

Chen, Y., Huang, W.C., Sejourne, J., Clipperton-Allen, A.E., and Page, D.T. (2015b). Pten Mutations Alter Brain Growth Trajectory and Allocation of Cell Types through Elevated beta-Catenin Signaling. J Neurosci *35*, 10252-10267.

Chenn, A. (2008). Wnt/beta-catenin signaling in cerebral cortical development. Organogenesis *4*, 76-80.

Chenn, A., and Walsh, C.A. (2002). Regulation of cerebral cortical size by control of cell cycle exit in neural precursors. Science *297*, 365-369.

Chiang, C., Litingtung, Y., Lee, E., Young, K.E., Corden, J.L., Westphal, H., and Beachy, P.A. (1996). Cyclopia and defective axial patterning in mice lacking Sonic hedgehog gene function. Nature *383*, 407-413.

Cobos, I., Calcagnotto, M.E., Vilaythong, A.J., Thwin, M.T., Noebels, J.L., Baraban, S.C., and Rubenstein, J.L. (2005). Mice lacking Dlx1 show subtype-specific loss of interneurons, reduced inhibition and epilepsy. Nat Neurosci *8*, 1059-1068.

Cordero, D.R., Brugmann, S., Chu, Y., Bajpai, R., Jame, M., and Helms, J.A. (2011). Cranial neural crest cells on the move: their roles in craniofacial development. Am J Med Genet A *155A*, 270-279.

Courchesne, E., Mouton, P.R., Calhoun, M.E., Semendeferi, K., Ahrens-Barbeau, C., Hallet, M.J., Barnes, C.C., and Pierce, K. (2011). Neuron number and size in prefrontal cortex of children with autism. JAMA *306*, 2001-2010.

Courchesne, E., Pierce, K., Schumann, C.M., Redcay, E., Buckwalter, J.A., Kennedy, D.P., and Morgan, J. (2007). Mapping early brain development in autism. Neuron *56*, 399-413.

Creanga, A., Glenn, T.D., Mann, R.K., Saunders, A.M., Talbot, W.S., and Beachy, P.A. (2012). Scube/You activity mediates release of dually lipid-modified Hedgehog signal in soluble form. Genes Dev *26*, 1312-1325.

Curtin, E., Hickey, G., Kamel, G., Davidson, A.J., and Liao, E.C. (2011). Zebrafish wnt9a is expressed in pharyngeal ectoderm and is required for palate and lower jaw development. Mech Dev *128*, 104-115.

Custo Greig, L.F., Woodworth, M.B., Galazo, M.J., Padmanabhan, H., and Macklis, J.D. (2013). Molecular logic of neocortical projection neuron specification, development and diversity. Nat Rev Neurosci *14*, 755-769.

Danesin, C., Peres, J.N., Johansson, M., Snowden, V., Cording, A., Papalopulu, N., and Houart, C. (2009). Integration of telencephalic Wnt and hedgehog signaling center activities by Foxg1. Dev Cell *16*, 576-587.

De Ferrari, G.V., and Moon, R.T. (2006). The ups and downs of Wnt signaling in prevalent neurological disorders. Oncogene *25*, 7545-7553.

De Rubeis, S., He, X., Goldberg, A.P., Poultney, C.S., Samocha, K., Cicek, A.E., Kou, Y., Liu, L., Fromer, M., Walker, S.*, et al.* (2014). Synaptic, transcriptional and chromatin genes disrupted in autism. Nature *515*, 209-215.

Deguchi, K., Inoue, K., Avila, W.E., Lopez-Terrada, D., Antalffy, B.A., Quattrocchi, C.C., Sheldon, M., Mikoshiba, K., D'Arcangelo, G., and

Armstrong, D.L. (2003). Reelin and disabled-1 expression in developing and mature human cortical neurons. J Neuropathol Exp Neurol *62*, 676-684.

Deng, Y.Q., Zhang, N.N., Li, C.F., Tian, M., Hao, J.N., Xie, X.P., Shi, P.Y., and Qin, C.F. (2016). Adenosine Analog NITD008 Is a Potent Inhibitor of Zika Virus. Open Forum Infect Dis *3*, ofw175.

Desai, A.R., and McConnell, S.K. (2000). Progressive restriction in fate potential by neural progenitors during cerebral cortical development. Development *127*, 2863-2872.

Di Lullo, E., and Kriegstein, A.R. (2017). The use of brain organoids to investigate neural development and disease. Nat Rev Neurosci *18*, 573-584.

Dorsky, R.I., Moon, R.T., and Raible, D.W. (1998). Control of neural crest cell fate by the Wnt signalling pathway. Nature *396*, 370-373.

Dougherty, M., Kamel, G., Grimaldi, M., Gfrerer, L., Shubinets, V., Ethier, R., Hickey, G., Cornell, R.A., and Liao, E.C. (2013). Distinct requirements for wnt9a and irf6 in extension and integration mechanisms during zebrafish palate morphogenesis. Development *140*, 76-81.

Durak, O., Gao, F., Kaeser-Woo, Y.J., Rueda, R., Martorell, A.J., Nott, A., Liu, C.Y., Watson, L.A., and Tsai, L.H. (2016). Chd8 mediates cortical neurogenesis via transcriptional regulation of cell cycle and Wnt signaling. Nat Neurosci *19*, 1477-1488.

Durand, C.M., Betancur, C., Boeckers, T.M., Bockmann, J., Chaste, P., Fauchereau, F., Nygren, G., Rastam, M., Gillberg, I.C., Anckarsater, H.*, et al.* (2007). Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders. Nat Genet *39*, 25-27.

Earl, R.K., Turner, T.N., Mefford, H.C., Hudac, C.M., Gerdts, J., Eichler, E.E., and Bernier, R.A. (2017). Clinical phenotype of ASD-associated DYRK1A haploinsufficiency. Mol Autism *8*, 54.

Echelard, Y., Epstein, D.J., St-Jacques, B., Shen, L., Mohler, J., McMahon, J.A., and McMahon, A.P. (1993). Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity. Cell *75*, 1417-1430.

Edison, R.J., and Muenke, M. (2004a). Central nervous system and limb anomalies in case reports of first-trimester statin exposure. The New England journal of medicine *350*, 1579-1582.

Edison, R.J., and Muenke, M. (2004b). Mechanistic and epidemiologic considerations in the evaluation of adverse birth outcomes following gestational exposure to statins. Am J Med Genet A *131*, 287-298.

Eiraku, M., Takata, N., Ishibashi, H., Kawada, M., Sakakura, E., Okuda, S., Sekiguchi, K., Adachi, T., and Sasai, Y. (2011). Self-organizing optic-cup morphogenesis in three-dimensional culture. Nature *472*, 51-56.

Ericson, J., Muhr, J., Placzek, M., Lints, T., Jessell, T.M., and Edlund, T. (1995). Sonic hedgehog induces the differentiation of ventral forebrain neurons: a common signal for ventral patterning within the neural tube. Cell *81*, 747-756.

Ernst, C. (2016). Proliferation and Differentiation Deficits are a Major Convergence Point for Neurodevelopmental Disorders. Trends Neurosci *39*, 290-299.

Fan, C.M., and Tessier-Lavigne, M. (1994). Patterning of mammalian somites by surface ectoderm and notochord: evidence for sclerotome induction by a hedgehog homolog. Cell *79*, 1175-1186.

Fattahi, F., Steinbeck, J.A., Kriks, S., Tchieu, J., Zimmer, B., Kishinevsky, S., Zeltner, N., Mica, Y., El-Nachef, W., Zhao, H.*, et al.* (2016). Deriving human ENS lineages for cell therapy and drug discovery in Hirschsprung disease. Nature *531*, 105-109.

Faundes, V., Newman, W.G., Bernardini, L., Canham, N., Clayton-Smith, J., Dallapiccola, B., Davies, S.J., Demos, M.K., Goldman, A., Gill, H.*, et al.* (2018). Histone Lysine Methylases and Demethylases in the Landscape of Human Developmental Disorders. Am J Hum Genet *102*, 175-187.

Fontebasso, A.M., Papillon-Cavanagh, S., Schwartzentruber, J., Nikbakht, H., Gerges, N., Fiset, P.O., Bechet, D., Faury, D., De Jay, N., Ramkissoon, L.A.*, et al.* (2014). Recurrent somatic mutations in ACVR1 in pediatric midline high-grade astrocytoma. Nat Genet *46*, 462-466.

Frantz, G.D., and McConnell, S.K. (1996). Restriction of late cerebral cortical progenitors to an upper-layer fate. Neuron *17*, 55-61.

Fukuchi-Shimogori, T., and Grove, E.A. (2001). Neocortex patterning by the secreted signaling molecule FGF8. Science *294*, 1071-1074.

Gallagher, D., Voronova, A., Zander, M.A., Cancino, G.I., Bramall, A., Krause, M.P., Abad, C., Tekin, M., Neilsen, P.M., Callen, D.F.*, et al.* (2015). Ankrd11 is a chromatin regulator involved in autism that is essential for neural development. Dev Cell *32*, 31-42.

Garcez, P.P., Loiola, E.C., Madeiro da Costa, R., Higa, L.M., Trindade, P., Delvecchio, R., Nascimento, J.M., Brindeiro, R., Tanuri, A., and Rehen, S.K. (2016). Zika virus impairs growth in human neurospheres and brain organoids. Science *352*, 816-818.

Gaspard, N., Bouschet, T., Hourez, R., Dimidschstein, J., Naeije, G., van den Ameele, J., Espuny-Camacho, I., Herpoel, A., Passante, L., Schiffmann, S.N.*, et al.* (2008). An intrinsic mechanism of corticogenesis from embryonic stem cells. Nature *455*, 351-357.

Gilman, S.R., Iossifov, I., Levy, D., Ronemus, M., Wigler, M., and Vitkup, D. (2011). Rare de novo variants associated with autism implicate a large

functional network of genes involved in formation and function of synapses. Neuron *70*, 898-907.

Glessner, J.T., Wang, K., Cai, G., Korvatska, O., Kim, C.E., Wood, S., Zhang, H., Estes, A., Brune, C.W., Bradfield, J.P.*, et al.* (2009). Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. Nature *459*, 569-573.

Gompers, A.L., Su-Feher, L., Ellegood, J., Copping, N.A., Riyadh, M.A., Stradleigh, T.W., Pride, M.C., Schaffler, M.D., Wade, A.A., Catta-Preta, R.*, et al.* (2017). Germline Chd8 haploinsufficiency alters brain development in mouse. Nat Neurosci *20*, 1062-1073.

Gongal, P.A., French, C.R., and Waskiewicz, A.J. (2011). Aberrant forebrain signaling during early development underlies the generation of holoprosencephaly and coloboma. Biochim Biophys Acta *1812*, 390-401.

Gonzalez, F., Zhu, Z., Shi, Z.D., Lelli, K., Verma, N., Li, Q.V., and Huangfu, D. (2014). An iCRISPR platform for rapid, multiplexable, and inducible genome editing in human pluripotent stem cells. Cell Stem Cell *15*, 215-226.

Gregory, G.D., Vakoc, C.R., Rozovskaia, T., Zheng, X., Patel, S., Nakamura, T., Canaani, E., and Blobel, G.A. (2007). Mammalian ASH1L is a histone methyltransferase that occupies the transcribed region of active genes. Mol Cell Biol *27*, 8466-8479.

Grzadzinski, R., Huerta, M., and Lord, C. (2013). DSM-5 and autism spectrum disorders (ASDs): an opportunity for identifying ASD subtypes. Mol Autism *4*, 12.

Guerrero, I., and Chiang, C. (2007). A conserved mechanism of Hedgehog gradient formation by lipid modifications. Trends Cell Biol *17*, 1-5.

Hahm, K., Sum, E.Y., Fujiwara, Y., Lindeman, G.J., Visvader, J.E., and Orkin, S.H. (2004). Defective neural tube closure and anteroposterior patterning in mice lacking the LIM protein LMO4 or its interacting partner Deaf-1. Mol Cell Biol *24*, 2074-2082.

Hammond, V.E., So, E., Cate, H.S., Britto, J.M., Gunnersen, J.M., and Tan, S.S. (2010). Cortical layer development and orientation is modulated by relative contributions of reelin-negative and -positive neurons in mouse chimeras. Cereb Cortex *20*, 2017-2026.

Hanners, N.W., Eitson, J.L., Usui, N., Richardson, R.B., Wexler, E.M., Konopka, G., and Schoggins, J.W. (2016). Western Zika Virus in Human Fetal Neural Progenitors Persists Long Term with Partial Cytopathic and Limited Immunogenic Effects. Cell reports *15*, 2315-2322.

Harfe, B.D., Scherz, P.J., Nissim, S., Tian, H., McMahon, A.P., and Tabin, C.J. (2004). Evidence for an expansion-based temporal Shh gradient in specifying vertebrate digit identities. Cell *118*, 517-528.

Harrison, S.J., Nishinakamura, R., Jones, K.R., and Monaghan, A.P. (2012). Sall1 regulates cortical neurogenesis and laminar fate specification in mice: implications for neural abnormalities in Townes-Brocks syndrome. Dis Model Mech *5*, 351-365.

Hatini, V., Tao, W., and Lai, E. (1994). Expression of winged helix genes, BF-1 and BF-2, define adjacent domains within the developing forebrain and retina. J Neurobiol *25*, 1293-1309.

Hazlett, H.C., Gu, H., Munsell, B.C., Kim, S.H., Styner, M., Wolff, J.J., Elison, J.T., Swanson, M.R., Zhu, H., Botteron, K.N.*, et al.* (2017). Early brain development in infants at high risk for autism spectrum disorder. Nature *542*, 348-351.

He, J., Qi, W.B., Wang, L., Tian, J., Jiao, P.R., Liu, G.Q., Ye, W.C., and Liao, M. (2013). Amaryllidaceae alkaloids inhibit nuclear-to-cytoplasmic export of ribonucleoprotein (RNP) complex of highly pathogenic avian influenza virus H5N1. Influenza Other Respir Viruses *7*, 922-931.

Hernandez-Munoz, I., Lund, A.H., van der Stoop, P., Boutsma, E., Muijrers, I., Verhoeven, E., Nusinow, D.A., Panning, B., Marahrens, Y., and van Lohuizen, M. (2005). Stable X chromosome inactivation involves the PRC1 Polycomb complex and requires histone MACROH2A1 and the CULLIN3/SPOP ubiquitin E3 ligase. Proc Natl Acad Sci U S A *102*, 7635-7640.

Hindson, C.M., Chevillet, J.R., Briggs, H.A., Gallichotte, E.N., Ruf, I.K., Hindson, B.J., Vessella, R.L., and Tewari, M. (2013). Absolute quantification by droplet digital PCR versus analog real-time PCR. Nat Methods *10*, 1003-1005.

Hirabayashi, Y., Itoh, Y., Tabata, H., Nakajima, K., Akiyama, T., Masuyama, N., and Gotoh, Y. (2004). The Wnt/beta-catenin pathway directs neuronal differentiation of cortical neural precursor cells. Development *131*, 2791-2801.

Hirabayashi, Y., Suzki, N., Tsuboi, M., Endo, T.A., Toyoda, T., Shinga, J., Koseki, H., Vidal, M., and Gotoh, Y. (2009). Polycomb limits the neurogenic competence of neural precursor cells to promote astrogenic fate transition. Neuron *63*, 600-613.

Honig, L.S. (1981). Positional signal transmission in the developing chick limb. Nature *291*, 72-73.

Houart, C., Caneparo, L., Heisenberg, C., Barth, K., Take-Uchi, M., and Wilson, S. (2002). Establishment of the telencephalon during gastrulation by local antagonism of Wnt signaling. Neuron *35*, 255-265.

Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. Cell *157*, 1262-1278.

Huang, P., Nedelcu, D., Watanabe, M., Jao, C., Kim, Y., Liu, J., and Salic, A. (2016). Cellular Cholesterol Directly Activates Smoothened in Hedgehog Signaling. Cell *166*, 1176-1187 e1114.

Huang, Y.C., Shih, H.Y., Lin, S.J., Chiu, C.C., Ma, T.L., Yeh, T.H., and Cheng, Y.C. (2015). The epigenetic factor Kmt2a/Mll1 regulates neural progenitor proliferation and neuronal and glial differentiation. Dev Neurobiol *75*, 452-462.

Iossifov, I., Levy, D., Allen, J., Ye, K., Ronemus, M., Lee, Y.H., Yamrom, B., and Wigler, M. (2015). Low load for disruptive mutations in autism genes and their biased transmission. Proc Natl Acad Sci U S A *112*, E5600-5607.

Iossifov, I., O'Roak, B.J., Sanders, S.J., Ronemus, M., Krumm, N., Levy, D., Stessman, H.A., Witherspoon, K.T., Vives, L., Patterson, K.E.*, et al.* (2014). The contribution of de novo coding mutations to autism spectrum disorder. Nature *515*, 216-221.

Jamain, S., Quach, H., Betancur, C., Rastam, M., Colineaux, C., Gillberg, I.C., Soderstrom, H., Giros, B., Leboyer, M., Gillberg, C.*, et al.* (2003). Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism. Nat Genet *34*, 27-29.

Jessell, T.M. (2000). Neuronal specification in the spinal cord: inductive signals and transcriptional codes. Nat Rev Genet *1*, 20-29.

Jo, J., Xiao, Y., Sun, A.X., Cukuroglu, E., Tran, H.D., Goke, J., Tan, Z.Y., Saw, T.Y., Tan, C.P., Lokman, H.*, et al.* (2016). Midbrain-like Organoids from Human Pluripotent Stem Cells Contain Functional Dopaminergic and Neuromelanin-Producing Neurons. Cell Stem Cell *19*, 248-257.

Johnson, M.B., Wang, P.P., Atabay, K.D., Murphy, E.A., Doan, R.N., Hecht, J.L., and Walsh, C.A. (2015). Single-cell analysis reveals transcriptional heterogeneity of neural progenitors in human cortex. Nat Neurosci *18*, 637-646.

Jung, E.M., Moffat, J.J., Liu, J., Dravid, S.M., Gurumurthy, C.B., and Kim, W.Y. (2017). Arid1b haploinsufficiency disrupts cortical interneuron development and mouse behavior. Nat Neurosci *20*, 1694-1707.

Kadoshima, T., Sakaguchi, H., Nakano, T., Soen, M., Ando, S., Eiraku, M., and Sasai, Y. (2013). Self-organization of axial polarity, inside-out layer pattern, and species-specific progenitor dynamics in human ES cell-derived neocortex. Proc Natl Acad Sci U S A *110*, 20284-20289.

Kalkman, H.O. (2012). A review of the evidence for the canonical Wnt pathway in autism spectrum disorders. Mol Autism *3*, 10.

Kamel, G., Hoyos, T., Rochard, L., Dougherty, M., Kong, Y., Tse, W., Shubinets, V., Grimaldi, M., and Liao, E.C. (2013). Requirement for frzb and fzd7a in cranial neural crest convergence and extension mechanisms during zebrafish palate and jaw morphogenesis. Dev Biol *381*, 423-433.

Kanai, M.I., Okabe, M., and Hiromi, Y. (2005). seven-up Controls switching of transcription factors that specify temporal identities of Drosophila neuroblasts. Dev Cell *8*, 203-213.

Kazmin, A., Garcia-Bournissen, F., and Koren, G. (2007). Risks of statin use during pregnancy: a systematic review. J Obstet Gynaecol Can *29*, 906-908.

Kelava, I., and Lancaster, M.A. (2016). Stem Cell Models of Human Brain Development. Cell Stem Cell *18*, 736-748.

Kiecker, C., and Niehrs, C. (2001). A morphogen gradient of Wnt/beta-catenin signalling regulates anteroposterior neural patterning in Xenopus. Development *128*, 4189-4201.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol *14*, R36.

Kim, W.Y., Wang, X., Wu, Y., Doble, B.W., Patel, S., Woodgett, J.R., and Snider, W.D. (2009). GSK-3 is a master regulator of neural progenitor homeostasis. Nat Neurosci *12*, 1390-1397.

Koemans, T.S., Kleefstra, T., Chubak, M.C., Stone, M.H., Reijnders, M.R.F., de Munnik, S., Willemsen, M.H., Fenckova, M., Stumpel, C., Bok, L.A.*, et al.* (2017). Functional convergence of histone methyltransferases EHMT1 and KMT2C involved in intellectual disability and autism spectrum disorder. PLoS Genet *13*, e1006864.

Kohwi, M., and Doe, C.Q. (2013). Temporal fate specification and neural progenitor competence during development. Nat Rev Neurosci *14*, 823-838.

Kohwi, M., Lupton, J.R., Lai, S.L., Miller, M.R., and Doe, C.Q. (2013). Developmentally regulated subnuclear genome reorganization restricts neural progenitor competence in Drosophila. Cell *152*, 97-108.

Krumm, N., O'Roak, B.J., Shendure, J., and Eichler, E.E. (2014). A de novo convergence of autism genetics and molecular neuroscience. Trends Neurosci *37*, 95-105.

Krumm, N., Turner, T.N., Baker, C., Vives, L., Mohajeri, K., Witherspoon, K., Raja, A., Coe, B.P., Stessman, H.A., He, Z.X.*, et al.* (2015). Excess of rare, inherited truncating mutations in autism. Nat Genet *47*, 582-588.

Kurabayashi, N., and Sanada, K. (2013). Increased dosage of DYRK1A and DSCR1 delays neuronal differentiation in neocortical progenitor cells. Genes Dev *27*, 2708-2721.

Kuwabara, T., Hsieh, J., Muotri, A., Yeo, G., Warashina, M., Lie, D.C., Moore, L., Nakashima, K., Asashima, M., and Gage, F.H. (2009). Wnt-mediated activation of NeuroD1 and retro-elements during adult neurogenesis. Nat Neurosci *12*, 1097-1105.

Kwan, K.Y., Lam, M.M., Krsnik, Z., Kawasawa, Y.I., Lefebvre, V., and Sestan, N. (2008). SOX5 postmitotically regulates migration, postmigratory differentiation, and projections of subplate and deep-layer neocortical neurons. Proc Natl Acad Sci U S A *105*, 16021-16026.

Lai, T., Jabaudon, D., Molyneaux, B.J., Azim, E., Arlotta, P., Menezes, J.R., and Macklis, J.D. (2008). SOX5 controls the sequential generation of distinct corticofugal neuron subtypes. Neuron *57*, 232-247.

Lakoma, J., Garcia-Alonso, L., and Luque, J.M. (2011). Reelin sets the pace of neocortical neurogenesis. Development *138*, 5223-5234.

Lamb, A.N., Rosenfeld, J.A., Neill, N.J., Talkowski, M.E., Blumenthal, I., Girirajan, S., Keelean-Fuller, D., Fan, Z., Pouncey, J., Stevens, C.*, et al.* (2012). Haploinsufficiency of SOX5 at 12p12.1 is associated with developmental delays with prominent language delay, behavior problems, and mild dysmorphic features. Hum Mutat *33*, 728-740.

Lancaster, M.A., Corsini, N.S., Wolfinger, S., Gustafson, E.H., Phillips, A.W., Burkard, T.R., Otani, T., Livesey, F.J., and Knoblich, J.A. (2017). Guided self-organization and cortical plate formation in human brain organoids. Nat Biotechnol *35*, 659-666.

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

Lancaster, M.A., and Knoblich, J.A. (2014b). Organogenesis in a dish: modeling development and disease using organoid technologies. Science *345*, 1247125.

Lancaster, M.A., Renner, M., Martin, C.A., Wenzel, D., Bicknell, L.S., Hurles, M.E., Homfray, T., Penninger, J.M., Jackson, A.P., and Knoblich, J.A. (2013). Cerebral organoids model human brain development and microcephaly. Nature *501*, 373-379.

Lee, S.M., Tole, S., Grove, E., and McMahon, A.P. (2000). A local Wnt-3a signal is required for development of the mammalian hippocampus. Development *127*, 457-467.

Lewis, P.M., Dunn, M.P., McMahon, J.A., Logan, M., Martin, J.F., St-Jacques, B., and McMahon, A.P. (2001). Cholesterol modification of sonic hedgehog is required for long-range signaling activity and effective modulation of signaling by Ptc1. Cell *105*, 599-612.

Li, H., Saucedo-Cuevas, L., Regla-Nava, J.A., Chai, G., Sheets, N., Tang, W., Terskikh, A.V., Shresta, S., and Gleeson, J.G. (2016). Zika Virus Infects Neural Progenitors in the Adult Mouse Brain and Alters Proliferation. Cell Stem Cell *19*, 593-598.

Li, H., Tierney, C., Wen, L., Wu, J.Y., and Rao, Y. (1997). A single morphogenetic field gives rise to two retina primordia under the influence of the prechordal plate. Development *124*, 603-615.

Lui, J.H., Hansen, D.V., and Kriegstein, A.R. (2011). Development and evolution of the human neocortex. Cell *146*, 18-36.

Lupo, G., Harris, W.A., and Lewis, K.E. (2006). Mechanisms of ventral patterning in the vertebrate nervous system. Nat Rev Neurosci *7*, 103-114.

Manuel, M., Martynoga, B., Yu, T., West, J.D., Mason, J.O., and Price, D.J. (2010). The transcription factor Foxg1 regulates the competence of telencephalic cells to adopt subpallial fates in mice. Development *137*, 487-497.

Marcus, E.A., Kintner, C., and Harris, W. (1998). The role of GSK3beta in regulating neuronal differentiation in Xenopus laevis. Mol Cell Neurosci *12*, 269-280.

Mariani, J., Coppola, G., Zhang, P., Abyzov, A., Provini, L., Tomasini, L., Amenduni, M., Szekely, A., Palejev, D., Wilson, M.*, et al.* (2015). FOXG1-Dependent Dysregulation of GABA/Glutamate Neuron Differentiation in Autism Spectrum Disorders. Cell *162*, 375-390.

Maroof, A.M., Keros, S., Tyson, J.A., Ying, S.W., Ganat, Y.M., Merkle, F.T., Liu, B., Goulburn, A., Stanley, E.G., Elefanty, A.G.*, et al.* (2013). Directed differentiation and functional maturation of cortical interneurons from human embryonic stem cells. Cell Stem Cell *12*, 559-572.

Martynoga, B., Morrison, H., Price, D.J., and Mason, J.O. (2005). Foxg1 is required for specification of ventral telencephalon and region-specific regulation of dorsal telencephalic precursor proliferation and apoptosis. Dev Biol *283*, 113-127.

McClellan, J., and King, M.C. (2010). Genetic heterogeneity in human disease. Cell *141*, 210-217.

McConnell, S.K. (1985). Migration and differentiation of cerebral cortical neurons after transplantation into the brains of ferrets. Science *229*, 1268-1271.

McConnell, S.K. (1988). Fates of visual cortical neurons in the ferret after isochronic and heterochronic transplantation. J Neurosci *8*, 945-974.

McGlinn, E., and Tabin, C.J. (2006). Mechanistic insight into how Shh patterns the vertebrate limb. Curr Opin Genet Dev *16*, 426-432.

Mecharles, S., Herrmann, C., Poullain, P., Tran, T.H., Deschamps, N., Mathon, G., Landais, A., Breurec, S., and Lannuzel, A. (2016). Acute myelitis due to Zika virus infection. Lancet *387*, 1481.

Merchan, P., Bardet, S.M., Puelles, L., and Ferran, J.L. (2011). Comparison of Pretectal Genoarchitectonic Pattern between Quail and Chicken Embryos. Front Neuroanat *5*, 23.

Merkle, F.T., Ghosh, S., Kamitaki, N., Mitchell, J., Avior, Y., Mello, C., Kashin, S., Mekhoubad, S., Ilic, D., Charlton, M.*, et al.* (2017). Human pluripotent stem cells recurrently acquire and expand dominant negative P53 mutations. Nature *545*, 229-233.

Merkle, F.T., Maroof, A., Wataya, T., Sasai, Y., Studer, L., Eggan, K., and Schier, A.F. (2015). Generation of neuropeptidergic hypothalamic neurons from human pluripotent stem cells. Development *142*, 633-643.

Miles, J.H., Takahashi, T.N., Hong, J., Munden, N., Flournoy, N., Braddock, S.R., Martin, R.A., Bocian, M.E., Spence, M.A., Hillman, R.E.*, et al.* (2008). Development and validation of a measure of dysmorphology: useful for autism subgroup classification. Am J Med Genet A *146A*, 1101-1116.

Mlakar, J., Korva, M., Tul, N., Popovic, M., Poljsak-Prijatelj, M., Mraz, J., Kolenc, M., Resman Rus, K., Vesnaver Vipotnik, T., Fabjan Vodusek, V.*, et al.* (2016). Zika Virus Associated with Microcephaly. The New England journal of medicine *374*, 951-958.

Monuki, E.S., Porter, F.D., and Walsh, C.A. (2001). Patterning of the dorsal telencephalon and cerebral cortex by a roof plate-Lhx2 pathway. Neuron *32*, 591-604.

Muguruma, K., Nishiyama, A., Kawakami, H., Hashimoto, K., and Sasai, Y. (2015). Self-organization of polarized cerebellar tissue in 3D culture of human pluripotent stem cells. Cell reports *10*, 537-550.

Munji, R.N., Choe, Y., Li, G., Siegenthaler, J.A., and Pleasure, S.J. (2011). Wnt signaling regulates neuronal differentiation of cortical intermediate progenitors. J Neurosci *31*, 1676-1687.

Niida, A., Hiroko, T., Kasai, M., Furukawa, Y., Nakamura, Y., Suzuki, Y., Sugano, S., and Akiyama, T. (2004). DKK1, a negative regulator of Wnt signaling, is a target of the beta-catenin/TCF pathway. Oncogene *23*, 8520-8526.

Nowakowski, T.J., Bhaduri, A., Pollen, A.A., Alvarado, B., Mostajo-Radji, M.A., Di Lullo, E., Haeussler, M., Sandoval-Espinosa, C., Liu, S.J., Velmeshev, D.*, et al.* (2017). Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. Science *358*, 1318-1323.

O'Leary, D.D., Chou, S.J., and Sahara, S. (2007). Area patterning of the mammalian cortex. Neuron *56*, 252-269.

Ockeloen, C.W., Willemsen, M.H., de Munnik, S., van Bon, B.W., de Leeuw, N., Verrips, A., Kant, S.G., Jones, E.A., Brunner, H.G., van Loon, R.L.*, et al.* (2015). Further delineation of the KBG syndrome phenotype caused by ANKRD11 aberrations. Eur J Hum Genet *23*, 1176-1185.

Okamoto, N., Miya, F., Tsunoda, T., Kato, M., Saitoh, S., Yamasaki, M., Kanemura, Y., and Kosaki, K. (2017). Novel MCA/ID syndrome with ASH1L mutation. Am J Med Genet A *173*, 1644-1648.

Oliver, G., Mailhos, A., Wehr, R., Copeland, N.G., Jenkins, N.A., and Gruss, P. (1995). Six3, a murine homologue of the sine oculis gene, demarcates the most anterior border of the developing neural plate and is expressed during eye development. Development *121*, 4045-4055.

Olliaro, P., and Mussano, P. (2003). Amodiaquine for treating malaria. Cochrane Database Syst Rev, CD000016.

Packer, A. (2016a). Enrichment of factors regulating canonical Wnt signaling among autism risk genes. Mol Psychiatry.

Packer, A. (2016b). Neocortical neurogenesis and the etiology of autism spectrum disorder. Neurosci Biobehav Rev *64*, 185-195.

Parikshak, N.N., Luo, R., Zhang, A., Won, H., Lowe, J.K., Chandran, V., Horvath, S., and Geschwind, D.H. (2013). Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. Cell *155*, 1008-1021.

Parra, B., Lizarazo, J., Jimenez-Arango, J.A., Zea-Vera, A.F., Gonzalez-Manrique, G., Vargas, J., Angarita, J.A., Zuniga, G., Lopez-Gonzalez, R., Beltran, C.L.*, et al.* (2016). Guillain-Barre Syndrome Associated with Zika Virus Infection in Colombia. The New England journal of medicine.

Pasca, S.P., Portmann, T., Voineagu, I., Yazawa, M., Shcheglovitov, A., Pasca, A.M., Cord, B., Palmer, T.D., Chikahisa, S., Nishino, S.*, et al.* (2011). Using iPSC-derived neurons to uncover cellular phenotypes associated with Timothy syndrome. Nat Med *17*, 1657-1662.

Pearson, B.J., and Doe, C.Q. (2003). Regulation of neuroblast competence in Drosophila. Nature *425*, 624-628.

Piunti, A., and Shilatifard, A. (2016). Epigenetic balance of gene expression by Polycomb and COMPASS families. Science *352*, aad9780.

Puelles, L., Harrison, M., Paxinos, G., and Watson, C. (2013). A developmental ontology for the mammalian brain based on the prosomeric model. Trends Neurosci *36*, 570-578.

Qian, X., Nguyen, H.N., Song, M.M., Hadiono, C., Ogden, S.C., Hammack, C., Yao, B., Hamersky, G.R., Jacob, F., Zhong, C.*, et al.* (2016). Brain-Region-Specific Organoids Using Mini-bioreactors for Modeling ZIKV Exposure. Cell *165*, 1238-1254.

Quadrato, G., Brown, J., and Arlotta, P. (2016). The promises and challenges of human brain organoids as models of neuropsychiatric disease. Nat Med *22*, 1220-1228.

Quadrato, G., Nguyen, T., Macosko, E.Z., Sherwood, J.L., Min Yang, S., Berger, D.R., Maria, N., Scholvin, J., Goldman, M., Kinney, J.P.*, et al.* (2017). Cell diversity and network dynamics in photosensitive human brain organoids. Nature *545*, 48-53.

Rallu, M., Corbin, J.G., and Fishell, G. (2002a). Parsing the prosencephalon. Nat Rev Neurosci *3*, 943-951.

Rallu, M., Machold, R., Gaiano, N., Corbin, J.G., McMahon, A.P., and Fishell, G. (2002b). Dorsoventral patterning is established in the telencephalon of

mutants lacking both Gli3 and Hedgehog signaling. Development *129*, 4963-4974.

Rasmussen, S.A., Jamieson, D.J., Honein, M.A., and Petersen, L.R. (2016). Zika Virus and Birth Defects--Reviewing the Evidence for Causality. The New England journal of medicine *374*, 1981-1987.

Redin, C., Brand, H., Collins, R.L., Kammin, T., Mitchell, E., Hodge, J.C., Hanscom, C., Pillalamarri, V., Seabra, C.M., Abbott, M.A*., et al.* (2017). The genomic landscape of balanced cytogenetic abnormalities associated with human congenital anomalies. Nat Genet *49*, 36-45.

Renner, M., Lancaster, M.A., Bian, S., Choi, H., Ku, T., Peer, A., Chung, K., and Knoblich, J.A. (2017). Self-organized developmental patterning and differentiation in cerebral organoids. EMBO J *36*, 1316-1329.

Rochard, L., Monica, S.D., Ling, I.T., Kong, Y., Roberson, S., Harland, R., Halpern, M., and Liao, E.C. (2016). Roles of Wnt pathway genes wls, wnt9a, wnt5b, frzb and gpc4 in regulating convergent-extension during zebrafish palate morphogenesis. Development *143*, 2541-2547.

Roessler, E., and Muenke, M. (2010). The molecular genetics of holoprosencephaly. Am J Med Genet C Semin Med Genet *154C*, 52-61.

Ronemus, M., Iossifov, I., Levy, D., and Wigler, M. (2014). The role of de novo mutations in the genetics of autism spectrum disorders. Nat Rev Genet *15*, 133-141.

Sagner, A., and Briscoe, J. (2017). Morphogen interpretation: concentration, time, competence, and signaling dynamics. Wiley Interdiscip Rev Dev Biol *6*.

Sanders, S.J., He, X., Willsey, A.J., Ercan-Sencicek, A.G., Samocha, K.E., Cicek, A.E., Murtha, M.T., Bal, V.H., Bishop, S.L., Dong, S*., et al.* (2015). Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci. Neuron *87*, 1215-1233.

Sanders, T.A., Llagostera, E., and Barna, M. (2013). Specialized filopodia direct long-range transport of SHH during vertebrate tissue patterning. Nature *497*, 628-632.

Sasai, Y., Eiraku, M., and Suga, H. (2012). In vitro organogenesis in three dimensions: self-organising stem cells. Development *139*, 4111-4121.

Schizophrenia Working Group of the Psychiatric Genomics, C. (2014). Biological insights from 108 schizophrenia-associated genetic loci. Nature *511*, 421-427.

Schmitz, S.U., Albert, M., Malatesta, M., Morey, L., Johansen, J.V., Bak, M., Tommerup, N., Abarrategui, I., and Helin, K. (2011). Jarid1b targets genes regulating development and is involved in neural differentiation. EMBO J *30*, 4586-4600.

Sebat, J., Lakshmi, B., Malhotra, D., Troge, J., Lese-Martin, C., Walsh, T., Yamrom, B., Yoon, S., Krasnitz, A., Kendall, J.*, et al.* (2007). Strong association of de novo copy number mutations with autism. Science *316*, 445-449.

Seltzer, L.E., Ma, M., Ahmed, S., Bertrand, M., Dobyns, W.B., Wheless, J., and Paciorkowski, A.R. (2014). Epilepsy and outcome in FOXG1-related disorders. Epilepsia *55*, 1292-1300.

Shcheglovitov, A., Shcheglovitova, O., Yazawa, M., Portmann, T., Shu, R., Sebastiano, V., Krawisz, A., Froehlich, W., Bernstein, J.A., Hallmayer, J.F.*, et al.* (2013). SHANK3 and IGF1 restore synaptic deficits in neurons from 22q13 deletion syndrome patients. Nature *503*, 267-271.

Shen, T., Ji, F., Yuan, Z., and Jiao, J. (2015). CHD2 is Required for Embryonic Neurogenesis in the Developing Cerebral Cortex. Stem Cells *33*, 1794-1806.

Shimamura, K., and Rubenstein, J.L. (1997). Inductive interactions direct early regionalization of the mouse forebrain. Development *124*, 2709-2718.

Shinya, M., Eschbach, C., Clark, M., Lehrach, H., and Furutani-Seiki, M. (2000). Zebrafish Dkk1, induced by the pre-MBT Wnt signaling, is secreted from the prechordal plate and patterns the anterior neural plate. Mech Dev *98*, 3-17.

Shiraishi, A., Muguruma, K., and Sasai, Y. (2017). Generation of thalamic neurons from mouse embryonic stem cells. Development *144*, 1211-1220.

Srivastava, A., Ritesh, K.C., Tsan, Y.C., Liao, R., Su, F., Cao, X., Hannibal, M.C., Keegan, C.E., Chinnaiyan, A.M., Martin, D.M.*, et al.* (2016). De novo dominant ASXL3 mutations alter H2A deubiquitination and transcription in Bainbridge-Ropers syndrome. Hum Mol Genet *25*, 597-608.

Sterneckert, J.L., Reinhardt, P., and Scholer, H.R. (2014). Investigating human disease using stem cell models. Nat Rev Genet *15*, 625-639.

Stoner, R., Chow, M.L., Boyle, M.P., Sunkin, S.M., Mouton, P.R., Roy, S., Wynshaw-Boris, A., Colamarino, S.A., Lein, E.S., and Courchesne, E. (2014). Patches of disorganization in the neocortex of children with autism. The New England journal of medicine *370*, 1209-1219.

Suga, H., Kadoshima, T., Minaguchi, M., Ohgushi, M., Soen, M., Nakano, T., Takata, N., Wataya, T., Muguruma, K., Miyoshi, H.*, et al.* (2011). Self-formation of functional adenohypophysis in three-dimensional culture. Nature *480*, 57-62.

Sweet, D.J. (2017). iPSCs Meet GWAS: The NextGen Consortium. Cell Stem Cell *20*, 417-418.

Taguchi, N., Rubin, E.T., Hosokawa, A., Choi, J., Ying, A.Y., Moretti, M.E., Koren, G., and Ito, S. (2008). Prenatal exposure to HMG-CoA reductase

inhibitors: effects on fetal and neonatal outcomes. Reprod Toxicol *26*, 175-177.

Tang, H., Hammack, C., Ogden, S.C., Wen, Z., Qian, X., Li, Y., Yao, B., Shin, J., Zhang, F., Lee, E.M.*, et al.* (2016). Zika Virus Infects Human Cortical Neural Progenitors and Attenuates Their Growth. Cell Stem Cell *18*, 587-590.

Tian, H., Jeong, J., Harfe, B.D., Tabin, C.J., and McMahon, A.P. (2005). Mouse Disp1 is required in sonic hedgehog-expressing cells for paracrine activity of the cholesterol-modified ligand. Development *132*, 133-142.

Topol, A., Tran, N.N., and Brennand, K.J. (2015). A guide to generating and using hiPSC derived NPCs for the study of neurological diseases. J Vis Exp, e52495.

Touma, J.J., Weckerle, F.F., and Cleary, M.D. (2012). Drosophila Polycomb complexes restrict neuroblast competence to generate motoneurons. Development *139*, 657-666.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol *28*, 511-515.

Tukachinsky, H., Kuzmickas, R.P., Jao, C.Y., Liu, J., and Salic, A. (2012). Dispatched and scube mediate the efficient secretion of the cholesterol-modified hedgehog ligand. Cell reports *2*, 308-320.

Vulto-van Silfhout, A.T., Rajamanickam, S., Jensik, P.J., Vergult, S., de Rocker, N., Newhall, K.J., Raghavan, R., Reardon, S.N., Jarrett, K., McIntyre, T.*, et al.* (2014). Mutations affecting the SAND domain of DEAF1 cause intellectual disability with severe speech impairment and behavioral problems. Am J Hum Genet *94*, 649-661.

Willsey, A.J., Sanders, S.J., Li, M., Dong, S., Tebbenkamp, A.T., Muhle, R.A., Reilly, S.K., Lin, L., Fertuzinhos, S., Miller, J.A.*, et al.* (2013). Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. Cell *155*, 997-1007.

Winterfeld, U., Allignol, A., Panchaud, A., Rothuizen, L.E., Merlob, P., Cuppers-Maarschalkerweerd, B., Vial, T., Stephens, S., Clementi, M., De Santis, M.*, et al.* (2013). Pregnancy outcome following maternal exposure to statins: a multicentre prospective study. BJOG *120*, 463-471.

Wolpert, L. (1969). Positional information and the spatial pattern of cellular differentiation. J Theor Biol *25*, 1-47.

Xu, M., Lee, E.M., Wen, Z., Cheng, Y., Huang, W.K., Qian, X., Tcw, J., Kouznetsova, J., Ogden, S.C., Hammack, C.*, et al.* (2016). Identification of small-molecule inhibitors of Zika virus infection and induced neural cell death via a drug repurposing screen. Nat Med.

Yu, C., Mannan, A.M., Yvone, G.M., Ross, K.N., Zhang, Y.L., Marton, M.A., Taylor, B.R., Crenshaw, A., Gould, J.Z., Tamayo, P.*, et al.* (2016). High-throughput identification of genotype-specific cancer vulnerabilities in mixtures of barcoded tumor cell lines. Nat Biotechnol *34*, 419-423.

Zeng, X., Goetz, J.A., Suber, L.M., Scott, W.J., Jr., Schreiner, C.M., and Robbins, D.J. (2001). A freely diffusible form of Sonic hedgehog mediates long-range signalling. Nature *411*, 716-720.

Zhang, Y., and Alvarez-Bolado, G. (2016). Differential developmental strategies by Sonic hedgehog in thalamus and hypothalamus. J Chem Neuroanat *75*, 20-27.

Zmurko, J., Marques, R.E., Schols, D., Verbeken, E., Kaptein, S.J., and Neyts, J. (2016). The Viral Polymerase Inhibitor 7-Deaza-2'-C-Methyladenosine Is a Potent Inhibitor of In Vitro Zika Virus Replication and Delays Disease Progression in a Robust Mouse Infection Model. PLoS Negl Trop Dis *10*, e0004695.