PRENATAL ALCOHOL EXPOSURE AND THE DEVELOPING BRAIN

LONG NON-CODING RNA MECHANISMS & TRANSCRIPTOMIC CHANGES

A Dissertation

by

NIHAL ASHRAF MOHAMMED MOHAMMED ELFARGHALY SALEM

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee	Rajesh C. Miranda
Committee Members	Farida Sohrabji
	Michael C. Golding
	C. Jane Welsh
Head of Department	Michael Smotherman

May 2021

Major Subject: Neuroscience

ABSTRACT

Fetal alcohol exposure is the leading non-genetic cause of neurodevelopmental disabilities. In the first part of this work, I utilize a neural stem cell culture model to study the effects of alcohol (ethanol) exposure on a novel pseudogene family member of the pluripotency factor Oct4/Pou5f1. This pseudogene locus, mOct4pg9, is transcribed into a long non-coding RNA. Alcohol exposure resulted in its upregulation in neural progenitor cells and I show that its upregulation mediates alcohol's pro-maturation effects on neural progenitor cells. In the second part, I utilize a mouse model of single binge dose prenatal alcohol exposure and study cell-type specific effects of alcohol exposure on fetal cortex development. I identified fetal sex-specific cell-type specific prenatal alcohol effects in the fetal developing cortex, including effects mediated by the loss of a long non-coding RNA, Xist, resulting in loss of X-chromosome inactivation

DEDICATIONS

I dedicate this dissertation to my family. To my late mother, Omnia Taher, I remember how proud she was when I got accepted to the Ph.D. program. Her words to me, “to remember my skills and not to be afraid,” remained a driving force in my journey even when she was no longer present to say them. To my father, Ashraf Salem, he supports me every day telling me that there is nothing I cannot achieve; he flew thousands of miles to stay with my children when I traveled to conferences. He calls me after each exam or presentation to make sure I did well. To my brother, Mahmoud Salem, who patiently and knowledgeably guided me through my journey to learn how to code, which had a significant impact on my graduate career, I wouldn’t have been able to learn how to program without his help! To my husband, Mohamed Elkholy, for his continuous support through the journey and for encouraging me to pursue my dream. To my children, AbdAllah and MennatAllah, I often get asked, “how do you do graduate school with two children?” my answer was usually, I wouldn’t have done it without them. They filled every day of this journey with joy and fun, supported me happily and patiently through exams’ and deadlines’ stress, and celebrated with me every achievement. They have been the best partners during this wonderful journey, and I owe them a lot.

ACKNOWLEDGMENTS

I am grateful for every moment of my Ph.D. journey and everyone who made this journey possible, special, and rewarding.

I want to thank my mentor Dr. Rajesh Miranda for believing in me and giving me the opportunity to be a scientist. I walked to his office to interview him for a course during my Masters, in which he took the time to talk to me about his scientific journey and his research; during this interview, I knew right away that I want to join his laboratory as a directed studies student. In his laboratory, I grew a passion for research and decided to pursue my Ph.D. He has provided me incredible mentorship, support, and encouragement; he encouraged and supported me to pursue every opportunity that would further my career. His door is always open to his students. He patiently spent the time discussing ideas and results with me, answering my questions, and teaching me. The environment he creates in his laboratory encourages collaborations and good collegiality and makes all the laboratory members a big family supporting one another.

I want to thank my committee members, Drs. Farida Sohrabji, Michael Golding, and Jane Welsh for their continuous support. I had the privilege to interact with Dr. Sohrabji, learn from her daily, and rotate in her lab. I took the fetal and embryo physiology course with Dr. Golding, and I thought this was the best-taught course I took during my graduate studies. I had the privilege to interact with Dr. Welsh during her term as Chair for the TAMIN program. Their continuous feedback on my projects enhanced the projects and taught me how to engage in productive scientific discussion and think critically about my data. I am grateful for the incredible support they provided me.

I want to thank Texas A&M Institute for Neuroscience for giving me the opportunity to be a member of its cohort of Ph.D. students and for providing all the support during my Ph.D. journey & the Texas A&M Master of Biotechnology Program for opening the doors to my graduate school journey.

Daily, I had the privilege to meet and interact with students, staff, and faculty in the department of Neuroscience and Experimental Therapeutics, all of which I called family, and no day passed without receiving encouraging and positive words from them. Everyone in the department provided help and advice without hesitation when I asked them.

I want to thank Dr. Miranda's current and former laboratory members who became family and friends for life. Dr. Amanda Mahnke patiently taught me, among a lot of skills and techniques, how to use a confocal microscope, how to do in-vivo experiments, and how to give a compelling presentation. Alexander Tseng, Dae Chung, Marisa Pinson, and Amy Adams continuously provided help and support inside and outside the lab. I would also like to thank all the former and current undergraduate students in the laboratory, specially Cadianna Garcia & Hooman Jahromi.

CONTRIBUTORS AND FUNDING SOURCES

This work was supervised by a dissertation committee consisting of Dr. Rajesh Miranda (advisor) of the Department of Neuroscience and Experimental Therapeutics, Dr. Farida Sohrabji of the Department of Neuroscience and Experimental Therapeutics, Dr. Michael Golding of Department of Veterinary Physiology, and Dr. Jane Welsh of the Department of Veterinary Integrative Biosciences.

The work conducted for the dissertation was completed by the student independently.

Funding sources:

Graduate study was supported by a grant from the NIH (R01AA024659 to Dr. Rajesh Miranda), funding from the Texas A&M Institute for Neuroscience, and a NIH Blueprint Diversity Specialized Predoctoral to Postdoctoral Advancement in Neuroscience (D-SPAN) fellowship (F99NS113423) to Nihal Salem

TABLE OF CONTENTS

ABSTRACT ii

DEDICATIONS iii

ACKNOWLEDGMENTS iv

CONTRIBUTORS AND FUNDING SOURCES vi

LIST OF FIGURES ix

LIST OF TABLES x

1 INTRODUCTION 1

1.1 Genetic, Epigenetic and Environmental causes of Neurodevelopmental disorders 1
1.2 Neurodevelopmental & intrauterine growth restriction 1
1.3 Genetics & Epigenetics of neurodevelopmental disorders 3
1.4 Environment & neurodevelopmental disorders 6
1.5 Prenatal alcohol exposure and brain development 10
1.6 Long non-coding RNA and brain development 12
1.7 Single cell RNA sequencing to decipher the complexity of the developing brain and the specificity of environmental hits 14
1.8 Conclusion and overall dissertation question: 18

2 A NOVEL OCT4/POU5F1-LIKE NON-CODING RNA CONTROLS NEURAL MATURATION AND MEDIATES DEVELOPMENTAL EFFECTS OF ETHANOL. 20

2.1 Introduction 20
2.2 Methods 23
2.3 Results 34
2.4 Discussion 44
2.5 Figures 51
2.6 Tables 60

3 CELL TYPE AND SEX-SPECIFIC TARGETS OF PRENATAL ALCOHOL EXPOSURE IN THE MOUSE DEVELOPING CORTEX: XIST-DEPENDENT FEMALE SPECIFIC MECHANISMS. 64

3.1 Summary 64
3.2 Introduction 64
3.3 Methods 67
3.4 Results 73
3.5 Discussion 84
3.6 Figures 91
3.7 Tables 102

4 CONCLUSIONS 108

REFERENCES 113

APPENDIX 1 137

APPENDIX 2 155

LIST OF FIGURES

Figure 1-1 Effects of prenatal alcohol exposure on cortical development 19

Figure 2-1: Identification and characterization of Oct4 pseudogenes 51

Figure 2-2: Characterization of mOct4pg9 regulation of Oct4/Pou5f1 53

Figure 2-3: mOct4pg9 lncRNA association with miRNAs and Argonaut chaperone proteins 55

Figure 2-4: Effects of mOct4pg9 ncRNA manipulation on neurosphere size and DNA synthesis..... 57

Figure 2-5 RNA expression level of key stem cell and differentiation-associated genes in lysates from of alcohol treated neurosphere cultures..... 58

Figure 3-1: scRNA sequencing reveals the heterogenicity of ED 14.5 murine developing cortex..... 91

Figure 3-2: Pseudotime analysis orders the cells based on sequence of gene expression changes through neuronal maturation progression 93

Figure 3-3: Effects of prenatal alcohol exposure on cells trajectories and phases 94

Figure 3-4: Prenatal alcohol exposure results in gene dysregulation in murine developing cortex in lineage and sex specific manner..... 96

Figure 3-5: Prenatal alcohol exposure disrupts gene expression modularity 98

Figure 3-6: Female specific responses to prenatal alcohol exposure are mediated by alcohol induced disruption of the female specific Xist gene 100

LIST OF TABLES

Table 2-1 : Showing mouse genomic loci with high sequence homology to mouse Pou5F1/Oct4 mRNA 60

Table 2-2: Human transcripts with high sequence homology to mouse Oct4pg9 ncRNA. 61

Table 2-3 Pathway overrepresentation analysis. 62

Table 2-4: Summary effects of ethanol exposure, Oct4pg9 overexpression, siRNA &CRISPR mediated Oct4pg9 knockdown on the expression of mRNA markers for neural stem cell identity and maturation. 63

Table 3-1: Dam and Sample Ids and maternal blood alcohol level of each sample 102

Table 3-2: Percentages of cells in each cell lineage 103

Table 3-3 Gene ontology enriched genes subclusters of cluster 1 104

Table 3-4: Significant upstream regulator across multiple clusters 105

Table 3-5:IPA function annotation for Eef1a1 combined modules genes (586 genes) 106

Table 3-6:IPA function annotation for Xist combined modules genes (352 genes) 107

1 INTRODUCTION

1.1 GENETIC, EPIGENETIC AND ENVIRONMENTAL CAUSES OF NEURODEVELOPMENTAL DISORDERS

Brain development starts in the first trimester during embryonic period, where genetic and environmental factors act either individually or together, to change brain development trajectories. Neurodevelopmental disorders, arise from genetic mutations, epigenetic changes, environmental factors that directly affect brain development (such as insecticides exposure, and some drugs of abuse), or other environmental factors that can result in intrauterine growth restriction, which would result in altered brain development. Various cellular mechanisms (RNA binding, methylation program, chromosomal deletions, chromatin remodeling) mediate the effects of environmental factors on brain development. Many of those disorders show sexual dimorphism in causes and presentation.

1.2 NEURODEVELOPMENTAL & INTRAUTERINE GROWTH RESTRICTION

Neurodevelopmental adversity can result from intrauterine growth restriction. Small for gestational age (SGA) children evaluated at two years old had significantly lower cognitive, motor and adaptive scores and higher significant risk of low scores in language and adaptive domains, noteworthy that those results were reported after correction of covariates such as socioeconomic status, smoking, and breastfeeding [1]. In this study, a sex difference was reported in one of the domains tested, with males accounting for a significant proportion of lower language and adaptive

scores. In a guinea pig model of intrauterine growth restriction (IUGR), growth-restricted animals had fewer CA1 pyramidal neurons and Purkinje neurons at 4-7 days of age[2]. MRI studies have shown that IUGR infants exhibit persistent deficits in brain structure, including decreased white matter in various cortical regions and cerebellum at 12 months of age, coupled with neurodevelopmental deficiencies in the fine motor coordination domains, compared to healthy term-born infants and appropriate for gestational age (AGA) preterm infants[3].

IUGR affects brain development not necessarily directly by hypoxia but through epigenetic and transcriptional mechanisms, which are sex-specific in many instances. Sex and tissue-specific changes in epigenetic determinants of chromatin structure [4] were shown in IUGR rat model. IUGR animals at birth exhibited decreased genome-wide, CPG island methylation in the brain with no sex differences, while sex differences emerged at postnatal day (PD) 21, where those methylation changes were no longer notable in males while they persisted in female pups. The hippocampus of male rats at PD21 exhibited decreased acetylation at histone H3 lysine K 9, while IUGR female rats shown increased acetylation. IUGR also resulted in impaired metabolism of catecholamines, in a sex-specific manner, low birth-weight piglets from an IUGR model showed increased levels of a serotonin metabolite in the hippocampus, while the low birth-weight females did not show this effect [5].

Fetal growth restriction can be caused by genetic or environmental causes. The Silver-Russel syndrome, SRS, is a growth disorder characterized by prenatal and postnatal growth retardation. SRS patients have characteristic facial features, such as a triangular face, narrow chin, prominent forehead, small jaw, in addition to learning disabilities, speech and language problems. SRS is caused either by imprinting defects (hypomethylation at 11p15 imprinting control regions (ICR) H19-IGF2-ICR), or by other non-epigenetic causes. For example, a mutation in IGF2 [6],

duplication in the NSD1 gene locus [7], or altered methylation at promoters of genes (HOXA4) [8] were all shown to result in SRS. In a sibling-controlled study, SRS patients present with moderate but significant impairment in the cognitive outcome [9]. The overlap between SRS and neural developmental disability can also be explained by reports of SRS association with early IUGR, which has detrimental effects on neurodevelopment [10]. Hypomethylation at the HoxA4 gene promoter was shown in SRS patients who did not otherwise exhibit disomy or imprinting defects. Hypomethylation change was also shown in healthy children with short stature. HoxA4 gene (also known as Hoxd3) has the miRNA miR-10b gene imbedded in its sequence. Hypomethylation at the HoxA4 gene promoter would be expected to affect miRNA miR-10b expression levels. miR10-b is an important regulator of BDNF, and miR-10b overexpression was shown to suppress BDNF[11], suggesting its important role in brain development.

1.3 GENETICS & EPIGENETICS OF NEURODEVELOPMENTAL DISORDERS

CHROMATIN REMODELING MECHANISMS

The SWI/SNF (Switch/Sucrose Non-Fermentable)/BAF (BRG1/BRM-Associated Factor) heteromeric chromatin remodeling, a complex of ∼10–15 different proteins, causes ATP-dependent disassembly of nucleosomes by dissociating histones from DNA, to facilitate chromatin remodeling and transcriptional activation[12]. Phenotype-genotype correlations in 85 patients with mutations in SWI/SNF complex "SWI/SNF-related ID syndromes" [13] indicate that all mutations were associated with agenesis of corpus callosum, speech impairment and syndromic intellectual disability. Mutations in different subunits of this complex were found in the Coffin-Siris syndrome, a genetic condition first described in the literature in 1970 [14], characterized by distinct facial features, underdeveloped fifth digit and variable degrees of

learning disabilities and developmental delays. Additionally, a syndrome characterized by severe intellectual disability, the Nicolaides–Baraitser syndrome, was found to be caused by a mutation in SMARCA2, a member of the SWI/SNF complex[15]. Mice heterozygous for a BRG1 null mutation exhibit exencephaly [16].

RNA BINDING PROTEINS MECHANISMS

Mutations in neuronal RNA binding protein genes are associated with neurodevelopmental disorders. Syndromes within this class, including Fragile-x syndrome (FMRP mutations), Rett syndrome (MeCP2), MeCP2 duplication syndrome are examples of disorders resulting from mutations or gene mis-dosing. Brain-derived MeCP2 associates with splicing factor Prpf3. Many of the mutations in MeCP2 causing Rett syndrome are within Prpf3 binding site disrupting the MeCP2-Prpf3 complex [17]. Additionally, in a mouse model of RETT syndrome, mutations in MeCP2 resulted in a loss of MeCP2 interaction with the transcriptional regulator DHX9 and the splicing factor recruiter LEDGF. Cortices of this mouse model shown dysregulation in splicing of hundreds of mRNAs. Among the outcomes described in this study, dysregulated mRNA splicing in some AMPAR genes, which mediates the majority of fast synaptic transmission in the central nervous system were identified [18]. This dysregulation explains synaptic dysfunction observed in Rett syndrome, since synaptic trafficking of AMPA receptors is altered in the hippocampus of Rett syndrome, resulting in altering synaptic plasticity [19] and reducing long-term potentiation. This is an area that provides therapeutic potential for improving one of the symptoms of RETT syndrome. For example, a small-molecule ligand for the BDNF receptor was shown to restore hippocampal LTP and improve hippocampal-dependent object location memory in female mutant MeCP2 mice[20]. Another MeCP2 related disorder is MeCP2 duplication syndrome, which results in severe intellectual disability, autism, anxiety and other neuropsychiatric symptoms[21]. This

syndrome affects males predominantly however female carriers of the MeCP2 duplication show intellectual disability [22]. Mouse models of MeCP2 duplications show initially high spine density in terminal dendritic branches, which have higher turnover than control mice, resulting in overall less spine density at postnatal week 12, which would correspond to the late onset of symptoms in MeCP2 duplication syndrome.[23]. Another study creating a mouse model of MeCP2 duplication syndrome model was able to recapitulate the behavioral deficits of the syndrome (learning and memory aspects) and electrophysiological measures shown impairment in excitatory synapse function in cultured hippocampal neurons rather than inhibitory[24]. Antisense nucleotide against MeCP2 in MeCP2 duplication mouse model rescued abnormal gene expression in hippocampus prominently after 4 weeks of treatment and robustly at 8 weeks of treatment in adult symptomatic MeCP2 duplication mice, abnormal motor behavior was reversed at week 7 of treatment, anxiety-related behavior and hypoactivity were reversed at week 10 of treatment[25].

Fragile X syndrome is an X-linked syndrome characterized by learning disabilities, cognitive impairment, and characteristic facial features; it might include autistic-like features and ADHD like symptoms. The syndrome occurs due to the loss of Fragile X Mental Retardation Protein (FMRP). The FMR1 gene contains CGG repeats in its 5' UTR. An unaffected, i.e., normal population, may have up to 55 of these trinucleotide repeats, while Fragile X syndrome patients may have more than 200 CGG repeats (Full mutation), resulting in promoter methylation and subsequent of silencing the FMR1 gene and the loss of FMRP. Individuals with repeats ranging from 50 to 200 are called premutation carriers, reviewed in[26]. Although carriers are not affected by Fragile X syndrome, they show some physical features such as prominent ears[27]. Elderly male carriers might show fragile X-associated tremor/ataxia syndrome[28], while female carriers experience Fragile X-associated primary ovarian insufficiency and often report attention deficit

hyperactivity disorder (ADHD), anxiety and depression, which is suggested to be secondary to ovarian dysfunction ([29]). Premutation effects can be explained by the impairment of FMR1 mRNA translation efficiency in the premutation carriers (rather than the complete loss of FMRP in full mutation carriers)[30, 31]. Fragile X syndrome presents in males more than females. Males with full mutation exhibit the behavioral, physical and cognitive features of the syndrome while not all of the females with the full mutation exhibit those features. FMRP is RNA-binding protein, loss of FMRP results in deficiencies in synaptic plasticity. FMRP associates with ribosomes, analysis of FMRP associated RNA identified a list of FRMP target mRNAs, the top ontology of those mRNAs were synaptic transmission and the pathway analysis for those transcripts reveals CREB signaling in neurons, long-term synaptic potentiation, axonal guidance, serotonin signaling among other neuronal transmission-related pathways as the top pathways associated with those transcripts. FMRP was found to bind to those transcripts to repress their translation, therefore loss of FMRP protein would increase the expression of dosage-sensitive transcripts, resulting in impaired synaptic transmission.[32].FMRP was also found to associate with miR-125b,miR-125b targets N-Methyl-D-aspartate (NMDA) receptor subunit NR2A, a process which impacts synaptic plasticity[33].FMR1 knockout mice show dysregulation of axonal growth and motility [34].

1.4 ENVIRONMENT & NEURODEVELOPMENTAL DISORDERS

MATERNAL SMOKING & NEURODEVELOPMENT

Maternal smoking impacts neurodevelopment through a variety of mechanisms. Nicotine results in intrauterine growth restriction [35, 36], which impacts neural development (discussed above). Prenatal nicotine exposure was shown to disrupt various aspects of the HPA axis, maternal serum corticosteroids increased at mid to late gestation after nicotine exposure in rats which would result in fetal adrenocortical dysfunction [37]. Moreover, prenatal nicotine upregulated hippocampal

GAD67 and induced HPA axis hypersensitivity [38]. Disrupted HPA axis is an area of sexual dimorphism, reported in the context of prenatal stress and prenatal ethanol exposure simultaneously, prenatal stress was found to increase plasma ACTH and corticosterone after stressing male and female offspring of stressed dams at the last week of pregnancy, however, male offspring were found to be unaffected[39]. Excessive fetal exposure to glucocorticoids (synthetic) resulted in decreased cortical thickness and resulted in child affective problems which are persistent at 6-10 years.[40]. Children prenatally exposed to second-hand smoke had decreased mental development index score when assessed at six months compared to controls [41], suggesting an increased risk of developmental delay.

Cigarette smoking during pregnancy was shown to decrease miR-16, 21 and miR-146 in the human placentae[42].In neural stem cell culture model, nicotine treatment was shown to increase miR-9 , miR-21, miR-153, miR335 and miR-140-3p[43], those miRNAs are important in neurogenesis and brain development [44-47]. Pyrosequencing of umbilical cords of newborns whose mothers were either smoking during pregnancy shown increased methylation of Cytosine–phosphate–guanine dinucleotide paira (CpGs) at the Insulin-like Growth Factor 2 (IGF2) DMR and not CpGs at H 19 compared to non smokers, those differences were more evident in males than females and were attributed to 21% of smoking-related low birth weight in males compared to only 2% in females [48].

Studying the fetal consequences of maternal smoking and second-hand smoking exposure only regarding nicotine effects misses other important compounds in cigarette smoke that impacts fetal growth. The effect of tobacco smoke extract (TSE) in an in-vitro model of neural differentiation was more detrimental than the effects of its nicotine and benzo(a)pyrene contents individually.

TSE and nicotine both promoted differentiation of PC12 cells (cell line model for neuronal differentiation) resulting in fewer but larger neuronal cells.[49].

Benzo[a]pyrene is an environmental carcinogen present in a varying amount in tobacco cigarette smoke[50]. In utero exposure to benzo[a]pyrene resulted in deficits in cortical neuronal NMDA responses[51].Polyaromatic hydrocarbon exposure resulted in decreased white matter and cognitive and behavioral deficits[52]. Benzopyrene impaired neurodifferentiation of PC12 cells by increasing cell numbers but decreasing neurotransmitter phenotypes (dopaminergic and cholinergic cells)[53]

PESTICIDES AND NEURODEVELOPMENTAL DISORDERS

Prenatal exposure to pesticides and insecticides results in neurodevelopmental consequences, some of which were shown to be sexually dimorphic. Occupational prenatal exposure to pesticides was shown to have an effect on brain stem electrophysiology in affected children with no sex differences, however, impaired neuropsychological function was evident only in females tested at age 6-11 years[54]. Prenatal exposure to Chlorpyrifos(CPF) resulted in decreased IQ and working memory scores when assessed at 7 years of age[55]. Prenatal exposure to CPF also resulted in delayed psychomotor and mental development and attention-deficit/hyperactivity disorder (ADHD) [56]. Prenatal CPF exposure resulted in brain anomalies, MRI scans show frontal and parietal cortical thinning in children highly exposed to CPF[57]. Transcriptomic analysis of murine offspring brains prenatally exposed to cypermethrin(CYP) showed developmental neurotoxicity resulting from changes in mitochondrial function and nucleosome assembly pathways gene expression.[58]

DRUGS OF ABUSE & NEURODEVELOPMENT

Various drugs of abuse have been associated with detrimental outcomes of pregnancy, varying from stillbirth to anomalies and delayed development. Drugs are often co-abused which complicates the maternal and fetal outcomes and complicates the studying of specific drug-specific outcome relationship.

Methamphetamine

Prenatal methamphetamine exposure results in behavioral defects in childhood and adolescence. Prenatally methamphetamine-exposed children evaluated at different ages had higher emotional reactivity and increased anxiety & depression problems (assessed at the age of 3), ADHD problems (evaluated at the age of 5) [59], increased cognitive problems (assessed at the age of 7.5) [60]. Effects of prenatal methamphetamine exposure extended into adolescence. At a follow up at 14 years of age, prenatally methamphetamine-exposed children had delayed school performance as indicated using mathematics points and grade retention[61]. Brain structural changes were also reported, prenatal exposure to methamphetamine resulted in decreased subcortical structures volume (putamen, globus pallidus, and hippocampus) which correlated with poor sustained attention and delayed verbal memory [62].

Marijuana

Prenatal exposure to cannabis was reported to increase aggressive behavior and attention problems in girls (not in boys) at 18-month-old children [63]. Prenatal cannabis exposure resulted in increased frontal cortical thickness when assessed in 6-8 years old children, one limitation of cannabis exposure studies that cannabis is often co-abused with tobacco, however children prenatally exposed to tobacco alone exhibited decreased cortical thickness, co-exposure to cannabis and tobacco resulted in opposite effect[64]. Behavioral effects resulting from prenatal

cannabis exposure weren't short term; increased hyperactivity, impulsivity and inattention symptoms were reported at ten years old children [65]. At the gene expression levels, Dopamine D2 gene was downregulated in the amygdala in human prenatally cannabis exposed fetal specimens, an effect which was significant only in males and not females[66]. On a neurophysiological level, fMRI study on young adults who were prenatally exposed to cannabis showed increased neural activity in the bilateral prefrontal cortex during response inhibition, indicating decreased ability to inhibit responding when required by a response inhibition task[67].

1.5 PRENATAL ALCOHOL EXPOSURE AND BRAIN DEVELOPMENT

Prenatal alcohol exposure (PAE) is common [68] and difficult to prevent due to high rates of untreated alcohol use disorders [69] and unplanned pregnancies [70]. In a recent state-wide study in Texas, 8.4% of assessed newborn infants were positive for a blood alcohol metabolite, indicative of at least 3rd-trimester PAE [71]. PAE results in microcephaly [72] and persistent brain growth deficits [73], and is the leading cause of intellectual and other neurobehavioral disabilities [74]. A recent prospective case ascertainment study showed that between 1.1 - 9% of school-aged children met diagnostic criteria for FASD [75].

PAE results in fetal alcohol spectrum disorders, a cluster of physical and neurobehavioral anomalies [76, 77]. PAE results in growth deficiency (height or weight below the 10th percentile), decreased head circumference, and facial features, including a thin upper lip, smooth philtrum, and shortened palpebral fissures [78, 79]. Neurobehavioral deficits include deficits in attention, executive function and cognitive performance, hyperactivity, and motor skills development, among others [80, 81].

CELL CULTURE MODELS OF PAE

The development of brain regions, including cerebral cortex, can be severely affected by prenatal alcohol exposure. In a Canadian study, Beaulieu and colleagues, using post-mortem studies in FAS patients have shown cerebral cortical dysgenesis along with abnormal neuronal and glial migration [81]. FASD patients have a generally thinner cerebral cortex [82] and decreased cortical neuron formation [83], implicating an effect of ethanol on neural stem cell function (NSC). The murine neuroepithelium derived neurospheres cell culture model was used to model the effect of prenatal ethanol exposure on the developing brain [84]. Alcohol exposure results in increased proliferation in human [85] and rodent neural stem cells [86]. Ethanol treatment of NSCs resulted in increasing both cell migration and the expression of migration-related genes during mitogen-withdrawal differentiation, with cellular adhesion to laminin[87]. Exposing NSCs to alcohol prior to differentiation resulted in decreased neuronal outgrowth and complexity after retinoic acid induced differentiation[86].

ANIMAL MODELS OF PAE

Simple organisms, rodents, and non-human primates' models of prenatal alcohol exposure have been extensively used. Each model provides different face validity advantage, and some challenges (reviewed in [88-90]).

Zebrafish (*Danio rerio*) – an example of a simple organism model-is used as a model to study the effects of embryonic exposure to ethanol on development. While the Zebrafish is evolutionary distanced from mammals , a review [91]discussed their face validity (similar behavioral effects across zebrafish and human), construct validity (similar mechanisms of alcohol induced defects) and their predictive validity (similar drug effects). Embryonic alcohol exposure model affected external and internal physical development of zebra fish -eye diameter

and heart rate- in a dose and time exposure dependent manner[92]. Zebrafish was also used to model neurobehavioral deficits resulting from prenatal alcohol exposure[93]. Zebrafish embryonic alcohol exposure resulted in retinal Wnt signaling defects which was rescued by folic acid or retinoic acid supplementation[94].

Immune activation occurs in the fetal white matter in the first trimester binge alcohol exposure in an ovine model [95]. Prenatal alcohol was additionally shown to affect ovine fetal cerebellar Purkinje cells irrespective of exposure time [96, 97]. Ovine models were additionally used to assess the outcome of potential nutritional and therapeutic interventions after PAE[95, 98, 99]. An ovine model of PAE has shown the potential of use circulating miRNAs as biomarkers for alcohol exposure during pregnancy, in both mother and neonates[100].

Mouse models, an example of rodent models, are a convenient model for studying the effects of prenatal alcohol exposure on embryonic development. Early work [101] showed craniofacial malformations in the offspring as a result of exposure to alcohol during gastrulation. Rodent models offer diverse behavioral phenotypes that can be utilized to test different social, cognitive and affective outcomes of prenatal alcohol exposure (reviewed in [102])

1.6 LONG NON-CODING RNA AND BRAIN DEVELOPMENT

Long non-coding RNAs(lncRNAs) are a class of non-coding RNAs with transcript length exceeding 200 nucleotides. lncRNAs lack the ability to code for functional proteins, but may influence the biology of cells through a number of alternate mechanisms. For instance, lncRNAs can act as decoys for miRNAs thus influencing translation, act as scaffolds to recruit chromatin remodelers, modulate mRNA splicing, or by serving as a source of endogenous small interfering RNAs (siRNAs)(reviewed in [103]).

The dysregulation of lncRNAs was shown in postmortem ASD brains[104], of those RNAs are DLX6 and Evf2. DLX6 is a crucial lncRNA for craniofacial development [105], while Evf2is an important interacting element with the BRG1/BAF170 chromatin remodeling complex, a mechanism which is implicated in FASD[106, 107]. One lncRNA was shown to be dysregulated by alcohol exposure in neural stem cell culture model[108]. Prenatal exposure to environmental toxins was shown to dysregulate lncRNAs resulting in developmental consequences. Prenatal exposure to arsenic alters brain the lncRNA Gas5 levels in age and sex-dependent manner which regulates glucocorticoid mediated gene transcription, resulting in learning and memory deficits and increased depressive-like behaviors[109]. A chromosomal translocation disrupting lnc-NR2F1 was identified in a family manifesting neurodevelopmental symptoms, further studies show that it enhances induced neuronal cell maturation and regulates transcription of neuronal genes, including autism-associated genes[110].

The Xist gene encodes a lncRNA which paints one X-chromosome randomly in females and is crucial for inactivation of one of the two female X chromosomes [111]. Tsix is a non-coding antisense to Xist lncRNA expressed from the active X-chromosome and regulates the early steps of X inactivation[112]. X-chromosome skewness (non-random X chromosome inactivation) has been reported in females with autism[113] and also in females with X-linked mental retardation [114] and Rett syndrome [115]. Decreased Xist, increased Tsix and decreased XCI-regulating factors were reported in female pups prenatally exposed to BPA and is a possible mechanism of BPA induced neurodevelopmental disorders[116]. Over-expression of Xist and X-linked genes was shown in female psychiatric patients (bipolar disorder or major depression)[117].

1.7 SINGLE CELL RNA SEQUENCING TO DECIPHER THE COMPLEXITY OF THE DEVELOPING BRAIN AND THE SPECIFICITY OF ENVIRONMENTAL HITS

DEVELOPMENT OF SINGLE CELL RNA SEQUENCING METHODS

Single-cell RNA sequencing approach has developed from low throughput technique to higher throughput more automated methods. Fluorescence-activated cell sorting (FACS) combined with RNA sequencing , e.g. massively parallel single-cell RNA-sequencing (MARS-Seq)[118], was one of the first endeavors to single cell RNA sequencing. Microwell based platforms, e.g. Cytoseq, based on large arrays containing thousands of microwells used to combine barcoded beads with single cells which are loaded to the microwells through laminar flow allowing for the barcoding of the contents of each cell[119, 120]. The ability only to specifically sequence a predefined set of genes is a limitation of this technique. Microfluidic based methods of single cell isolation and library preparation, like the Fluidigm C1 platform, resulted in increasing the throughput of number of sequenced cells[121]. The development of Dropseq[122] resulted in the advance of single cell genomics due to the ability to sequence larger number of cells, in this method droplets are used to compartmentalize cells into nanoliter-sized reaction chambers followed by molecular barcoding strategy to remember the cell-of-origin of each mRNA. Droplet-based encapsulation of single cells was later commercialized into more cost-effective and time saving 10X Genomics Chromium System(reviewed in [123]). The combination of the advancement in single cell RNA sequencing and the ability to visualize mRNA within tissue sections through spatial transcriptomics[124], allows for the identification map cellular and molecular signature in tissues[125]. Spatial transcriptomic survey of tissues has been applied in

neurodegenerative and Alzheimer's disease research[126-129] but its application in neurodevelopmental disorder research is still lacking.

SINGLE CELL TRANSCRIPTOMICS TO REVEAL TISSUE HETEROGENICITY

Single cell sequencing allowed for the identification of the heterogenicity in cellular composition during brain development and the validation of organoid models. scRNA sequencing of developing human prefrontal cortex from gestational weeks 8 to 26 mapped the timeline of neurogenesis of excitatory neurons in the prefrontal cortex and detect the presence of interneuron progenitors in early developing prefrontal cortex and shown that neurons developed from NPCs in early gestational weeks whereas OPCs and astrocytes differentiated from NPCs in later weeks, pseudotime analysis of the sequenced cells shown the order of development of specific interneurons subtypes[130].scRNA sequencing of oligodendrocyte progenitor cells isolated from neonatal mouse cortices identified previously unknown progenitors of oligodendrocytes resembling adult activated NSCs with OPC developmental trajectory [131]. scRNA sequencing of ganglionic eminence cells dissected from wild-type mouse embryos at embryonic day (E)13.5 (for the MGE) or E14.5 (for the CGE and LGE), identified a common developmental program of gene expression functions in the mitotic progenitors of all three ganglionic eminences, while Postmitotic cells from all eminences pass through distinct precursor states leading into the diversity of cortical interneurons [132]. Both scRNA sequencing and ATAC sequencing revealed the transcriptional and chromatin accessibility dynamics during the development of organoids modeling MGE and cortex specific organoids and identified multiple cell types formed during their development[133].

SINGLE CELL TO IDENTIFY MEDIATORS OF NEURODEVELOPMENTAL DISORDERS

In addition to the potential scRNA technology to decipher cellular heterogenicity, this technology can potentially identify important mediators of development. scRNA sequencing of striata from mice with a conditional knockout of Foxp1, a transcription factor strongly linked to autism and intellectual disability, show evidence for dysregulation of striatal projection neuron subtypes in FOXP1 deletion and identified Foxp1-regulated targets which overlapped with ASD-associated genes [134]. Cerebral cortical depletion of Foxp2, a gene linked to ASD and ADHD, was shown to alter D1R expressing cells subtype composition at postnatal day 7[135]. A layer V specific knockout of Tbr1, a transcription factor implicated in autism, at a stage similar to human mid-pregnancy fetal development, followed by scRNA sequencing at PD5 found evidence for altered expression of a subset of genes that control synaptogenesis, synaptic maturation, and microtubule assembly in addition to regulation of other ASD genes [136].

SINGLE CELL TRANSCRIPTOMICS ALLOW FOR UNDERSTANDING CELL-TYPE SPECIFIC ENVIRONMENTAL EXPOSURE EFFECTS

scRNA seq has been used to understand cell-specific effects of environmental exposure on development. [137] Using embryonic body model to study the effects of nicotine on embryonic development, it was demonstrated that nicotine caused lineage specific effects and dysregulated cell-to-cell communication of EBs, in addition to eliciting cell type specific responses in differentiated EBs. Cell type specific effects of prenatal tamoxifen exposure on the developing brain were shown using scRNA. Tamoxifen caused biphasic effects on cortical neurogenesis, suppressing cell cycle progression and inhibiting neural progenitor cells proliferation; and increasing the number of subventricular zone progenitors expressing neuronal genes, leading to precocious neuronal differentiation[138]. Single-cell RNA sequencing analysis of immune cells

harvested from Influenza A Virus-infected lungs of mice, pinpointed specific players mediating the prenatal arsenic exposure induced-enhanced inflammatory response[139]. Cell specific transcriptomic responses to prenatal alcohol exposure have not been widely investigated, [140] but found gene expression changes-specifically upregulation of Kcnn2- in prenatally alcohol exposed mice in a subset of motor neurons(M1) in the mature brain.

MODULARITY IN GENE EXPRESSION AND WEIGHTED CORRELATION NETWORK ANALYSIS (WGCNA)

Construction of weighted gene co-expression networks allows for the identification of hub genes modulating key processes in normal embryonic development [141, 142] and in neurodevelopmental disorders[143]. Transcriptomic analysis and WGCNA construction in post-mortem brain tissue from autism cases resulted in the identification of two gene co-expression modules associated with autism, one is enriched for astrocyte markers and activated microglia, highlighting a link between inflammation and autism spectrum disorders [144]. In this study [145], the authors identified in cerebral cortex, differentially methylated regions in genes associated with neurodevelopmental disorders (autism spectrum disorder, Rett syndrome, and Dup15q syndrome). They applied to WGCNA methodology to single cell RNA seq data [146] and published RNA-seq datasets [147] and showed that identified modules enriched in Dup15q DMR genes and RTT DMR genes, were also enriched for immune function, cell activation, and immune effector process microglial developmental time. WGCNA analysis on the transcriptomic signature from ethanol-treated differentiated and undifferentiated human embryonic stem cells identified altered modules associated with metabolic processes, oxidative stress, and neuronal properties of stem cells[148]. In another study, dysregulated modules were associated with autophagy, neuroactive ligand-receptor interaction, Jak-Stat pathway and Calcium signaling

pathways [149]. In an animal model of prenatal alcohol exposure and maternal separation [150], the authors identified a gene co-expression module in adult hippocampus, associated with prenatal alcohol exposure and the development of anxiety-like behavior, analysis of gene members of this module implicates altered RNA processing and transcriptional regulation.

1.8 CONCLUSION AND OVERALL DISSERTATION QUESTION:

Prenatal alcohol exposure (PAE) causes cell-type specific disruption of cell proliferation and subsequent layer-specific thickness changes in the ventricular and subventricular zone differently (summarized in Figure 1-1, and reviewed in [151]) In the first part (section 2) , I utilize a neural stem cell culture model-modelling mid-1st to 2nd human trimester, to study the effects of alcohol (ethanol) exposure on a novel pseudogene family member of the pluripotency factor Oct4/Pou5f1. This pseudogene locus, mOct4pg9, is transcribed into a long non-coding RNA, and I study how this lncRNA mediates the effect of alcohol exposure on neural stem cells proliferation and maturation. In the second part (section 3) , I utilize a mouse model of single binge dose prenatal alcohol exposure, at a time point equivalent to human mid-1st to 2nd trimester, and study cell-type specific effects of alcohol exposure on fetal cortex development. In this section, I identify fetal sex-specific cell-type specific prenatal alcohol effects in the fetal developing cortex.

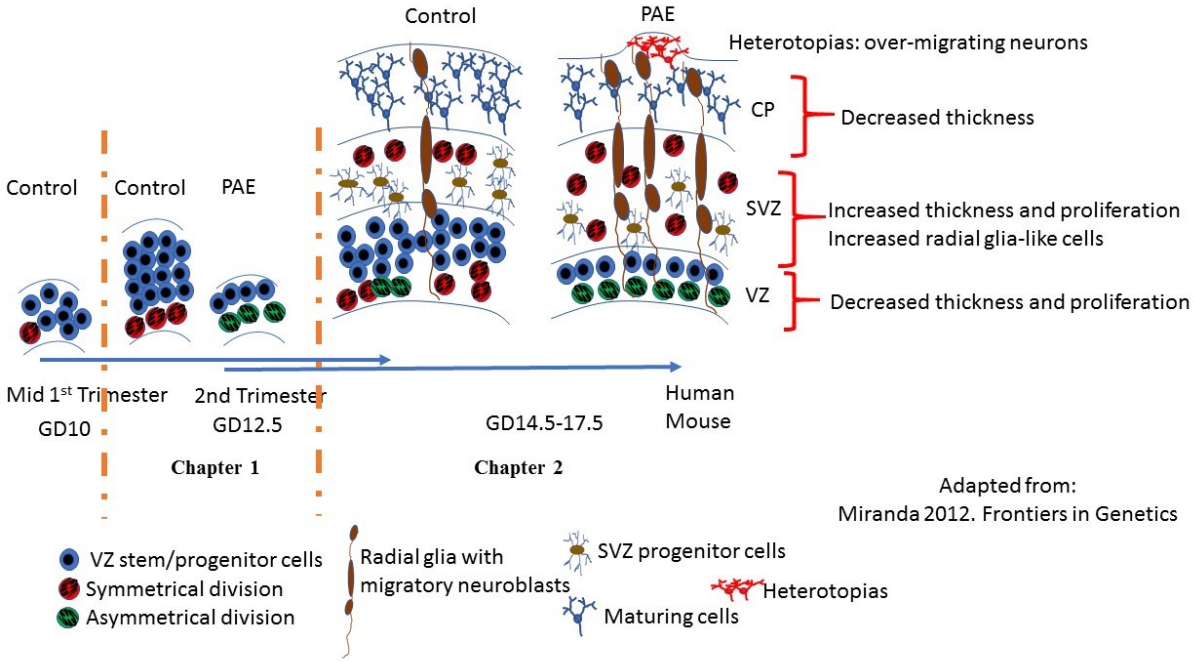


Figure 1-1 Effects of prenatal alcohol exposure on cortical development

2 A NOVEL OCT4/POU5F1-LIKE NON-CODING RNA CONTROLS NEURAL MATURATION AND MEDIATES DEVELOPMENTAL EFFECTS OF ETHANOL.*

2.1 INTRODUCTION

Prenatal alcohol exposure (PAE) can result in microcephaly [72] and persistent brain growth deficits [73], and is the leading cause of intellectual and other neurobehavioral disabilities [74], collectively termed Fetal Alcohol Spectrum Disorders (FASD), in the US [75] and worldwide [152]. The prevalence of PAE is high, particularly during the first and second trimesters [153-155], encompassing the peak period of human fetal neurogenesis in key brain regions like the cerebral cortex [156, 157]. Earlier studies in rodent models showed that PAE during this equivalent developmental period, resulted in microencephaly [158] and decreased neuron number [159], and that this loss of neurons was due to loss of proliferative capacity in the germinal ventricular zone [160, 161], i.e., the neural stem cell (NSC) niche. To assess the direct effects of ethanol on fetal neural stem and progenitor cells, we implemented an *ex vivo* neurosphere culture model of the micro-dissected mouse and rat fetal telencephalic neuroepithelium, which models the NSC niche. We reported that ethanol exposure did not kill NSCs, but rather, resulted in increased proliferation, the loss of cell-surface markers for neural stem cell identity and aberrant premature maturation [86, 87, 162-164]. Moreover, these effects of ethanol were explained in part, by alterations in epigenetic mechanisms, including chromatin

* Reprinted with permission from [228].

modeling [106, 165] and microRNA (miRNA)-mediated translation regulation. Manipulation of the miRNAs resulted in microencephaly [166] and premature maturation of NSCs [164, 167, 168]. These data indicate that ethanol exposure is not cytotoxic to NSCs, but instead results in premature maturation of NSCs, leading to their depletion.

However, apart from producing more mature, lineage-specific daughter cells, NSCs also engage in a process of self-renewal to replenish their numbers. Oct4/Pou5f1 (Octamer binding protein-4/POU domain, class 5, transcription factor 1) has been identified in mammals as a core member of a small network of transcription factors that control the pluripotency and self-renewal capacity of mammalian stem cells [169]. Not only is the formation of pluripotent stem cells in the mammalian embryo dependent on the expression of Oct4/Pou5f1 [170], but its enforced expression can cause somatic cells to revert back to a stem-cell like state [171], suggesting that this transcription factor is essential to maintaining stem cell identity. Moreover, overexpression of this transcription factor has been shown to result in the expansion of the neural plate, at the expense of ectodermal tissue, and to inhibit neurogenesis [172], while its loss, accelerates neural differentiation [173], suggesting that Oct4/Pou5f1facilitates the expansion of neural stem cell numbers and prevents maturation. Therefore, a possible loss of Oct4/Pou5f1 following PAE may explain the loss of NSC self-renewal capacity. This hypothesis is supported by our previous data showing that ethanol exposure resulted in a loss of the stem cell-enriched ATP-dependent transporter, ABCG2 [86], which is important for stem cell renewal [174] and induced by Oct4/Pou5f1 [175].

Assessing the effects of PAE on Oct4/Pou5f1-mediated self-renewal and maturation behavior of neural stem cells is complicated by the fact the Oct4/Pou5f1 gene locus has been duplicated numerous times during mammalian evolution. These gene duplication events have resulted in the

emergence of a number of 'pseudogenes' that appear to have lost open reading frames for protein translation [176, 177], but are transcribed to produce non-protein-coding RNAs (ncRNAs, [178]). Such pseudogene-encoded ncRNAs have been shown to have the capacity to sequester microRNAs (reviewed in [179]), and are therefore likely to interfere with the translational regulation of their parent protein-coding gene's expression [180]. However, these ncRNAs may have alternate roles as sources of small interfering RNAs, microRNAs [181], or as epigenetic regulators of gene transcription [182]. For Oct4/Pou5f1, three human Oct4 pseudogene-derived transcripts were predicted to have microRNA sequestration activity, were found to be expressed in neural stem-like cells at high levels, along with Oct4/Pou5f1, and their expression decreased with differentiation, along with Oct4/Pou5f1 [176], suggesting a complementary role for these pseudogenes in maintaining stem cell identity. Consistent with this hypothesis of functional facilitation, a human Oct4 pseudogene was shown to upregulate Oct4/Pou5f1 in endometrial and hepatocellular carcinomas, by sequestering the microRNA miR-145 [183, 184]. In contrast, a nuclear X-linked Oct4 pseudogene was shown to silence the Oct4/Pou5f1 gene through the interaction with Histone-lysine N-methyltransferase, to facilitate trimethylation of histone H3 (H3K9me3) and Heterochromatin protein 1 (HP1)α binding to the Oct4/Pou5f1 promoter [185]. Additionally, an antisense RNA transcript originating from an Oct4 pseudogene was shown to interact with the chromatin remodeling protein EZH2, to suppress the transcription of Oct4/Pou5f1 [182]. These data suggest that nuclear and cytoplasmic Oct4/Pou5f1-related ncRNAs titrate Oct4/Pou5f1 protein levels, and consequently, control stem cell identity.

In this study, we identified Oct4 pseudogene transcripts in mouse neural stem cells that are homologous to human Oct4 pseudogenes. We report that the expression of one pseudogene, encoded on mouse chromosome 9 (*m*Oct4pg9), is induced following ethanol exposure, and that

its overexpression results in suppression of Oct4/Pou5f1. The expression of ncRNAs from this pseudogene locus may mediate the loss of neural-stem cell renewal capacity and aberrant maturation. To our knowledge, this is the first study to link the activity of this ncRNA family - pseudogenes- to the neural growth deficits associated with PAE.

2.2 METHODS

NEURAL STEM CELL CULTURE MODEL AND ETHANOL TREATMENT PARADIGM

Neuroepithelial cells from the dorsal telencephalic vesicle, corresponding regional precursor of the murine iso-cortex, were micro-dissected from gestational day (GD) 12.5 C57bl/6J mouse fetuses obtained from timed pregnancies, as previously published [86, 162, 164, 167, 168]. All procedures were as approved by the Texas A&M University Committee on Animal Care (IACUC). Cells were maintained as non-adherent cultures in mitogenic serum-free media, resulting in the formation of neurospheres, an *ex vivo* model of the fetal neural stem cell niche [84]. For ethanol treatment experiments neurospheres were randomly assigned to control (0 mg/dL) or ethanol treatment groups (60 mg/dL, 120mg/dL or 320 mg/dL (13, 26 and 70mM, respectively)), capped tightly with phenolic caps, and sealed with parafilm to limit the loss of ethanol. This range of exposures was designed to bracket the range of consumption in human populations, from casual consumption to levels observed in persons with alcohol use disorders [186]. Cultures were treated with ethanol for 5 days to mimic exposure through the peak period of cortical neurogenesis, with culture media replaced after 3 days. RNA from neurospheres was isolated using the miRNeasy Mini Kit (Qiagen, Germantown, MD). cDNA synthesis was performed using qScript™ cDNA SuperMix (Qiagen). We assessed gene expression for key stemness and differentiation markers using primers we previously validated [168].

NEURAL STEM CELLS DIFFERENTIATION PARADIGM.

We previously showed that neurospheres cultured on a laminin substrate, and maintained in the absence of bFGF, became adherent. Over a 72-hour window this process results in the radial like migration of neuroblasts away from the parent neurosphere and the emergence of neuronal markers [87]. Neurosphere cultures were suspended in media lacking bFGF, and also supplemented with brain-derived neurotrophic factor (BDNF, 25 ng/ml, PeproTech Inc., Rocky Hill, NJ) to further stimulate neuronal maturation, and plated on laminin-coated (1.2 mg/ml, Thermo Fisher Scientific, Waltham, MA) 12-well plates. Culture medium was replaced every other day. Cells were harvested after either 24 or 72 hours after initiation of the mitogen-withdrawal/BDNF supplementation-induced differentiation program to capture early and later stages of maturation.

NUCLEAR/CYTOPLASMIC FRACTIONATION, RNA ISOLATION &DNASE TREATMENT:

Neurospheres were fractioned into nuclear and cytoplasmic compartments using NE-PER Nuclear and Cytoplasmic Extraction Kit (Thermo Fisher Scientific). RNA was isolated from nuclear and cytoplasmic compartments using a Trizol based method (Qiagen). To remove any genomic DNA contamination, isolated RNA was treated with DNase for 1hour at 37 C using 6 Units of Ambion TURBO DNase (Thermo Fisher Scientific). RNA was then purified from the DNase treatment reagents using RNA Clean & Concentrator-5 (Zymo Research, Irvine, CA).

IDENTIFICATION OF MOUSE OCT4 PSEUDOGENES AND SURVEY OF EXPRESSION AMONG DIFFERENT MOUSE TISSUES

Discontiguous megablast (NCBI, [187]) was used to identify 6 genomic loci with moderate to high homology (coverage from 46% to 100% and identity from 71% to 100%) to mouse Oct4/Pou5f1 mRNA (Mus musculus transcript variant 1, RefSeq NM_013633; GRCm38/mm10,Table 2-1). Based on discriminability from the parent Oct4/Pou5f1 mRNA

transcript, four presumptive Oct4 pseudogenes on mouse (*m*) chromosomes 9 (*m*Oct4pg9), 6 (*m*Oct4pg6), 14 (*m*OCT4pg14) and X (*m*Oct4pgX) were selected for further investigation. Primer pairs were designed for each of these selected pseudogenes and the parent Oct4/Pou5f1 using the NCBI primer Blast tool (Appendix 1 a). cDNA synthesized from DNase treated RNA isolated from embryonic stem cells (obtained from Texas A&M Institute for Genomic Medicine) and neural stem cells was used to determine pseudogene expression. Primers were tested using ViiA^{TM} 7 Real-Time PCR (RT-PCR, Thermo Fisher Scientific) system and SYBR Green Reagents (Quanta). Melting and dissociation curves were analyzed for the presence of single amplicons, the PCR products were size-fractionated on 1.5% agarose gel to assess amplicon size, and finally, the amplicons were sequenced (Gene Technology Lab, Texas A&M University) to confirm their identity. Cycle thresholds (CTs) were computed for RNA transcripts and normalized to the CT for beta-actin using the ΔΔCT method. Sequence data for the *m*Oct4pg9 amplicon has been deposited in GenBank (accession number, *BankIt2349944 Oct4pg9 MT551012*).

To study the expression of Oct4 pseudogenes in mouse embryonic tissues at different embryonic stages, we extracted RNA sequence data from the ENCODE project, GSE36025 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE36025), for select tissues from C57BL/6NJ mice. We applied a pipeline for identification of novel un-annotated RNA transcripts in which reads were mapped to the mouse genome (mm10) using BOWTIE [188], followed by transcript assembly using Cufflinks [189]. The assembled transcripts were then assessed for alignment with genomic sequences for Oct4 pseudogenes.

USING SINGLE CELL SEQUENCING TO IDENTIFY GENE EXPRESSION PROFILES IN OCT4 POSITIVE AND MOCT4PG9 POSITIVE CELLS IN NEUROSPHERES

We profiled neurosphere single cell RNA expression data ([168], data deposited at NCBI GEO, Accession #GSE155880) for the cells expressing *m*Oct4pg9. Since *m*Oct4pg9 is not currently an annotated Refseq RNA, we profiled the sorted BAM file resulting from the Cell Ranger pipeline and extracted the unique molecular identifiers (UMI) of the cells which contained transcripts overlapping with the genomic coordinates of *m*Oct4pg9. Those UMIs were then used to extract the gene expression matrix of the *m*Oct4pg9-positive cells. Additionally, we identified the cells which were Oct4/Pou5f1 positive to assess the overlap between Oct4/Pou5f1-positive cells and *m*Oct4pg9-positive cells. To identify the candidate RNA transcripts that discriminated between *m*Oct4pg9^{+ve} and Oct4/Pou5f1^{+ve} cells, and all the cells, we ranked expressed genes in descending order of expression (log normalized expression of each gene in each of the three groups). and extracted the top 2000 most highly expressed genes (~10% of the expressed genes) from each list, then determined the unique expressed genes for each of the *m*Oct4pg9^{+ve} and Oct4/Pou5f1^{+ve} cells eliminating the common genes between the three lists. We then performed IPA analysis on the unique expressed genes in the *m*Oct4pg9^{+ve} and Oct4/Pou5f1^{+ve} cells using the 2000 most highly expressed genes in each subset of cells as a reference dataset.

IDENTIFICATION OF HUMAN OCT4 PSEUDOGENES AND SURVEY OF EXPRESSION IN DEVELOPING TISSUES

We used NCBI's Blast tool to identify the human Oct4 pseudogenes homologues to mouse *m*Oct4pg9 (Table 2-2). We also used the GSE103715 dataset [190] to study the expression level of Oct4 and Oct4 pseudogenes during human embryonic stem cell differentiation neuronal cells. Additionally, we used ARCHS^4 (Massive Mining of Publicly Available RNA-seq Data from

Human and Mouse, [191], https://amp.pharm.mssm.edu/archs4/index.html), to survey the association of human Oct4 pseudogenes with phenotypes that overlapped with PAE-associated phenotypes.

MOCT4PG9 SIRNA MEDIATED KNOCKDOWN:

100 nM of Smart pool siRNA targeted against *m*Oct4pg9 or siGENOME Non-Targeting Control siRNA Pool #2 (Dharmacon, Lafayette, CO) (Sequences in Appendix 1 b) was transfected into the neural stem cells using lipofectamine RNAiMAX Transfection reagent (Thermo Fisher Scientific), after 48 hours the cells were harvested, and RNA or protein were isolated (see below). On average, siRNA-mediated knockdown resulted in a 31.67±8.79% decrease in mOct4pg9 RNA expression ($t_{(5)}$=3.599, p=0.016) relative to control.

MOCT4PG9 CRISPR-MEDIATED KNOCKDOWN:

We designed three single guide RNAs against non-overlapping regions of the *m*Oct4pg9 coding region (Sequences in Appendix 1 c). We assessed the efficiency and specificity of each of the sgRNAs by measuring both *m*Oct4pg9 lncRNA and its host gene, Sik3, using RT-PCR. We selected the guide RNA with the highest *m*Oct4pg9 knockdown and the least effect on Sik3 for CRISPR mediated knockdown in our experiments. sgRNA-2 met our criteria of the highest knockdown of *m*Oct4pg9 with the least effect on Sik3 mRNA (Appendix 1 d, average knockdown of *m*Oct4pg9 RNA, 44.13±7.59% ($t_{(22)}$=5.812,p=0.000008), compared to average knockdown of Sik3 host-gene mRNA, 21.50±9.51% ($t_{(22)}$=2.26 p=0.034)). Neurospheres were dissociated into single cells, filtered through 100 µm filter, diluted to $2X10^5$ cells/ml of media and 500 µl of the cell suspension was plated into each well of a 6-well plate. 30 pmol of the guide RNA (Synthego, Redwood City, CA) and 0.5 µl of Geneart CRISPR nuclease mRNA (Thermo Fisher Scientific) was transfected into each well using lipofectamine RNAiMAX (Thermo Fisher Scientific). After

48 hours, cells were harvested and RNA was isolated using miRNeasy micro kit (Qiagen,). DNase treatment, cDNA synthesis and RT PCR were performed as described above.

MOCT4PG9 PLASMID-MEDIATED OVEREXPRESSION

We designed a mammalian gene expression AAV vector expressing *m*Oct4pg9 under the control of the human eukaryotic translation elongation factor 1 α1 short form (EFS) promoter (Vector Builder, Chicago, IL). For the control expression plasmid, the *m*Oct4pg9 coding sequence was replaced with the GFP coding sequence lacking a translation start codon (Appendix 1 e). We electroporated 800 ng of either *m*Oct4pg9 overexpressing plasmid or the control plasmid per million cells and plated $1X10^6$ cells per well of a 6-well plate. Briefly, neurosphere cultures were dispersed into a single-cell suspension and filtered through 100 µm filter. The recovered cells were concentrated by centrifugation and resuspended at a concentration of $1X10^6$ cells/10µl PBS. Plasmids were transfected into cells by electroporation, using the Neon™ Transfection System Kit (Voltage: 1200 V, Pulse duration: 20 ms, number of pulses: 2, Thermo Fisher Scientific). Electroporated cells were diluted into 2 ml of prewarmed media and incubated for 72 hours, followed by cell harvesting for RNA and protein isolation.

ARGONAUTE (AGO) PROTEINS IMMUNOPRECIPITATION

Neurospheres were fixed with 0.2% paraformaldehyde (Electron Microscopy Sciences, Hatfield, PA) in 1X DPBS (Thermo Fisher Scientific) for 10 minutes, following which, the cross-linking reaction was stopped with 2M glycine (Sigma-Aldrich, St. Louis, MO) which was added to produce a final concentration of 0.125 M. Nuclear and cytoplasmic fractionation was performed, and the isolated cell fractions were incubated with 8 µl of either anti-Ago2 antibody (Cat#ab32381, Abcam, Cambridge, MA), anti-pan Ago (Cat.#MABE56, Sigma-Aldrich, St. Louis, MO) or rabbit IgG (Thermo Fisher Scientific) overnight, followed by incubation with 100 µl of Dynabeads

Protein G (Thermo Fisher Scientific) for 1 hour at 4°C. The Dynabeads were recovered using a magnetic stand, washed twice with 200 µl 0.05% Tween 20 (Sigma Aldrich) in 1X DPBS and once in 200 µl 0.02 % Tween 20 in 1X DPBS, and then resuspended in 100 µl 0.02% Tween 20 in 1X DPBS. Dynabead-linked samples were then digested with proteinase K (Thermo Fisher Scientific) and then un-crosslinked at 65°C for 4 hours with periodic agitation. Total RNA was then extracted with the TRIzol® Reagent (Qiagen), according to the manufacturer's instructions.

BIOTINYLATED OLIGONUCLEOTIDE MEDIATED CAPTURE OF MOCT4PG9

We designed nine oligonucleotide probes complementary to regions on the *m*Oct4pg9 ncRNA transcript which we assessed for chromatin accessibility using Capture Hybridization Analysis of RNA Targets (CHART, [192]). Briefly, 1 µg of neurosphere RNA was mixed with RNAse H (Thermo Fisher Scientific), and one of the antisense oligonucleotide probes (each oligonucleotide tested separately), followed by a one-hour incubation at 37°C. After the incubation, quenching buffer was added to stop the action of RNAse H, RNA was purified, and cDNA was synthesized and assessed for *m*Oct4pg9 ncRNA transcript expression by RT-PCR. Since RNAse-H specifically digests RNA: DNA hybrids, lower levels of *m*Oct4pg9 ncRNA indicate higher accessibility of that region of the ncRNA to the oligonucleotide probe. We selected 3 of oligonucleotide probes (sequences presented in Appendix 1 f) which showed the highest accessibility to *m*Oct4pg9 ncRNA as capture reagents. We prepared hybridization buffer using each of the 3 selected biotinylated oligonucleotide capture probes then we combined the buffer with neurosphere cell lysates and incubated overnight at room temperature with rotation. The probe-lncRNA hybrids were captured using Invitrogen Dynabeads™ MyOne™ Streptavidin C1 beads (Thermo Fisher Scientific), washed according to the manufacturer's protocol, and then hybrids were eluted and RNA isolated. To identify miRNAs associated with *m*Oct4pg9 ncRNA,

capture samples with the highest enrichment of *m*Oct4pg9 ncRNA over mismatch oligonucleotide mediated-capture were used for miRNA profiling.

PROFILING CHANGES IN MIRNA EXPRESSION AFTER MOCT4PG9 MANIPULATION AND MIRNA ASSOCIATION TO MOCT4PG9 NCRNA

MicroRNA PCR Human Panels I & II V4.M (Exiqon/Qiagen) were used to assess the changes in miRNA expression after *m*Oct4pg9 overexpression or knockdown, and to assess the association of miRNAs with *m*Oct4pg9 lncRNA. Briefly cDNA was prepared from purified RNA using Universal cDNA Synthesis Kit II using 40 ng of RNA as input. Master mix was prepared using cDNA(diluted to 5 ng/ul) and ExiLENT SYBR® Green Master Mix Kit and loaded onto PCR panels. RT-PCR was performed using the ViiA^{TM} 7 Real-Time PCR system (Thermo Fisher) using manufacturers recommended cycle parameters. The CT for each miRNA was normalized to the geometric mean of CTs for a group of miRNAs that were stably expressed in all samples and unaffected by *m*Oct4pg9 manipulation (for Oct4pg9 overexpression miRNA profiling : geometric mean of hsa-miR-21-5p, hsa-miR-362-5p, U6 snRNA, hsa-miR-361-5p, hsa-miR-20a-5p, hsa-miR-296-5p, hsa-miR-93-5p , for Oct4pg9 knockdown miRNA profiling : geometric mean of hsa-miR-412-3p, hsa-miR-135b-5p, hsa-miR-98-5p, hsa-miR-153-3p, hsa-miR-20a-5p were used for normalization). To assess the enrichment of miRNAs associated with *m*Oct4pg9 ncRNA, we compared anti-*m*Oct4pg9 oligonucleotide probe-captured miRNAs to miRNAs captured with mismatched oligonucleotide probes. For each miRNA we subtracted the CT of that miRNA from the mean CT of all amplified miRNAs in this sample ($\Delta CT = CT_{miRNA} - CT_{mean}$) captured by the anti-*m*Oct4pg9 oligonucleotide probe, and subtracted from that miRNAs ΔCT values for the mismatched oligonucleotide-mediated RNA capture control probe, to account for non-probe specific RNA binding, and the data were expressed as 2^-ΔΔCT to denote fold-

enrichment of association between miRNA and *m*Oct4pg9 ncRNA compared to mismatched capture probes. A miRNA was identified as 'associated with *m*Oct4pg9 ncRNA' if the fold-enrichment was higher in samples obtained from each of three unique *m*Oct4pg9 ncRNA-specific capture probes compared to the mismatched capture probes.

OCT4/POU5F1 3'-UTR TARGET ASSAY

The 3'- untranslated region (UTR) sequence for Mouse Oct4/Pou5f1 (NM_013633.2) was cloned into luciferase reporter vector (MmiT030962-MT06, GeneCopoeia, Inc.). We co-electroporated, using the Neon™ Transfection System according to manufacturer's protocol (as outlined above), 500 ng of either control or *m*Oct4pg9 overexpression plasmid, and 500 ng of either miRNA Target clone control vector (CmiT000001-MT06, Genecopoeia) or mouse Oct4/Pou5F1 3'-UTR expression clone per 1 million cells. The cells were then plated in 6-well plates, and harvested and lysed after 72 hours. Lysates (20 µl) were assessed for luciferase activity (Luc-Pair Duo-Luciferase Assay Kit 2.0, Genecopoeia), and quantified using the Synergy-2 multi-mode plate reader (BioTek, Winooski, VT). Firefly luciferase intensity from each sample was normalized to the internal control, Renilla luciferase, and all the ratios were normalized to the average of the control samples for the experiment. We performed three independent electroporation experiments and the normalized ratios were analyzed using two-way ANOVA, followed by post-hoc fisher's LSD.

METHYLATED AND HYDROXYMETHYLATED DNA ASSAY.

DNA was isolated from *m*Octpg9-overexpressing and control neurospheres, 72 hours after electroporation, using Nucleospin RNA/DNA/Protein kit (Macherey-Nagel, Bethlehem, PA), and 150 ng of DNA was used for the methylated and hydroxymethylated DNA assay according to the assay manufacturer's instructions (Abcam, Cambridge, MA). Briefly DNA sample in binding buffer was incubated for 90 minutes in high-affinity DNA-binding wells, washed, then incubated

with capture antibodies against methylated or hydroxymethylated DNA. Antibody binding was visualized by a colorimetric reaction and quantified spectrophotometrically by absorbance at 450 nm using Infinite M200 microplate reader (Tecan, Männedorf, Switzerland). The measured absorbance from the DNA samples along with the absorbance from a positive control was used to calculate the percentage of methylated and hydroxymethylated DNA.

PROTEIN ISOLATION AND WESTERN IMMUNOBLOT ASSAYS

Cultured neurospheres were collected by centrifugation and lysed in 1X RIPA lysis buffer (EMD Millipore, Burlington, MA) supplemented with Halt protease inhibitor cocktail (Thermo Fisher Scientific). Protein concentration was determined using Pierce BCA protein assay kit (Thermo Fisher Scientific). Cleared cell lysates were size fractionated on a 4-12% Bis-Tris gradient acrylamide gel (Thermo Fisher Scientific, Rockford, IL), at 150 volts for 60 min, and transferred to a PVDF membrane using the iBlot dry transfer system and manufacturer protocols (Thermo Fisher Scientific). The membrane was then blocked with either 5% nonfat dry milk and 2.5% goat serum in Tris-buffered saline containing Tween 20 (TTBS) or Odyssey® Blocking Buffer (LI-COR Biotechnology. Lincoln, Nebraska) in TBS for 1 hour, incubated overnight with primary antibody, washed and incubated with secondary antibody for 1 hour at room temperature (list of antibodies and dilutions are presented in Appendix 1 g), and subsequently developed either using PerkinElmer Western Lightning Plus Chemi-ECL (PerkinElmer, Waltham, MA) and visualized using a CCD camera (Fluorchem Q, Alpha Innotech, San Leandro, CA) or direct visualization of near infrared-conjugated secondary antibodies using Odyssey® CLx Imaging System (LI-COR Biotechnology. Lincoln, Nebraska).

LACTATE DEHYDROGENASE ASSAY:

After *m*Oct4pg9 manipulation (overexpression or knockdown) cells were separated from culture medium by centrifugation, and 50 µl of culture-conditioned medium was assessed in duplicates for lactate dehydrogenase (LDH, Pierce LDH cytotoxicity assay kit, Thermo Fisher Scientific) activity as a measure of cytotoxicity and cell death [167]. The colorimetric reaction was quantified spectrophotometrically at 490 nm and normalized to absorbance at 680 nm for background correction using the Infinite M200 microplate reader (Tecan).

CLICK IT EDU ASSAY

The rate of DNA synthesis was assayed using Click-iT™ EdU Alexa Fluor™ 488 Flow Cytometry Assay Kit (Thermo Fisher Scientific,). Briefly, 10 mM EdU was added to the culture media, to a final concentration of 10 µM, for one hour. Cells were then fixed, permeabilized, and incorporated EdU was coupled to Alexa Fluor® 488-picolyl azide and detected using the Gallios 2/5/3 Flow Cytometer (Beckman Coulter, Brea, CA). Data were analyzed using Kaluza software (Beckman Coulter).

STATISTICAL ANALYSES:

Data were analyzed in GraphPad Prism software, version 8.0.0 (GraphPad Software) and R software. Unpaired two-tailed student t-tests, one-way and two-way ANOVAs were computed as appropriate and an uncorrected Fisher's LSD was used for post-hoc analyses, and in all cases, p-values less than 0.05 were considered significant. Data were expressed as individual data points from independent cell cultures (with the exception of neurosphere size analyses, where individual points represent individual neurospheres pooled from independent cell cultures). Individual bars represent group means and error bars represent standard error of the mean

(S.E.M). Effect-sizes calculations (with Hedges' *g* correction for small sample sizes, [193]) were performed in miRNA expression experiments (effsize R Package).

2.3 RESULTS

DUPLICATION OF THE OCT4/POU5F1 GENE IN MOUSE GIVES RISE TO ALTERNATE RNA TRANSCRIPTS THAT ARE DIFFERENTIALLY EXPRESSED IN DEVELOPING TISSUES.

To identify presumptive murine loci for processed pseudogenes for Oct4/Pou5f1, we mapped the full-length Oct4/Pou5f1 mRNA transcript to the mouse genome (GRCm38/mm10 assembly). We found sequence alignment with significant homology to Oct4/Pou5f1 in 6 different regions of the mouse genome. None of these regions of duplication appear to have retained the open reading frame (ORF) for Oct4/Pou5f1, and therefore, we annotated these in our study as presumptive murine Oct4 pseudogene (*m*Oct4pg) loci and labeled them based on their chromosomal location (Table 2-1). We assessed the expression of both Oct4/Pou5f1 and these presumptive murine Oct4 pseudogenes in mouse embryonic and neural stem cells. Both C57Bl/6J embryonic stem cells and neural stem cells expressed Oct4/Pou5f1 mRNA. The embryonic stem cells expressed RNA transcripts unique to *m*Oct4pg6, *m*Oct4pg9 and *m*Oct4pg14 loci, whereas, mouse C57Bl/6J-derived neural stem cells only expressed transcripts from the *m*Oct4pg9 locus. Other presumptive Oct4 pseudogene loci did not give rise to detectable RNA transcripts in these two cell types.

We next analyzed murine C57BL/6NJ RNA-seq data available in the Gene Expression Omnibus [194, 195] for RNA transcripts from the developing brain that aligned with the genomic sequence of the predicted pseudogenes, including 2000 nucleotides of flanking 5' and 3' regions, and assessed expression (Fragments per kilobase of exon per million reads mapped, FPKM) differences. No transcripts were detected that aligned to the genomic locations of *m*Oct4pg3,

*m*Oct4pg6 or *m*Oct4pg14 pseudogenes in gestational day (GD) 11.5 brain tissues. However, brain RNA transcripts were detected that aligned to *m*Oct4pg9 (mouse Chromosome 9 (+ 46,160,992-46,162,166)) from GD11.5 onwards into the postnatal period (Figure 2-1b). Analysis of publicly available single-cell RNAseq data from the [196] showed that transcripts bracketing the region of *m*Oct4pg9 were present in astrocytes and oligodendrocytes, but not neurons from adult C57BL/6JN mice (Figure 2-1c), whereas mRNA for Oct4/Pou5f1 was not present in any of these adult neural cell subtypes.

MOCT4PG9 IS EXPRESSED IN NEURAL STEM CELLS AND IS DOWNREGULATED DURING NEURONAL DIFFERENTIATION

To assess the developmental stage-related expression of *m*Oct4pg9 and Oct4/Pou5f1, we examined the expression of these RNAs in NSCs cultured under mitogenic conditions, to favor self-renewal, and under growth factor-guided differentiation. *m*Oct4pg9 RNA is expressed in NSCs maintained in neurosphere cultures at an average 3.475-fold higher level than Oct4/Pou5f1 mRNA (+/- S.E.M = 0.3231). Within NSCs, *m*Oct4pg9 RNA transcripts are preferentially localized, at an average 10-fold higher level, to the nucleus relative to cytoplasm ($t_{(4)}$ =3.263, p=0.031, **Δ**Ct of Oct4pg9 in the nucleus =3.707+/- 0.47 , **Δ**Ct of Oct4pg9 in the cytoplasm = 6.954+/- 0.057). To assess the expression of *m*Oct4pg9 during neural differentiation, NSCs in neurosphere cultures were differentiated over a period of 24 to 72 hours, in the presence of laminin, in a mitogen-withdrawal, BDNF-supplementation paradigm. *m*Oct4pg9 RNA was significantly decreased in this differentiation paradigm ($F_{(2,13)}$ =21.47, p<0.0001), with a 50% decrease observed within 24 hours. In contrast, Oct4/Pou5f1 mRNA expression was not significantly altered ($F_{(2,14)}$ =2.151, not significant, *ns*), suggesting that a reserve capacity for NSC self-renewal is maintained during growth factor-guided differentiation (Figure 2-1d).

MOUSE OCT4PG9 IS HOMOLOGOUS TO TWO DEVELOPMENTALLY EXPRESSED HUMAN OCT4 PSEUDOGENES.

*m*Oct4pg9 shares significant sequence homology with three human Oct4 pseudogenes, POU5F1P3, POU5F1P4 and POU5F1P5 (Table 2-2,Appendix 1 h). Analysis of published RNAseq data from transcriptome profiling of human embryonic stem cells (hESCs) directed through neural differentiation course (GSE103715, [190]) showed that Oct4/Pou5f1 mRNAs is expressed at high levels in hESCs, but that expression decreased over the course of neural differentiation, POU5F1P3 and POU5F1P4 were expressed at lower levels than POU5F1 at the embryonic stem cell state , and their levels decreased upon differentiation, reaching undetectable levels of expression at an earlier differentiation state than Oct4/Pou5f1 (Figure 2-1e).

We utilized the Archs^4 web tool [191] to mine published RNA-seq data from human and mouse, to predict phenotypes associated human OCT4 pseudogenes in the Human Phenotype Ontology. Expression of human POU5F1P4 is predicted to be associated with failure to thrive in infancy (HP:0001531, z-score= 4.55), neonatal respiratory distress (HP:0002643, z-score=3.06), gait imbalance (HP:0002141, z-score=2.19), and abnormality of midbrain development (HP:0002418,z-score=2.1). POU5F1P5 expression is predicted to be associated with phenotypes involving 5th digit anomalies (HP:0009161, 0009376, 0004219, 0009237, 0004213 and 0006262, z-scores range 5.9-7.78), broad palm (HP:0001169, z-score=6.75), and tongue protrusion (HP:0010808, z-score=5.78).

ETHANOL CAUSES THE LOSS OF OCT4/POU5F1 PROTEIN, BUT INDUCES MOCT4PG9 IN NSCS

Since our previous data showed that ethanol exposure resulted in depletion of neural cells that expressed a variety of stem cell markers [86, 163], we assessed the effects of ethanol on

OCT4/POU5F1 protein compared to *m*Oct4pg9 RNA. Ethanol treatment for 5 days resulted in a significant, dose-related alteration in OCT4/POU5F1 protein expression in neurosphere cultures ($F_{(3,21)}$=5.719, p=0.005; overall R^2=0.45), and specifically, a significant decrease at the high dose of ethanol (320 mg/dl) compared to untreated controls (post-hoc Fisher LSD, p=0.032, Figure 2-2a, the complete, uncropped blot is presented in Appendix 1 i). In contrast, ethanol treatment resulted in a significant, dose related-increase in *m*Oct4pg9 RNA ($F_{(3,30)}$=4.403, p=0.011; overall R^2=0.31, Figure 2-2b).

SINGLE-CELL RNASEQ (SCRNA-SEQ) ANALYSIS SHOWS THAT MOCT4PG9 RNA AND OCT4/POU5F1 MRNA ARE EXPRESSED, EX VIVO, BY DIFFERENT NEURAL PROGENITOR POPULATIONS

Since Oct4/Pou5f1 and *m*Oct4pg9 were differentially regulated during NSC maturation and by ethanol exposure, we assessed their co-expression and coordinately expressed transcriptome using single cell RNA sequencing data from 11,200 neural progenitors obtained from neurosphere cultures of GD12.5 fetal mouse dorsal telencephalon. Our data show that *m*Oct4pg9 and Oct4/Pou5f1 RNA are expressed in 1% (112 cells) and 1.2% (136 cells) of assessed cells, respectively, but surprisingly, that their expression overlaps minimally (~2% of Oct/Pou5f1 mRNA expressing cells also express *m*Oct4pg9 RNA). To understand the phenotypes of *m*Oct4pg9^{+ve} and Oct4/Pou5f1^{+ve} cells relative to all cells, we ranked mRNA transcripts in each cell population from most to least abundant. then selected the 2000 most abundantly expressed genes in each of the three populations. We identified 52 genes that were uniquely expressed in Oct4/Pou5f1^{+ve} cells, 97 genes that were uniquely expressed by *m*Oct4pg9^{+ve} cells, as well as an additional 55 genes that were expressed in both Oct4/Pou5f1^{+ve} and *m*Oct4pg9^{+ve} cells, but not in other neural progenitors (Figure 2-2c). Pathway overrepresentation analysis showed that

Oct4/Pou5f1^{+} cells uniquely expressed genes related to Spermidine Biosynthesis, whereas *m*Oct4pg9^{+} cells uniquely expressed genes associated with Pyrimidine Deoxyribonucleotides *De Novo* Biosynthesis and Mitotic Roles of Polo-Like Kinase (Table 2-3). In contrast four interrelated and statistically significant pathways related to Mevalonate/Geranylgeranyldiphosphate-mediated posttranslational protein modifications and ATM (Ataxia Telangiectasia Mutated) signaling were enriched in both Oct4/Pou5f1^{+ve} and *m*Oct4pg9^{+ve} cells relative to all other neural progenitors.

MOCT4PG9 EXPRESSION INFLUENCES OCT4/POU5F1 AND REST PROTEINS.

Since our scRNAseq showed that *m*Oct4pg9 and Oct4/Pou5f1 transcripts were expressed in non-overlapping cohorts of cells, we hypothesized that *m*Oct4pg9 was a suppressor of Oct4/Pou5f1, i.e., that *m*Oct4pg9 overexpression would result in decreased expression of Oct4/Pou5f1 as well as REST (RE1 silencing transcription factor), the suppressor of neural differentiation [197, 198] that maintains self-renewal and levels of pluripotency factors like Oct4/Pou5f1 [199]. However, surprisingly and contrary to our hypothesis, *m*Oct4pg9 overexpression resulted in increased Oct4/Pou5f1 mRNA level ($t_{(14)}$=4.296, p=0.0007, Figure 2-2d) and protein-immunoreactivity, 72 hours after electroporation ($t_{(6)}$=4.039, p=0.0068, Figure 2-2e), with no change in either REST mRNA ($t_{(4)}$=0.1338, *ns*), or protein ($t_{(6)}$=0.6945, *ns,* Appendix 1 j). The siRNA-mediated *m*Oct4pg9 knockdown did not however, alter OCT4/POU5F1 mRNA($t_{(5)}$=1.769, *ns,* Figure 2-2d) or protein expression ($t_{(4)}$=1.329, p=0.255, Figure 2-2e), but did result in significant, 1.3-fold upregulation of Rest mRNA ($t_{(6)}$=4.87, p=0.0027), and a statistically significant 10% upregulation in REST protein ($t_{(4)}$=11.07, p=0.0004, Appendix 1 j). In contrast, ethanol treatment did result in a significant, dose-related increase in REST mRNA (One-way ANOVA, $F_{(3,15)}$=8.026, p=0.002), but no change in REST protein expression ($F_{(2, 6)}$=2.159, *ns*, Appendix 1 j).

MOCT4PG9 EFFECT ON GLOBAL DNA METHYLATION AND HYDROXYMETHYLATION.

To investigate potential effects of *m*Oct4pg9 ncRNA on DNA epigenetic modifications, we assessed global DNA methylation and hydroxymethylation, an intermediate step in DNA-demethylation and an important component of neural maturation [200], 72 hours after *m*Oct4pg9 overexpression. DNA methylation and hydroxymethylation were not altered by *m*Oct4pg9 overexpression ($t_{(7)}$=1.161, *ns*, and $t_{(9)}$=0.118, *ns*, respectively, Appendix 1 k).

MOCT4PG9 ASSOCIATES WITH NUCLEAR AND CYTOPLASMIC AGO2 AND WITH MIRNAS.

The 3' tail of *m*Oct4pg9 ncRNA shares 49.6% sequence homology with 3'UTR of Oct4/Pou5f1(Figure B-3 b). These data show that *m*Oct4pg9 associates with Ago2 in both the nucleus and cytoplasm of neural progenitor cells. To identify the miRNAs associated with *m*Oct4pg9 ncRNA, we designed and utilized biotinylated oligonucleotides complementary to regions in *m*Oct4pg9 ncRNA, to capture *m*Oct4pg9 ncRNA and its associated molecules. Profiling of the associated miRNA showed that *m*Oct4pg9 RNA was significantly and selectively associated with miRNAs miR-328-3p (MIMAT0000565) and miR-744-5p (MIMAT0004187), compared to the biotinylated capture control. None of the two miRNAs are predicted to target the Oct4 3'-UTR.

MOCT4PG9 CONTROLS OCT4/POU5F1 3' UTR-DEPENDENT TRANSLATION.

Since *m*Oct4pg9 associates with Ago-2 and miRNAs, and shares sequence homology with Oct4 3' UTR, we further predicted that *m*Oct4pg9 would facilitate translation of Oct4/Pou5f1 through sequestering miRNAs that repress Oct4/Pou5f1 translation. To test this predictison, we assessed luciferase activity in NSCs transfected with plasmids that contained firefly luciferase (*f*Luc) with the Oct4/Pou5f1 3'-UTR along with a plasmid to overexpress *m*Oct4pg9 ncRNA.

Overexpression of *m*Oct4pg9 lncRNA did not result in change in luciferase activity when co-transfected with the control plasmid, which did not contain a 3'-UTR. However, contrary to our prediction, *m*Oct4pg9 ncRNA overexpression resulted in a significant, 25% decrease in luciferase activity from a plasmid containing the Oct4/Pou5f1 3'-UTR downstream of *f*Luc (Two-Way ANOVA, interaction: $F_{(1,32)}=4.422$, $p=0.043$; post-hoc Fisher's LSD, $p<0.006$; Figure B-3 c). These data suggest that *m*Oct4pg9 ncRNA is a functional antagonist of Oct4/Pou5f1 translation, and is consistent with our scRNAseq data which showed that progenitor cells which express *m*Oct4pg9 ncRNA do not overlap with cells that express Oct4/Pou5f1 mRNA.

MOCT4PG9 OVEREXPRESSION RESULTED IN THE DYSREGULATION OF MIRNA EXPRESSION IN NEUROSPHERES.

Since mOct4pg9 is associated with Ago2 and miRNAs, we sought to identify the miRNAs regulated by mOct4pg9.We investigated the effects of mOct4pg9 manipulation (overexpression and knockdown) on the expression of miRNAs in neurospheres. CRISPR-mediated mOct4pg9 knockdown did not significantly affect the expression of any of the profiled miRNAs (Figure B 3d). However, plasmid-mediated overexpression of mOct4pg9 ncRNA resulted in the upregulation of two miRNAs, miR-760 (MIMAT0004957) & miR-98-5p (MIMAT0000096), and the downregulation of 16 miRNAs (Figure B 3e; Appendix 1 m). Interestingly, one downregulated miRNA, miR-328-3p, was also found to directly associate with mOct4pg9 ncRNA.

MOCT4PG9 INFLUENCES NEUROSPHERE SIZE.

A secondary measure of self-renewal and proliferative capabilities is assessment of neurosphere size, and may indicate a role for mOct4pg9 in proliferation that is not reflected in Nestin mRNA expression. Neurosphere cultures were transfected with either siRNA targeting mOct4pg9 or

plasmid containing the mOct4pg9 expression construct and the size of newly formed neurospheres was assessed. SiRNA-mediated knockdown of mOct4pg9 resulted in a dose-related decrease in diameter of the neurospheres formed after 48 hours ($F(2,649)=8.733$, $p=0.0002$), whereas Oct4pg9 overexpression significantly increased neurosphere size ($t(34) =9.324$, $p<0.0001$, Figure B 4a).

EFFECT OF MOCT4PG9 OVEREXPRESSION AND KNOCKDOWN ON CELL DEATH

Ethanol itself has been shown to induce cytotoxicity and consequently, the release of LDH from neural cells in culture [201] and *ex vivo* models [202], though we were unable to document similar cytotoxicity in cultured neural progenitor cells [86, 203]. However, since ethanol exposure resulted in the elevation of *m*Oct4pg9 transcript expression, we assessed cell survival after *m*Oct4pg9 overexpression and siRNA mediated knockdown by measuring LDH release from cells into culture medium. *m*Oct4pg9 overexpression did not alter LDH activity released into culture medium ($t_{(20)}$ =2.034,ns). In contrast, siRNA-mediated *m*Oct4pg9 knockdown resulted in decreased LDH release after 48 hours of transfection, $t_{(8)}=10.8$, $p=0.0001$) suggesting a suppression in baseline levels of cell death (Appendix 1 n).

MOCT4PG9 EFFECTS ON BASAL AND OCT4/POU5F1-MEDIATED DNA SYNTHESIS.

To assess mOct4pg9 effects on basal and Oct4/Pou5f1-mediated DNA synthesis, neurospheres were transfected with either non-targeting control siRNA or siRNAs targeting mOct4pg9 or Oct4/Pou5f1 or a mixture of siRNAs targeting both Oct4/Pou5f1 and mOct4pg9 median fluorescent intensity (MFI) of incorporated EdU was measured to assess the rate of DNA synthesis. There was a statistically significant interaction between siRNA-mediated knockdown of mOct4pg9 and Oct4/Pou5f1 (Two-way ANOVA, interaction: $F(1,30)=6.587$, $p=0.0155$, Figure B 4b). Oct4/Pou5f1 knockdown by siRNA, mimicking the suppressive effect of ethanol,

significantly increased DNA synthesis rate by 30% relative to control (post-hoc Fisher's LSD, $p<0.0001$). In contrast, siRNA-mediated mOct4pg9 knockdown did not significantly alter the DNA synthesis rate. However, simultaneous transfection with siRNAs to Oct4/Pou5f1 and mOct4pg9 prevented the increase in DNA synthesis observed with siRNA to Oct4/Pou5f1 alone. As with siRNA-mediated knockdown, mOct4pg9 overexpression also did not result in a significant change in DNA synthesis as measured by EdU fluorescence (Figure B 4b).

OVERLAP AND DIVERGENCE IN THE EFFECTS OF MOCT4PG9 AND ETHANOL ON NSC IDENTITY AND MATURATION MARKERS.

Stem/progenitor cell marker: Ethanol exposure resulted in a dose-related decrease in the expression of mRNA for the NSC-specific intermediate filament protein, nestin ($F(3, 16)=3.451$, $p=0.0417$), whereas neither mOct4pg9 overexpression ($t(14)=1.486$, ns) or siRNA mediated knockdown ($t(6)=0.1852$, ns) had any effect on its expression. However, in contrast to siRNA-mediated knockdown, a more permanent, CRISPR-mediated deletion of the mOct4pg9 locus resulted in a decrease in nestin mRNA expression $t(22)=2.839$, $p=0.0095$,Figure 2 5).

Neuronal markers: Ethanol exposure resulted in a dose-related increase in mRNAs for three neuron-specific proteins, the microtubule-associated proteins, Map2 ($F(3,16)=18.69$, $p<0.0001$), and Dcx ($F(3,16)=6.111$, $p=0.0057$) as well as NeuN/Rbfox3 ($F(3,16)=13.21$, $p=0.0001$). mOct4pg9 over-expression did not alter Map2 mRNA ($t(14)=0.995$, ns), or Dcx mRNA ($t(14)=1.949$, ns), but significantly increased Rbfox3/NeuN mRNA ($t(14)=2.698$, $p=0.017$). In contrast, siRNA-mediated mOct4pg9 knockdown resulted a significant decrease in Map2 mRNA ($t(6)=4.924$, $p=0.002$), and a significant, 80% reducstion in Dcx mRNA ($t(6)=9.614$, $p<0.0001$), but no change in NeuN/Rbfox3 mRNA ($t(6)= 0.8934$, ns). CRISPR mediated mOct4pg9 deletion also resulted in significant decrease in Map2 ($t(22)=2.647$, $p=0.0147$) and Dcx mRNAs

(t(22)=2.52, p=0.019), but as with siRNA-mediated knockdown, no change was observed in NeuN/Rbfox3 (t(22)=1.685, ns, Figure 2 5).

Astrocytic markers: Ethanol exposure resulted in a dose-related increase in mRNA transcripts for both astrocytic markers, Gfap and Glast/Slc1a3 ($F(3, 15)=23.82$, $p<0.0001$ and $F(3,16)=82.69$, $p<0.0001$ respectively). mOct4pg9 overexpression also resulted in a greater than two-fold, statistically significant increase in Gfap mRNA (t(14)=4.095, p=0.0011), GFAP protein immunoreactivity (t(6)=2.569, p=0.042), and GFAP immunofluorescence in neurosphere cultures (t(21)=2.617, p=0.0161, Appendix 1 o), but did not affect Glast/Slc1a3 mRNA levels (t(14)=0.559, ns). siRNA-mediated mOct4pg9 knockdown resulted in no change in Gfap mRNA expression (t(6)=0.511, ns), however it resulted in 25% increase in Glast/Slc1a3 (t(6)=3.319, p=0.016). However, CRISPR-mediated mOct4pg9 deletion downregulated Gfap mRNA (t(22)=2.628, p=0.015), but did not affect Glast/Slc1a3 mRNA levels (t(22)=1.676, ns, Figure 2 5).

Oligodendrocytic markers: Pdgfra was significantly upregulated by ethanol ($F(3,16)=9.357$, $p=0.0008$). Like ethanol, mOct4pg9 overexpression also resulted in upregulation of Pdgfra mRNA (t(14)=2.589,p=0.0219). However, neither mOct4pg9 siRNA-mediated knockdown (t(6)=1.19, ns) or CRISPR mediated knockdown (t(14)=1.1592, ns) altered Pdgfra mRNA. Olig2 mRNA levels were significantly downregulated by ethanol ($F(3, 16)=5.851$, $p=0.0068$). However, neither mOct4pg9 overexpression (t(8)=0.99, ns), CRISPR mediated knockdown (t(14)=0.7672, ns) or siRNA mediated knockdown (t(6)=0.727, ns,Figure 2 5) resulted in alterations in Olig2 mRNA levels. The effects of ethanol exposure, mOct4pg9 overexpression and knockdown on NSC identity and maturation markers are summarized in Table 2 4.

2.4 DISCUSSION

Mechanisms underlying the emergence of non-genetic congenital anomalies including those associated with FASD are often complex and causal mechanisms are difficult to untangle. Here we assessed the contribution of a novel ncRNA, encoded within a pseudogene locus, to the developmental effects of prenatal ethanol exposure. Pseudogenes are as numerous as protein coding genes but have, until recently, been characterized as genetic debris and non-functional transcripts, and their contribution to biology and the pathogenesis of disease, little understood [204]. The contribution of pseudogene biology to the etiology of disorders of fetal development, like FASD, has also not previously been described. However recent reports suggest that pseudogenes are in fact, functional ncRNAs [204], and we hypothesized, as with other ncRNAs like HOTAIR [205], that they play an important role in developmental programming, and therefore, in the pathogenesis of developmental diseases.

Pseudogene family members of the transcription factor Oct4/Pou5f1 are important candidates for investigation, since Oct4/Pou5f1 itself is essential for stem cell identity and self-renewal [171]. Consequently, ncRNAs that modify Oct4/Pou5f1's function are positioned to significantly modify developmental programs. This hypothesis is supported, in part, by recent reports that a number of Oct4 processed pseudogenes are active in human cancers. For example, human OCT4PG5/POU5F1P5 and OCT4PG4/POU5F1P4 have been shown to contribute to aggressive growth and proliferation in gastric [206], cervical [207] and hepatocellular [184, 208] tumors. These data suggest that, at least in tumor tissues, Oct4 pseudogenes facilitate growth. Though nothing is known about their contribution to development, or to the pathologies associated with development, it is interesting to note that our examination of co-expression networks assembled

from publicly available RNAseq data predicted an association between the expression of human POU5F1P4 and POUF5F1P5 and congenital defects including failure to thrive in infancy, as well as defects in brain, mid-face, pulmonary system and extremities, that partly overlap with FASD phenotypes.

Based on sequence homology, the C57bl/6 mouse genome is predicted to contain 6 processed Oct4 pseudogene loci that lack Oct4/Pou5f1 open reading frames, suggesting that any transcription from these loci likely gives rise to ncRNAs. Whereas mouse embryonic stem cells expressed RNA transcripts from 3 out of 6 presumptive loci, mouse-derived NSCs expressed RNA transcripts from only the *m*Oct4pg9 locus. Moreover, analysis of published RNAseq data shows that *m*Oct4pg9-derived RNA transcripts continue to be expressed in adult glial but not neuronal cells, and in the absence of Oct4/Pou5f1. These data suggest: *firstly*, that transcription from Oct4 pseudogene loci is actively regulated by as-yet-unknown developmental and cell-type specific mechanisms; *secondly*, that the expression of *m*Oct4pg9 ncRNAs is tied to glial rather than neuronal maturation. Therefore, whereas, *m*Oct4pg9 exhibited significant sequence homology to human POU5F1P4 and to POU5F1P5, we hypothesized that it was unlikely that *m*Oct4pg9 ncRNAs shared the tumor growth promoting effects attributed to human POU5F1P4 and POU5F1P5 [209-211], but perhaps shared their linkage to predicted congenital birth defects instead, and play a role in stem cell maturation.

Somatic stem cells generally exhibit limited proliferation to balance the demands for self-renewal and maturation [212], and in the ventricular zone (VZ) of the fetal brain, their low rate of proliferation does not account for the neuron numbers that need to be generated [213]. Rather, a

proliferative burst, termed 'transit amplification' in intermediate progenitors in a secondary neurogenic region, the subventricular zone (SVZ), generates the required numbers of neurons [214]. Consequently, a phenotype of increased proliferation is more likely to be associated with NSC maturation rather than self-renewal. Here, we report that ethanol exposure resulted in a dose-related loss of OCT4/POU5F1 protein, suggesting a loss of NSC self-renewal capacity, as well as an increase in *m*Oct4pg9 ncRNA. Moreover, siRNA-mediated Oct4/Pou5f1 knock-down resulted in an increased rate of DNA synthesis, an outcome which is consistent with the progression of stem cell maturation through transit-amplification. Moreover, the effect of Oct4/Pou5f1 knockdown was abolished by simultaneously knocking down *m*Oct4pg9 ncRNA. These data suggest that, unlike the synergy observed in tumor cells, *m*Oct4pg9 serves as a *functional antagonist* to Oct4/Pou5f1 in NSCs micro-dissected from fetal mouse brain, and acts as a factor in maturation of NSCs to transit-amplifying intermediate progenitors. It is also interesting to note that in contrast to the overexpression of *m*Oct4pg9 ncRNA which resulted in a loss of Oct4/Pou5f1, *m*Oct4pg9 inhibition did not alter Oct4/Pou5f1 expression, suggesting that loss of *m*Oct4pg9 is not, by itself, sufficient to promote stem cell identity.

The overexpression of *m*Oct4pg9 ncRNA itself also resulted in increased growth of neurosphere cultures, and although we were unable to directly document stimulation of DNA synthesis by *m*Oct4pg9, pathway overrepresentation analyses of our single cell RNAseq data showed that *m*Oct4pg9^{+} cells were enriched for the polo-kinase pathway that promotes mitosis and the appearance of transit amplifying intermediate progenitors in the developing nervous system [215]. Moreover, loss of function analyses, as discussed above, showed that mOct4pg9 does control cell proliferation during transit amplification. These data are consistent with previous studies which showed that developmental ethanol exposure resulted in decreased proliferation in the VZ (the

location of NSCs), but increased proliferation in the SVZ [160], and our own data showing that ethanol exposure resulted in the loss of CD117, CD133, Sca-1 and ABCG2, cell-surface markers for stem cell identity, increased asymmetric cell division and proliferation [86], specifically in $CD24^{+ve}$ transit-amplification cells [163], and maturation [164, 168].

The mechanisms that are mediated by *m*Oct4pg9 have not previously been identified. ncRNAs can recruit polycomb repressor protein complexes to silence chromatin [216], but despite its localization to both nucleus and cytoplasm, *m*Oct4pg9 ncRNA did not regulate global DNA methylation or hydroxymethylation. Gene-specific effects of *m*Oct4pg9 ncRNA on chromatin packing cannot be ruled out, and will need to be assessed in future, However, *m*Oct4pg9 ncRNA may alternatively regulate gene networks by serving as a miRNA sponge, like human POU5F1PG4 [184]. We therefore assessed the capacity of *m*Oct4pg9 ncRNA to sequester and regulate cellular miRNAs. Our data show that *m*Oct4pg9 ncRNA does co-immunoprecipitate with both Ago2, the catalytic component of the RNA-induced silencing (RISC) complex, and select miRNAs, specifically, miR-328-3p and miR-744-5p. Moreover, *m*Oct4pg9 ncRNA overexpression resulted in decreased expression of 16 miRNAs, including its binding partner, mmu-miR-328-3p. Therefore, *m*Oct4pg9 ncRNA does indeed act as a selective miRNA sponge for at least one miRNA but may also indirectly inhibit the expression of a larger number of miRNAs. These data support the hypothesis that *m*Oct4pg9 ncRNA contributes an additional layer of control to a miRNA-based regulatory network within NSCs, to diminish the effects of anti-proliferative miRNAs like miR-328-3p [217] on key downstream pathways, which may include Wnt signaling pathways [218] which facilitate neural development [219].

It has also been reported that ncRNAs protect mRNAs from translational repression by sequestering miRNAs [220, 221]. The 3' tail of *m*Oct4pg9 ncRNA has retained 49.6% sequence identity to the 3'UTR of Oct4/Pou5f1 (**Error! Reference source not found.**a) and may therefore, at least partly, protect Oct4/Pou5f1 translation by sequestering miRNAs. The possibility of translational disinhibition was supported by data showing that *m*Oct4pg9 ncRNA overexpression resulted in increased expression of OCT4/POU5F1 mRNA and protein, an outcome that was also reported for human Oct4 pseudogenes like POU5F1P4 which has been shown to protect OCT4/POU5F1 from translational repression [184]. However, both mouse and human Oct4/Pou5f1 have short 3'-UTRs of ~200 nucleotides. Short 3'-UTRs typically enable mRNA transcripts to escape translational repression by miRNAs [222], and therefore, Oct4/Pou5f1 has a decreased probability of being regulated by miRNAs compared to human nanog, another core transcription factor for stem cell renewal, which has a long 3'-UTR of >4kb. Moreover, our single-cell RNAseq data showed that *m*Oct4pg9 and Oct4/Pou5f1 transcripts were not expressed by the same cell, and contrary to the data published for human POU5F1P4, *m*Oct4pg9 overexpression resulted in decreased Oct4/Pouf51 3'-UTR-controlled luciferase expression. These data indicate that *m*Oct4pg9 ncRNA, acting at the 3'-UTR, is a translational repressor for Oct4/Pou5f1. This result explains why *m*Oct4pg9 and Oct4/Pou5f1 are not expressed in the same neural cells and is consistent overall with the effects of ethanol exposure. It is possible that *m*Oct4pg9 ncRNA inhibits 3'-UTR-dependent Oct4/Pou5f1 translation by miRNA-independent mechanisms. For instance, the lncRNA OCC-1 destabilizes RNA-binding proteins, which has the effect of destabilizing mRNA transcripts and consequently, inhibiting gene translation [223].

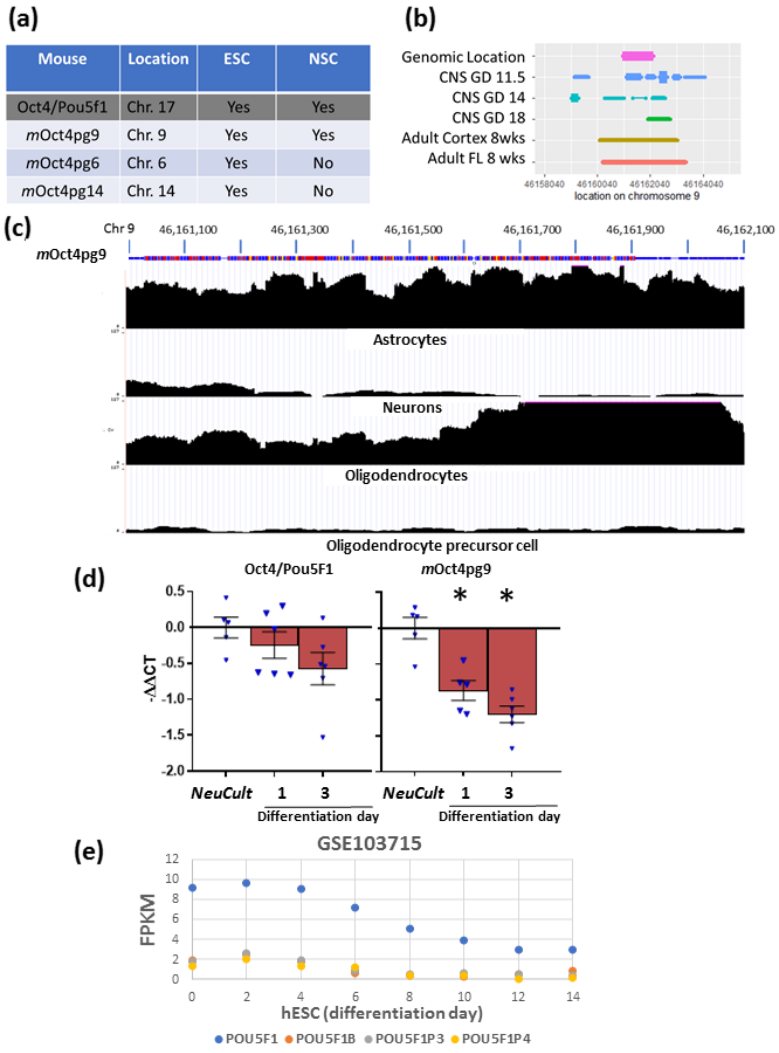
One explanation for the non-congruence between our observations that *m*Oct4pg9 ncRNA decreased 3'-UTR-dependent Oct4/Pou5f1 translation, but increased OCT4/POU5F1 protein immunoreactivity, is that intercellular communication between cells may result in signals from *m*Oct4pg9 ncRNA^{+ve} cells that induce Oct4/Pou5F1 expression in neighboring cells. This type of intercellular communication, recently characterized in mammalian tissues as analogous to 'quorum-sensing' behavior in bacterial colonies [224], is required for maintaining the regenerative capacity of stem cell niches, and may serve to also maintain stem cell populations in the developing fetal brain. An intercellular communication hypothesis is further supported by the analysis of gene transcripts that are shared by *m*Oct4pg9 ncRNA^{+ve} and Oct4/Pou5F1^{+ve} cells, but not by other neighboring neural progenitors. This unique shared subset of genes was overrepresented within the Mevalonate/Geranylation pathway which facilitates the prenylation of proteins, a post-translation lipid modification that is needed for attaching receptors and cell signaling molecules to cell membranes to facilitate intercellular communication [225]. Moreover, the ATM (Ataxia Telangiectasia Mutated) pathway, a second shared pathway is also activated by phrenylated signaling proteins [226] and promotes neural stem cell proliferation [227]. It remains to be determined whether shared signaling pathways facilitate communication between these otherwise distinct NSC sub-populations as a quorum sensing mechanism to inhibit renewal capacity.

In conclusion, we find that one of the 6 screened mouse Oct4/Pou5f1 pseudogene loci gives rise to a novel ethanol-sensitive RNA, *m*Oct4pg9 ncRNA, in fetal NSCs. *m*Oct4pg9 ncRNA has the capacity to selectively sequester miRNAs and Ago2, but unconventionally, inhibits Oct4/Pou5f1 3'-UTR-dependent translation in NSCs. Gain- and loss-of-function analyses show that its expression is associated with increased cell proliferation and maturation, and may contribute to

the pro-maturation effects of ethanol exposure that results in the loss of neural stem cells, and consequently, decreased brain growth. To our knowledge, this is the first report that implicates pseudogenes as mediators of teratogenic processes.

2.5 FIGURES

FIGURE 2-1: IDENTIFICATION AND CHARACTERIZATION OF OCT4 PSEUDOGENES. REPRINTED WITH PERMISSION FROM [228].

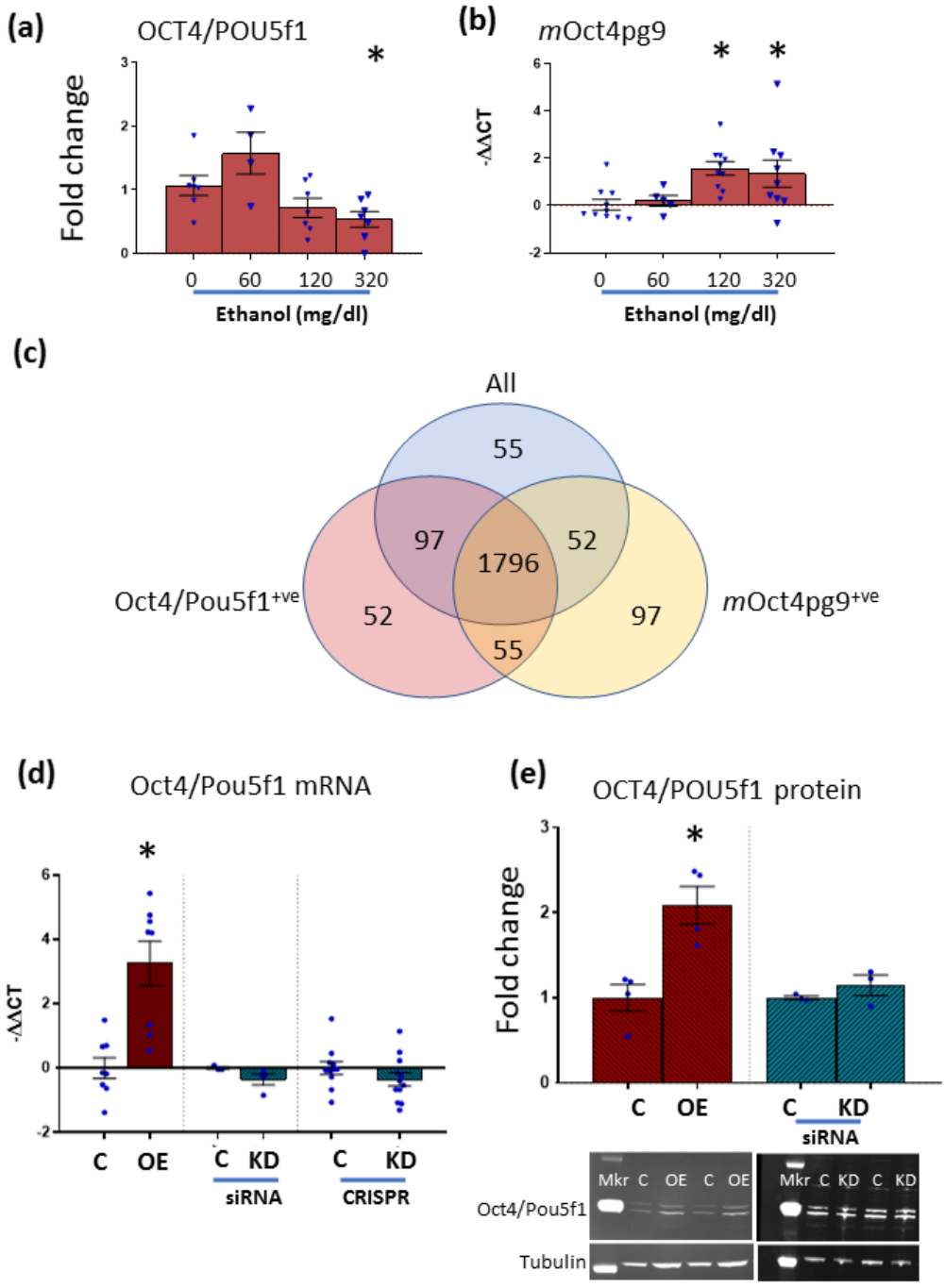


(a) Genomic location and expression status of Oct4 and Oct4 pseudogenes in embryonic stem cells (ESC) and neural stem cells (NSC) in neurosphere culture. (b) RNA transcripts aligning with the genomic location of *m*Oct4pg9, ± 2000 nucleotides, in different brain tissues at various developmental time points. Thickness of the bar indicates expression level (FPKM) of the transcript. (c) Analysis of publicly available single-cell RNAseq data from the Tabula Muris consortium compendium (Schaum et al., 2018) showing RNA transcripts bracketing the region of *m*Oct4pg9 in astrocytes, neurons, oligodendrocytes, and oligodendrocyte progenitor cells (OPC) from adult C57bl/6JN mice. (d) Expression levels of Oct4/Pou5f1 and *m*Oct4pg9 during the time course of differentiation of neurospheres, with data expressed as -ΔΔCT relative to neurosphere culture (*NeuroCult*), bars represent mean ± SEM, individual points are independent biological

Figure 2.1 continued:

replicates, * indicates Fisher LSD post hoc relative to neurosphere culture (*NeuroCult*) p-value <0.05. (e) Expression of POU5F1, POU5F1B, POUF1P3 and POU5F1P4 through neural differentiation time-course of human embryonic stem cells (hESCs) (Pluripotency (day 0); differentiation initiation (days 2, 4, and 6); neural commitment (days 8–10) and neural progenitor cell proliferation (days 12, and 14, GSE103715, (Li et al., 2017)).

FIGURE 2-2: Characterization of mOct4pg9 regulation of Oct4/Pou5f1. *REPRINTED WITH PERMISSION FROM [228].*

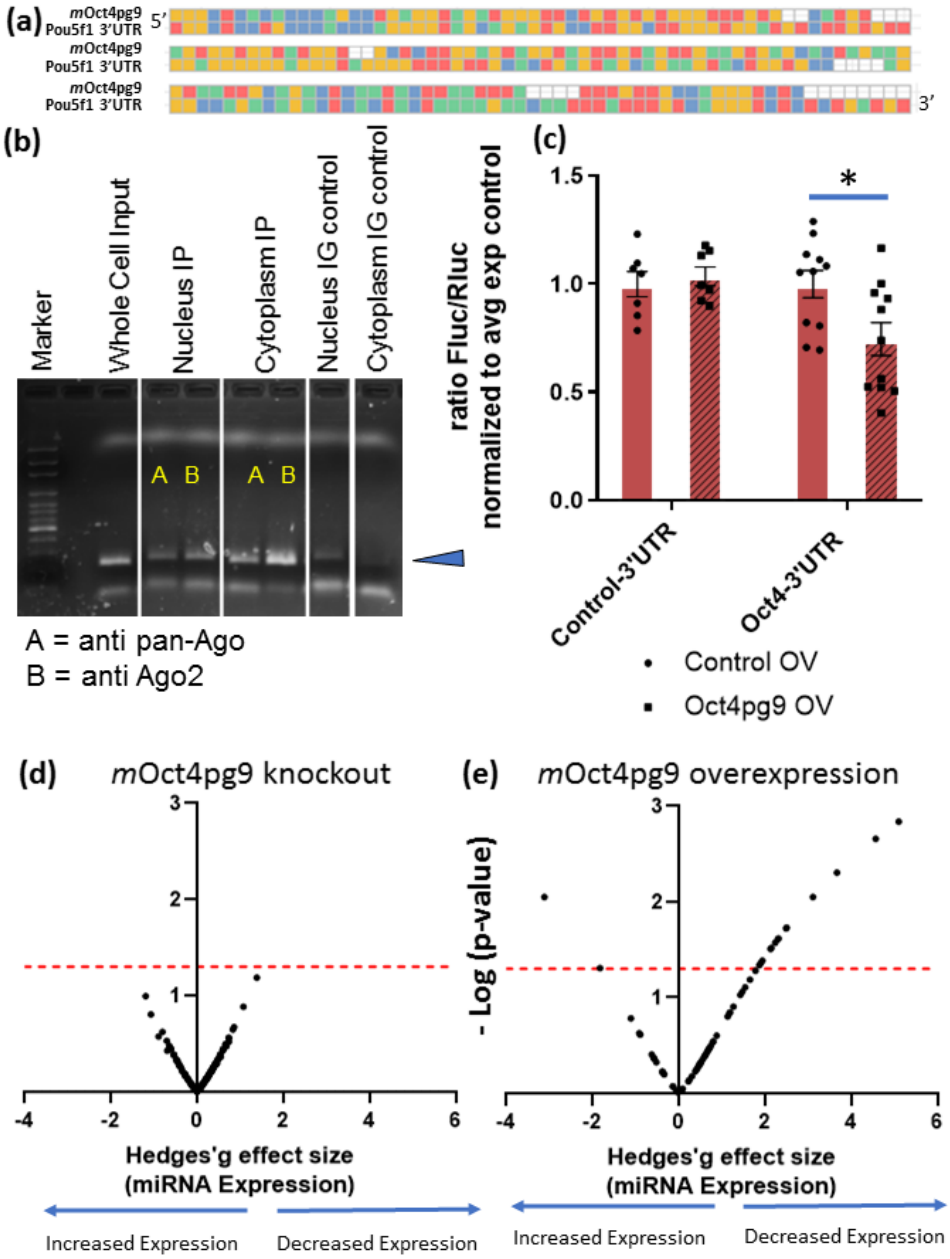


(a) Analysis of western immunoblot assay for OCT4/POU5F1 protein and β-actin proteins in lysates of neurospheres treated with 0, 60, 120 or 320 mg/dl of ethanol. Bar-graph represents mean expression of Oct4/Pou5f1, normalized to β-actin expression, in each group compared to average

Figure 2.2 continued:

expression in the controls. (b) *m*Oct4pg9 ncRNA expression, in ethanol-treated neurosphere cultures, expressed as -ΔΔCT, where a difference of 1 ΔΔCT unit represents an ~2-fold change in expression. (c) Venn diagram showing the overlap between the 2000 most abundantly expressed genes in Oct4/Pou5f1^{+ve}, and *m*Oct4pg9^{+ve} compared to the most abundant gene transcripts in all cells. (d,e) RT-PCR analysis of Oct4/Pou5f1 mRNA (d) and western immunoblot assays for the expression of OCT4/POU5F1 protein (e) in lysates from *m*Oct4pg9 overexpression (OE) and control neurospheres and *m*Oct4pg9 knockdown (KD) and control neurosphere cultures. Top panel (e) shows the quantification of the immunoblots normalized to loading control protein (tubulin) represented as a ratio to the average intensity in the control. Bars represent mean ± SEM, individual points represent independent biological replicates, * indicates unpaired two-tailed t-test p-value< 0.05 (full uncropped Western immunoblots are presented in Appendix 1 j).

Figure 2-3: mOct4pg9 lncRNA association with miRNAs and Argonaut chaperone proteins

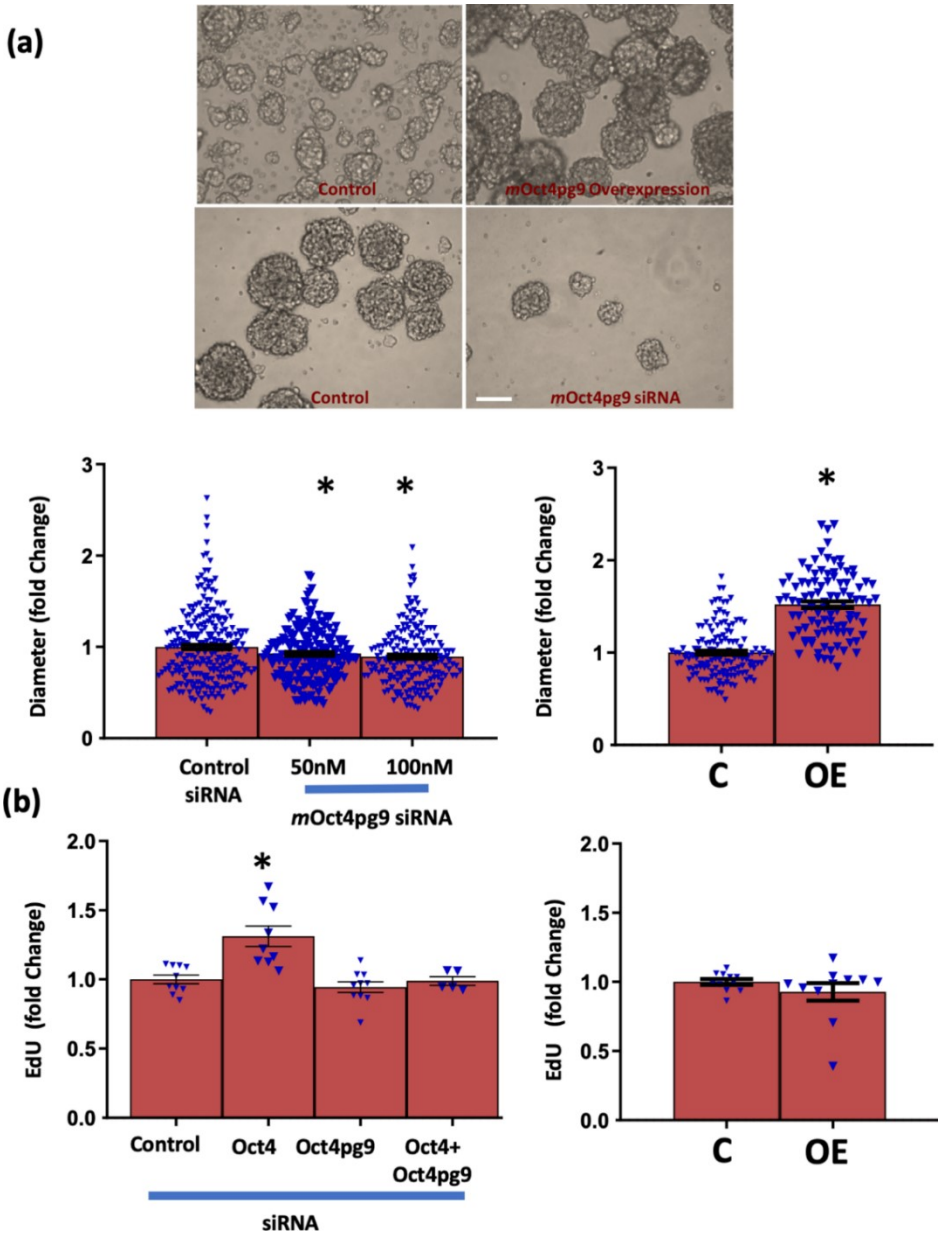


(a) Schematic diagram (ggmsa R package) showing the sequence homology between Oct4/Pou5f1 3'-UTR and *m*Oct4pg9 (Red, Blue, Orange and Green represent A, C, G and T nucleotides respectively). (b) Electrophoretic fractionation of PCR amplicons obtained from reverse-transcribed RNA amplified with primers for *m*Oct4pg9 following immunoprecipitation with anti-pan-Ago or anti-Ago2 antibodies, or control IgG (IG) of subcellular fractionated lysates from neurosphere cultures. (c) Ratio of *f*luc/Rluc intensity resulting from miRNA Target clone control vector and Oct4/Pou5f1 3'-UTR luciferase reporter vector co-electroporated with either *m*Oct4pg9 overexpression or control overexpression plasmid. * indicates p-value <0.05, Fisher LSD post hoc test. (d, e) Volcano plots showing statistical significance (-log[p-value]) compared to Hedges's g

Figure 2.3 continued:

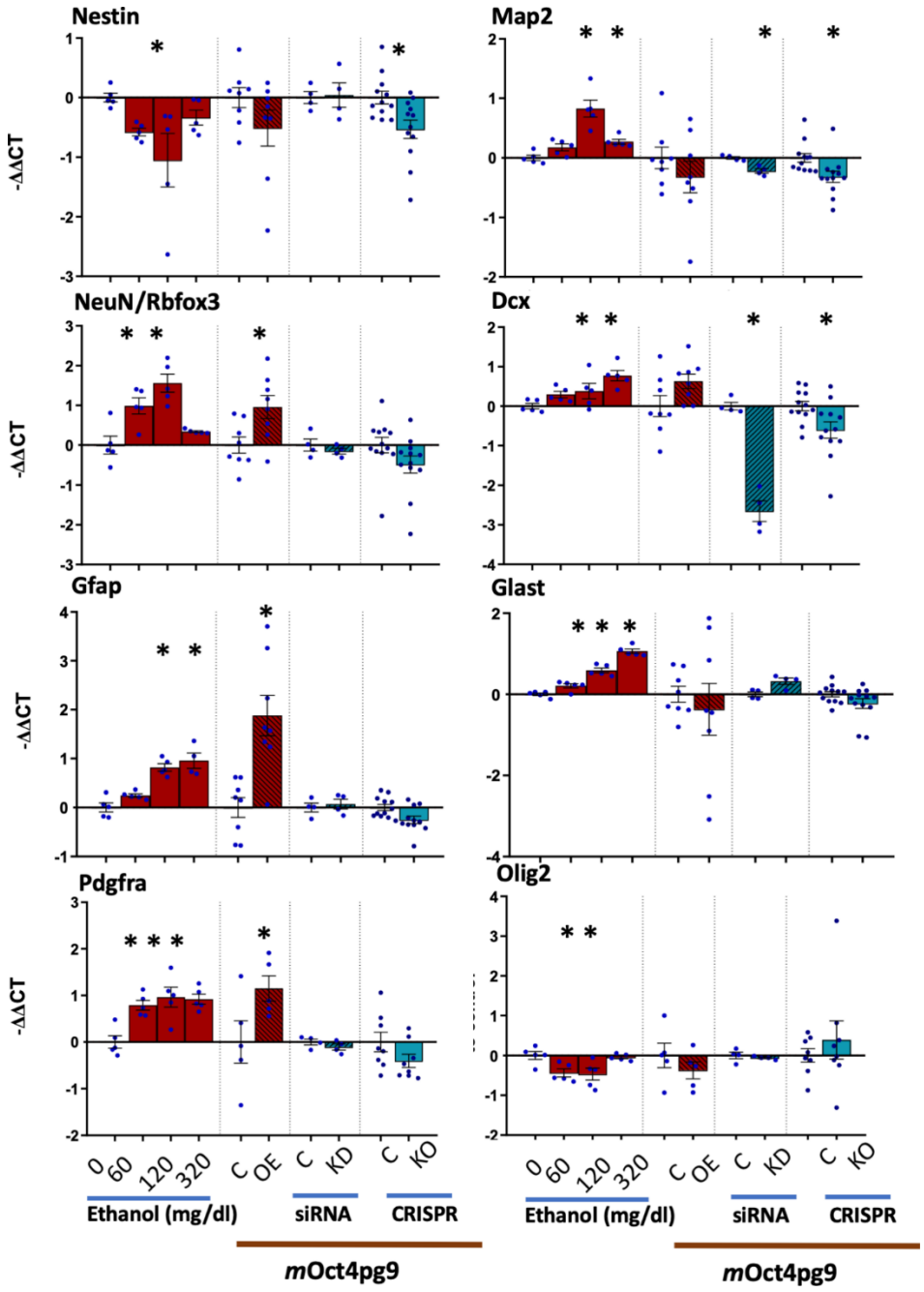
effect size of the change in miRNA expression following *m*Oct4pg9 knockdown (d) and overexpression (e). Red line marks the threshold of $p<0.05$ (unpaired two-tailed t-test).

FIGURE 2-4: EFFECTS OF MOCT4PG9 NCRNA MANIPULATION ON NEUROSPHERE SIZE AND DNA SYNTHESIS. REPRINTED WITH PERMISSION FROM [228].



(a) Representative images and quantification of sizes of neurospheres after *m*Oct4pg9 siRNA mediated knockdown and overexpression (OE). (b) Rate of DNA synthesis, expressed as mean fluorescence intensity (MFI) of EdU, after either knockdown of Oct4/Pou5f1 and/or *m*Oct4pg9 or following *m*Oct4pg9 overexpression. * indicates Fisher LSD post-hoc (left) or unpaired two tailed t-test (right), p-value<0.05 compared to control. Scale bar, 100 um.

FIGURE 2-5: RNA EXPRESSION LEVEL OF KEY STEM CELL AND DIFFERENTIATION-ASSOCIATED GENES IN LYSATES FROM OF ALCOHOL TREATED NEUROSPHERE CULTURES. REPRINTED WITH PERMISSION FROM [228].



RNA expression level of key stem cell and differentiation-associated genes in lysates from of neurosphere cultures treated with 0, 60, 120 or 320 mg/dl of ethanol, overexpressing *m*Oct4pg9 (OE), or treated with siRNA to knockdown (KO) or CRISPR sgRNA to knockdown (KO)

Figure 2.5 continued:

*m*Oct4pg9. Data is expressed as -ΔΔCT compared to the control of each experiment. Bars represent mean ± SEM, individual points represent independent biological replicates, * indicates p-value <0.05 for either Fisher LSD post hoc (compared to the control) or unpaired two-tailed t-test.

2.6 TABLES

TABLE 2-1 : SHOWING MOUSE GENOMIC LOCI WITH HIGH SEQUENCE HOMOLOGY TO MOUSE POU5F1/OCT4 MRNA. REPRINTED WITH PERMISSION FROM [228].

Pseudogene	**Coverage**	**E-Value**	**Identity**	**Accession**	**Expression in**	
					ESCs	NSCs
mOct4pg3	100%	0.0	85%	NR_033594.1	NA*	NA*
mOct4pg14	78%	0.0	84%	NR_144493.1	Yes	No
mOct4pgX	98%	0.0	77%	NR_144496.1	No	No
mOct4pg1	60%	0.0	100%	NR_144489.1	NA*	NA*
mOct4pg6	46%	5e-166	84%	NR_144494.1	Yes	No
mOct4pg9	75%	5e-103	71%	NG_020980.1	Yes	Yes

*: The design of primers specific to mOct4pg3 and mOct4pg1 was not feasible due to the high sequence homology to mouse Pou5F1/Oct4 mRNA

TABLE 2-2: HUMAN TRANSCRIPTS WITH HIGH SEQUENCE HOMOLOGY TO MOUSE OCT4PG9 NCRNA. REPRINTED WITH PERMISSION FROM [228].

Homo Sapiens Homologs	**Accession**	**Coverage**	**E-Value**	**Identity**
POU class 5 homeobox 1 (POU5F1), transcript variant 1, mRNA	NM_002701.5	88%	4e-72	69%
POU class 5 homeobox 1 pseudogene 3 (POU5F1P3), non-coding RNA	NR_036440.1	73%	2e-58	68%
POU class 5 homeobox 1 pseudogene 4 (POU5F1P4), non-coding RNA	NR_034180.1	68%	9e-56	68%
POU class 5 homeobox 1 pseudogene 5 (POU5F1P5), non-coding RNA	NR_131184.1	42%	6e-26	70%

TABLE 2-3 PATHWAY OVERREPRESENTATION ANALYSIS. REPRINTED WITH PERMISSION FROM [228].

The first two columns show pathways associated with genes that were uniquely expressed in Oct4/Pou5f1^{+ve} cells (52 unique genes) or mOct4pg9^{+ve} cells (97 unique genes). The last column lists significant pathways associated with genes that are expressed in common by Oct4Pou5f1^{+ve} cells and mOct4pg9^{+ve} cells (55 unique genes) but not in neighboring neural progenitor cells (see Figure 2C and supplementary file 2 for additional details).

Oct4/Pou5f1^{+ve} cells	**mOct4pg9^{+ve} Cells**	**Shared Pathways**
-1D-myo-inositol Hexakis-phosphate Biosynthesis	-Pyrimidine Deoxyribonucleotides De Novo Biosynthesis I	-Chromosomal Replication Control
-Spermidine Biosynthesis	-Mitotic Roles of Polo-Like Kinase	-Mevalonate/Geranylgeranyl-diphosphate Pathway
-Systemic Lupus Erythematosus in B Cell Signaling		-ATM Signaling
-Role of BRCA1 in DNA Damage Response		-Cholesterol Biosynthesis

TABLE 2-4: SUMMARY EFFECTS OF ETHANOL EXPOSURE, OCT4PG9 OVEREXPRESSION, SIRNA &CRISPR MEDIATED OCT4PG9 KNOCKDOWN ON THE EXPRESSION OF MRNA MARKERS FOR NEURAL STEM CELL IDENTITY AND MATURATION. REPRINTED WITH PERMISSION FROM [228].

				Oct4pg9 knockdown	
		Ethanol Exposure	Oct4pg9 Overexpression	siRNA mediated	CRISPR mediated
Stem Cell	Nestin	Downregulated	No Change	No Change	Downregulated
Neuronal	DCX	Upregulated	No Change	Downregulated	Downregulated
	NeuN	Upregulated	Upregulated	No Change	No Change
	MAP2	Upregulated	No Change	Downregulated	Downregulated
Oligo-dendrocyte	Olig2	Downregulated	No Change	No Change	No Change
	PDGFRa	Upregulated	Upregulated	No Change	No Change
Early Astrocyte	GFAP	Upregulated	Upregulated	No Change	Downregulated
Late astrocyte	GLAST	Upregulated	No Change	Upregulated	No Change

3 CELL TYPE AND SEX-SPECIFIC TARGETS OF PRENATAL ALCOHOL EXPOSURE IN THE MOUSE DEVELOPING CORTEX: XIST-DEPENDENT FEMALE SPECIFIC MECHANISMS.

3.1 SUMMARY

Prenatal alcohol exposure (PAE) results in cerebral cortex dysgenesis; microencephaly, decreased neuron number and disrupted neuronal migration. Here we implement single cell RNA sequencing to decipher the cell-type specific effects of PAE on murine developing cortex in a model of single dose binge alcohol exposure during early fetal cortical neurogenesis. The 139,195 cells were segregated into 33 clusters which were further grouped into 8 lineage subtypes. We identified gene co-expression networks disrupted by alcohol exposure. In female fetal brains, Xist, was a significant module hub gene, whose downregulation in PAE was accompanied by a loss of X inactivation. In addition, genes associated with autism spectrum disorders are targeted by PAE in males and females, indicating overlapping mechanisms mediating neurodevelopmental disorders. This work presents evidence that fetal sex is an important contributor to neural response to maternal alcohol exposure.

3.2 INTRODUCTION

Congenital neurodevelopmental disabilities are increasingly common. For instance, a recent National Health Interview Survey (NHIS) found that 17.8% of children surveyed in 2015-2017 were reported to have at least one developmental disability, specifically attention-deficit/hyperactivity disorder (ADHD, 9.5%), autism spectrum disorders (ASD, 2.5%) and

intellectual disabilities (ID, 1.2%), and moreover, those percentages were higher than those reported in 2009-2011 and 2012-2014 surveys [229]. Perturbations in the maternal-fetal environment are likely to be a significant, but unresolved contributor to the emergence of these developmental disorders [230]. Recent state-wide biomarker-based studies show that prenatal alcohol exposure (PAE) in the US accounts for ~8% of births [71, 231], making alcohol a common maternal-fetal perturbagen, and a leading cause of congenital anomalies and neurodevelopmental disabilities. The resulting umbrella diagnosis of Fetal Alcohol Spectrum disorders (FASD) is prevalent in between 1.1-9.8% of school- aged children in the US [75]. Moreover, individuals with diagnoses along the FASD continuum are also documented to be co-morbid for ASD, ADHD [232-237] and a number of other diagnostically distinct neuropsychiatric conditions [238]. The prevalence of these co-morbidities in FASD populations is significantly higher than in the general population [238], which suggests that many congenital neurodevelopmental disorders have a shared etiology, with origins in adverse maternal-fetal environments.

Heavy PAE is a well-established cause of brain growth deficits in human populations [239] and these outcomes have been replicated in mouse models of PAE (e.g., [240]). We hypothesize that PAE during the critical period of telencephalic neurogenesis which extends from the mid-first trimester through the second trimester in humans [241], is likely to be a particularly important cause of brain growth deficits. A number of birth-dating studies [242, 243] have shown that the pseudostratified epithelium of the fetal ventricular zone (VZ) generates most neurons of the cortical plate between gestational days 11 and 17, with founder neural stem cell (NSC) populations completing approximately 50% of neurogenic cycles by GD14.5. Mouse models of PAE between gestational days 12.5 to 14.5, corresponding to the first half of the cortical

neurogenic period, and the generation of layer V and VI neurons [243], resulted in ventricular dilation and decreased thickness of the developing cortical plate [164, 244]. We also observed in *ex vivo* models of the murine cortical neuroepithelium, that while ethanol exposure did not kill neural stem cells, it did result in increased asymmetric cell proliferation, loss of cells expressing markers of stem cell identity and increased indices of premature maturation [86, 87, 162, 163]. These data suggested the hypothesis that PAE during the period of neurogenesis results in diminished brain growth, because it results in premature maturation and consequently, loss of NSCs.

In the current study, we used a single-cell RNAseq approach (scRNAseq) to directly assess the persistent effects of a single episode of binge-like maternal ethanol exposure, on the developmental trajectory of fetal cerebral cortical neuroepithelial progenitors. ScRNAseq enables the characterization different cell types based on individual cell transcriptional profiles [122, 245], and has been used to successfully assess cell-type specific mediators of ASD [134-136] and cell-type-specific gene expression and allele-specific X-chromosome activation status in human Rett syndrome brains and in mouse models [246]. We also assessed whether the effects of PAE on cortical development was dependent on fetal sex, since a limited amount of published data in rodent models does suggest that sex differences in PAE-associated behaviors appear later in life [247, 248], and may be mediated by early molecular adaptations in neural cells.

Our data, focused on the VZ, SVZ and layer V/VI neurons, indicate that each of these subpopulations are molecularly and developmentally heterogenous, with up to 22 different cell-types identified within the VZ and SVZ and 8 different neuronal cell-types assigned to layers V/VI. PAE resulted in significant increases in cell cycle, mainly in SVZ-type cells and an exposure and sex-dependent shift in developmental trajectory in VZ, SVZ and neuronal

populations. Moreover, we identified significant sex differences in the response to PAE, biased towards females and linked to loss of X-inactivation and differential expression of ASD-associated genes.

3.3 METHODS

ESTABLISHMENT OF C57/BL6J TIMED-PREGNANCIES:

All animal procedures were approved by the Texas A&M University's Institutional Animal Care and Use taCommittee. Young adult C57Bl/6J nulligravid females were time-mated in-house with C57Bl/6J males. 1-2 females in proestrus/estrus phase were placed with a proven breeder male for a period of 90 minutes, one hour after the start of the dark cycle (at 7:00 pm), and the next morning was designated as gestational day (GD) 0.5 of pregnancy. Pregnancy was confirmed by daily weight gain starting from GD 8.5.

PRENATAL ALCOHOL EXPOSURE (PAE) PARADIGM:

On GD12.5, at the neurogenic time period for the genesis of layer V neurons [243], pregnant dams were placed in either control or alcohol vapor chambers (La Jolla Alcohol Research Inc, quad passive e-vape system), on absorbent pads, to avoid liquid accumulation in the chamber. Ethanol (95%) was injected into the system at a rate of 6 strokes/minute, with an air-flow of between 4 and 5 L/minute. The exposure to control air or ethanol vapor occurred during the dam's subjective night (initiated at 9:00am), over a period of 50 minutes, and animals were thereafter placed under a heat lamp until locomotor activity resumed. At 30 minutes following the exposure period, 20ul of tail-vein blood was aspirated for analysis of blood alcohol content (BAC) by gas chromatography (ThermoFisher, TRACE™ 1310), as we previously published [86, 87, 106, 167, 168]. Pregnant mice were single-housed after ethanol or control air exposures until fetal collection on GD 14.5. Food consumption during the dam's subjective day (6:00pm to

6:00 am) was not altered by this exposure paradigm (p-value > 0.3, *ns*), and there were no statistically significant differences in weight gain between control and PAE dams (p-value > 0.1, *ns*).

FETAL NEOCORTEX TISSUE EXTRACTION AND SINGLE CELL SUSPENSION PREPARATION

On GD14.5, pregnant dams were anesthetized with ketamine (0.09 mg/gram)/xylazine (0.106 mg/gram) and fetal mice delivered by laparotomy as we previously published [86, 87, 163, 167]. The dorsal telencephalon corresponding to the future isocortex of each mouse fetus in the litter was rapidly micro-dissected and immediately placed in chilled PBS in individual microcentrifuge tubes and centrifuged at 200xg for 3 minutes. The PBS was aspirated and tissue mechanically dissociated by trituration in 200 ul 0.25% EDTA free trypsin in PBS for 5 minutes at room temperature, followed by Trypsin inactivated with 1ml of 10% FBS in PBS. The cell suspension was washed twice in PBS and then filtered through a 40-micron filter, centrifuged and resuspended in 20 ul PBS. Pre-chilled methanol (80 ul) was added slowly, dropwise while mixing, to avoid cell clumping. The cell suspension was then placed on ice for 15 minutes before storage at -80^{o}C. To determine the sex of each of the pups, we performed alkaline lysis (25mM NaOH and 0.2 mM EDTA) on tail snips from each pup, at 95^{o}C for 15 minutes, followed by neutralization in 75 ul of Tris-HCl (40mM). Tissue debris were separated from the lysate by centrifugation and the supernatant, diluted 1:10, was assessed for sex-specific genes, using PCR with primers specific for Xist (X-chromosome) and SRY (Y-chromosome) genomic DNA.

CELL COUNTING AND ALIQUOT PREPARATION FOR SINGLE CELL SEQUENCING

Since methanol fixation may result in cell loss, we counted cells after fixation and re-counted immediately prior to combining aliquots for library preparation. To count the cells, we combined

10 ul of cell aliquot with 10 ul DMSO and 80 ul water, then we mixed 10 ul of this diluted aliquot with 10 ul DAPI in mounting media and manually counted on fluorescent microscope (Olympus BX60). We utilized experiments where the pregnant dams had at least 2 male and 2 female fetuses. For each dam we combined cells from 2 male fetuses into one aliquot and cells from 2 female fetuses into a second aliquot, each aliquot containing approximately 100,000 cells.

10X SINGLE CELL LIBRARY PREPARATION AND SEQUENCING

Single cell libraries for 10,000 cells/sample were prepared using Chromium Single Cell 3' Kits (v3 Chemistry) following manufacturer's protocol. Libraries were sequenced per 10x Genomics recommended parameters and depth of ~50,000 reads/cell on an Illumina NovaSeq 6000 (North Texas Genome Center). Libraries were demultiplexed using Cell Ranger (3.1.0 ,10x Genomics, https://www.10xgenomics.com/) *mkfastq* command.

CELL RANGER PIPELINE

The output demultiplexed *fastq* files for each sample were processed using Cell Ranger software pipeline on Texas A&M High Performance Research Computing clusters. First *cellranger count* was used to perform alignment (transcriptome:mm10), filtering, and barcode and UMI counting. The output was then processed through *cellranger aggr* which aggregates the outputs for *cellranger count*, normalizing those runs to the same sequencing depth. The output of those pipelines is an experiment wide feature (gene)-barcode matrix, TSNE, U-MAP plots of the experiment cells. The software performs principal component analysis on the normalized filtered feature barcode matrix. Based on the cells' projection into PCA space, cells that have similar expression profiles are grouped into clusters (*graphclust*). Raw and processed datafiles for each sample and the aggregate samples are deposited in GEO under accession number GSE158747.

IDENTIFICATION OF CLUSTERS AND DETERMINING THE EFFECTS OF ALCOHOL AND FETAL SEX ON CLUSTER MEMBERSHIP.

We used published expression data of cell-type markers in GD14.5 mouse as identified in [249] to define the class identity of each of the cell clusters. For those clusters for which class identity was not directly determinable from patterns of marker RNA expression, we determined the genes that were significantly upregulated or downregulated compared to an identified neighboring cluster. We then used gene ontology analysis to determine the function of the significantly different genes, thus determining the identity of that un-assigned cluster. We used the Seurat [250, 251] *FindClusters* feature to re-cluster cluster #1 which exhibited ambiguity in the expression of known cell identity markers into sub-clusters. For each sample, we quantified the percentage of cells in each of the clusters, or subclusters. To investigate the effects of alcohol exposure and sex on the cell population cluster identities, we performed a two-way ANOVA using the percentage of the cells in a cluster as a dependent variable and both sex and exposure as independent variables, we performed this analysis for each of the clusters.

DETERMINATION OF THE CELL CYCLE PHASE OF EACH CELL AND THE EFFECTS OF ALCOHOL ON PERCENTAGE OF CELLS IN EACH CLUSTER CELL CYCLE PHASES

We used the *Seurat* package cell cycle scoring function to assign each cell a score based on its expression of G2/M and S phase markers, then each cell is classified as G1, S or G2/M. We followed by quantifying, for each sample, the percentage of cells in each phase within each of the clusters/subclusters, then for each sample as a whole, we quantified the percentage of cells in each of the three phases. To investigate whether alcohol exposure or sex resulted in the change of percentages of cells in each phase, we performed, for each cluster and for each phase, a two-

way ANOVA using the percentage of cells in a phase as dependent variable, and alcohol exposure and sex as independent variables, with the Fisher LSD for post-hoc analysis.

CONSTRUCTION OF SINGLE-CELL TRAJECTORY AND ASSIGNMENT OF PSEUDOTIME SCORES

We utilized Monocle3 [252-256] to perform a trajectory analysis on UMAP clustered single cells from all the treatment groups. After trajectory construction the pipeline orders the cells based on the learnt trajectory-pseudo time-. We manually picked the root node of the trajectory graph and *pseudo-time* scores, representing the distance between a cell and the start of the trajectory, measured along the shortest path, were calculated for each cell. Cells that were unreachable from the root node of the trajectory graph were assigned an infinite *pseudotime* score. To investigate whether alcohol exposure or sex resulted in the change of *pseudotime* scores in each cluster, we performed, for each cluster, a two-way ANOVA using the *pseudotime* scores as dependent variable, and alcohol exposure and sex as independent variables.

DETERMINING DIFFERENTIALLY EXPRESSED GENES DUE TO ALCOHOL EXPOSURE.

To identify differentially regulated genes by alcohol in each cluster, or sub-cluster in the case of cluster #1, we used *DEsingle* package to analyze the gene counts in each cluster, comparing control males to PAE males and control females to PAE females. Using Zero-inflated binomial model, DEsingle has the advantage of estimating the proportion of real (not expressed gene) and dropout zero expression (technical zero) values [257]. *DEsingle* shows high accuracy and sensitivity and a better tradeoff between true positive rates and precision, compared to other differential expression analysis tools designed for scRNA or bulk RNA sequencing analyses [258]. *DEsingle* subdivides differentially expressed genes into 3 types, DEs, DEa and DEg. A differentially expressed gene-type is DEs ‘different expression status’, if a gene shows

significant difference in the proportion of real zeros between the two groups, but did not exhibit a significant difference in the overall expression abundance between the two groups. A gene is classified as DEa is for 'DE abundance', when it is significantly differentially expressed between the groups without significant difference in the proportion of real zeros. DEg or 'general differential expression' refers to genes that have significant difference in both the proportions of real zeros and the expression abundances between the two groups. *DESingle* adjusts p-values for multiple comparisons, with our corrected cutoff of α=0.05. We conducted three types of analyses, comparing alcohol exposed to controls in each of males and female samples and comparing control male samples to control female samples. The differential expression tests were run on each of the clusters, in addition to each of the sub-clusters that were obtained by re-clustering of cluster 1.

CONSTRUCTION OF WEIGHTED GENE CO-EXPRESSION NETWORKS, IDENTIFYING NETWORKS MOSTLY AFFECTED BY ALCOHOL EXPOSURE

We selected the top 20% variant genes in each of control male and control female clusters, and then for each cluster in each of the two groups, we utilized the *WGCNA R package* [259] to construct gene co-expression networks. We utilized the "dynamic tree cut" to identify modules of co-expressed genes, and identified gene hubs for each module. To identify the modules highly affected by PAE, we quantified the number of differentially expressed genes (upregulated or downregulated) in each module and calculated a '% differentially expressed' relative to the number of genes in each module.

IDENTIFYING CANDIDATE PAE DYSREGULATED PATHWAYS AND THEIR UPSTREAM REGULATORS BY IPA ANALYSIS

We used Ingenuity pathway analysis (Qiagen) to perform canonical pathway and upstream analyses on the differentially regulated genes (FDR p-value <0.05) identified in each of the clusters in each of the males and females' comparisons (Control vs PAE). Fold-changes of differentially expressed genes were used to predict the z-score of activation of significant pathways or regulators. We identified candidate up-stream transcriptional regulators of dysregulated genes (p-value of overlap <0.05) and candidate pathways significantly enriched for the genes altered by PAE.

3.4 RESULTS

SAMPLE CHARACTERISTICS

Pregnant dams exhibited incomplete sedation and impaired mobility with loss of righting reflex when exposed to ethanol vapor for 45 minutes, and achieved a blood alcohol content (BAC) between 210 and 260mg/dl at 30 minutes following the end of the exposure period (Table 3-1). All dams regained their righting reflex and mobility by 90 minutes following the exposure period and were returned to their home-cage.

We sequenced 139,195 single cells, micro-dissected from GD 14.5 fetal mouse dorsal telencephalon obtained from 6 dams. For each dam we combined two female pups to form one female sample, and two male pups to form one male sample. The number of single cells sequenced ranged from 7475 to 18612 cells/sample. The mean number of reads per cell ranged from 21,662 to 60,579, and the number of genes detected per sample averaged 19,933 in all samples . The percentage of reads mapping to mitochondrial genes in all cells averaged 6.74%. (percentage of mitochondrial reads in each cluster are presented in Appendix 2)

Single cell RNA sequence captures the heterogenicity of the murine fetal developing cortex

A graph-based clustering algorithm grouped the 139,195 individual cells into 33 separate clusters (

Figure 3-1). We used published markers of cell identity to identify and annotate each of the 33 clusters. The clusters were identified as one of 9 cell types. Five clusters expressed high levels of ventricular zone (VZ) markers Prom1 [260] and Pax6 [261]. NeuroG2 [262] and overlapping Pax6 and Eomes [261] expression marked the clusters at the junction between SVZ/VZ zones. Since Sox2 expression was shown to decrease from VZ to SVZ to cortical plate [263], we ordered the clusters within VZ, SVZ/VZ and VZ in order of decreasing mean cluster Sox2 expression. Radial glia markers were expressed in VZ, VZ/SVZ, and some of SVZ clusters. We used markers summarized in [245] to annotate clusters belonging to Layer I, interneurons, Lateral and Rostral layer V-VI, microglia and endothelial cells (

Figure 3-1-b, percentages of each cell type and are shown in Table 3-2 and

Figure 3-1-d). We used the expression of Slc4a1 [264], Alas2 [265] and Hemoglobin genes [266] to identify erythrocytes. Cluster 1 cells (6.5 % of the cells) expressed markers of both immature and mature cell lineages (cluster marked by *

Figure 3-1-b). To better understand this cluster, which included the largest number of cells, we further subclustered it into 8 distinct subclusters (

Figure 3-1-c). We used the significantly enriched genes in each cluster, identified by the *FindMarkers* tool in the *Seurat* pipeline, to identify functions of distinct sub-clusters in Cluster #1(Heatmap for expression of representative enriched genes in each subcluster is presented in

Figure 3-1-c and Appendix 2 -b). Collectively the sub-cluster analysis indicate that groups of Cluster #1 cells exhibit identities of path-finding cells (subclusters 2 and 7), migrating interneuron-like cells (subcluster 5), cells with the capacity to form neuronal projections and expressing markers of glutamatergic synapses (Subcluster 4), Subclusters 1 and 7 were primarily composed of dividing cells

Figure 3-1-c, Table 3-3).

We used published cell identity markers to further group the 33 clusters into 9 major annotated lineages[245]. Cell cycle composition (proportion of cells in G1, S or G2/M phases) and pseudotime score distribution was used to further order each cell cluster within its major annotated lineage. In the case of cluster #1, we also quantified the percentage of cell in each cell cycle phase in each of the sub-clusters. All VZ-annotated clusters were predominantly in S and G2M phases, and 3 out of 6 SVZ clusters were also predominantly in S and G2/M phases. Two SVZ clusters, clusters 13 and 20, which shared profile of cell identity markers varied in their cell cycle phases composition, cluster 13 cells were predominantly in G2M phase while cluster 20 was predominantly in S phase. In contrast, more mature, neuronal lineage committed cells including those within clusters defined as Layer V-VI neurons, or as interneuron were predominantly in G1 phase, as expected. Non-neural cell types, fetal erythrocytes, microglia and endothelial cells, were composed of cells in all three phases of cell cycle (

Figure 3-1-d).

We utilized the UMAP algorithm and ordered the cells based on their progression across a UMAP-initialized trajectory. The UMAP algorithm resulted in one major contiguous cluster, for which a trajectory was constructed and pseudotime was computed, and two additional minor

clusters. We then generated a pseudotime score for all the cells in the contiguous cluster based on their progression across a UMAP-initialized trajectory (Figure 3-2a), and the pseudotime score distribution in each graph-based cluster was used to order that cluster within each of the lineages contributing to the contiguous UMAP cluster (Figure 3-2b). The pattern of expression of VZ, SVZ, and neuronal markers, Pax6, NeuroG2 & Eomes, and Tbr1 respectively [245, 261],along the pseuodtime continuum indicates that progression across pseudotime does reflect a maturation process. Gas1, the proliferative radial glial marker [267], shows high expression in VZ cells and decreasing pattern of expression as SVZ cells progress through pseudotime. Plxna4 a gene crucial for neuronal radial migration [268] shows increased expression at pseudotime axis corresponding to the position of SVZ-type cells and expression was observed to taper-off with pseudotime progression towards cells defined as neurons of lateral layer V-VI (Figure 3-2c).

The two UMAP identified additional non-contiguous minor clusters were composed of two smaller interneurons, two transient progenitor cell clusters, graph-based cluster 4 and 18, and 2 SVZ clusters, graph-based cluster 13 and 20.

We identified 4 clusters that shared the interneuron markers, Dlx1, Dlx2 & Sst. Two of the interneuron clusters-including the largest of the four clusters- expressed markers of MGE derived cortical interneurons in addition to higher levels of Gad1, Gad2 compared to the other interneuron clusters. We used the t-test based *FindMarkers* feature in the *Seurat* package to identify genes enriched in each of the interneuron clusters compared to the total interneuron population.

PAE DID NOT ALTER THE CLUSTER COMPOSITION OF FETAL TELENCEPHALIC CELLS, BUT DID RESULT IN SEX- AND CLUSTER-SPECIFIC ALTERATION IN DEVELOPMENT TRAJECTORY AND CELL CYCLE.

We asked if an episode of PAE resulted in altered cell identity or developmental trajectory. We calculated the percentage of cells from each sample in each of the 33 clusters, neither alcohol exposure or the fetus sex affected the composition of the clusters (Two-way ANOVA dependent variable: number of cells in each sample, independent variables: exposure, sex). To study changes in the developmental trajectory, we asked if alcohol exposure and fetal sex resulted in change in the maturation trajectory as represented by the *pseudotime* score composition of each cluster. We found that 8 clusters exhibited a significant interaction effect between sex and PAE, 4 clusters exhibited a main effect of PAE and 2 clusters exhibited a main effect of sex (Figure 3-3a). Two VZ clusters, 1 SVZ cluster and two lateral layer V-VI clusters exhibited alcohol exposure-induced left-shift in pseudotime exposure, i.e., towards diminished maturity in males, but right-shift, indicative of increased maturity, in female clusters. PAE exposure resulted in right-shift, higher pseudotime scores, in two VZ clusters, but a left-shift, lower *pseudotime* scores, in one lateral Layer V-VI cluster. (Figure 3-3b and supplementary 2)

We next tested if PAE or fetal sex affected the percentage of cells assigned to each of the cell cycle phases in each cluster (Figure 3-3-a). Cluster 10 (Lateral Layer V-VI) exhibited a main effect of alcohol exposure on the percentage of cells in G1 phase, and main effect of both sex and exposure on the percentage of cells in G2M phase. There was main effect of alcohol exposure on the percentage of cells in both G1 and S phase in cluster 19 (SVZ cells) and an interaction effect of sex and alcohol exposure on the percentage of cells in G1 phase. Microglia

(cluster 30) exhibited a main effect of alcohol exposure on the percentage of cells in S phase (Figure 3-3c & Appendix 2 c).

PAE EFFECTS ON CELL DEATH POTENTIAL IN THE DEVELOPING CORTEX.

We next investigated whetsher PAE resulted in differential expression of caspases [269]).Our data shows that at GD 14.5, developing cortical cells in all clusters expressed Caspases 2, 3 and 6 (Expression levels in various clusters are shown in Appendix 2 c-b), whereas other caspases like caspase 1 were expressed in few to no cells. Out of 33 clusters, Caspase 2 was downregulated by PAE in 1 female cluster, Caspase 3 was downregulated by PAE in 2 male clusters and 3 female clusters, whereas Caspase 6 was upregulated by PAE in 1 female cluster. These data collectively show, that PAE does not have a significant and uniform effect on the potential for execution of apoptosis.

PRENATAL ALCOHOL EXPOSURE RESULTS IN DISTINCT SEX-SPECIFIC PATTERNS OF GENE EXPRESSION DYSREGULATION PRIMARILY IN MIGRATING CELL POPULATIONS.

To investigate the effects of PAE on transcription profiles in males and females, and to identify the baseline expression differences between male and female developing cortices, we utilized the *Desingle* package to identify three types of differentially expressed genes (detailed in methods section). Alcohol exposure resulted in more differentially expressed genes in female clusters compared to male clusters. A SVZ cluster and lateral layer V-VI neuron cluster exhibited the highest numbers of alcohol-dysregulated genes in females (Figure 3-4a), with majority of genes being DEg-type downregulated genes (Appendix 2 d). The highest number of dysregulated genes in male clusters were observed in rostral layer V-VI neurons. Few to no differentially expressed genes were observed in non-neuronal cell types in both males and females (Boxplots of magnitude of fold change of significant genes in each cluster are presented in Appendix 2 e).

We next investigated whether alcohol exposure dysregulates the same genes in males and females, and calculated the percentage of overlap between differentially expressed (DE) genes in PAE relative to controls, in males and females, in each cluster. Most clusters exhibited minimal overlap between DE genes in males and females (ranging from 0-31.1%, average = 6.9% overlap). However, cluster #2, with a radial glial identity exhibited the highest percentage 31.1%, of overlap between males and females (Appendix 2 d-b).

We identified frequently dysregulated genes, which are downregulated or upregulated in >50% of the clusters (Figure 3-4-b), Ubiquitin-b (Ubb) was an example of a gene with a sexually dimorphic response to PAE. Ubb transcripts were frequently upregulated in female clusters (17 clusters) but frequently downregulated in male clusters (21 clusters; Appendix 2 g). Xist was significantly downregulated by alcohol exposure in 19 female clusters (out of 33 clusters) (discussed later, Appendix 2 f)

To further understand the biological pathways enriched for DE genes, we performed IPA analysis on DE genes separately for male and female cells in each cluster and compared the significant pathways to identify common dysregulated pathways across lineages and pathways that were dysregulated in a lineage specific or sex specific pattern. Our analyses show that PAE resulted in lineage-wide disruption in translation control (EIF2, eI4) and stress/nutrient (mTOR) pathways in both male and female fetuses. Other pathways like 14-3-3 mediated signaling, epithelial adherens junction signaling and phagosome maturation were identified as significant in similar clusters in males and females, but were more dysregulated in male clusters (Figure 3-4-c). Sex-specific responses to PAE were observed at more selective stages of maturation. For instance, while pathways associated with mitochondrial dysfunction and disruption in oxidative

phosphorylation were identified in male migrating cells clusters and rostral layer V-VI clusters, these same pathways were identified as disrupted in females in Lateral Layer V-VI clusters.

Overall, interneuron clusters were among the least affected by alcohol exposure. However, interestingly one interneuron cluster, cluster 11, was specifically affected by PAE in both males and females, but in opposite directions. IPA analysis showed sex dimorphism in the pathways enriched for the differentially expressed genes, such that while EIF2, synaptogenesis, estrogen receptor signaling pathways and oxidative phosphorylation were predicted to be activated in female cells in this cluster, they were predicted to be inhibited in male cells in cluster 11 (Figure 3-4-c).

Using the Upstream regulator analysis tool in IPA, we further identified significant candidate upstream transcriptional regulators of differentially regulated genes in each of the male and female clusters, and specifically, transcriptional regulators which were frequently identified as significant upstream regulator across multiple clusters (Table 3-4). A subset of those upstream regulators (BCl11a, Ctnnb1, Htt, Appendix 2 f) was also dysregulated by PAE in some clusters.

We investigated the extent to which genes implicated in autism spectrum disorders (ASD, https://gene.sfari.org/database/human-gene/; 913 annotated genes) were dysregulated by PAE. In female clusters up to 26 % of ASD-linked genes were downregulated, mainly in VZ clusters and Lateral Layer V-VI neurons. Fewer ASD-related genes (up to 14%) were dysregulated in male clusters, in one VZ cluster and the rostral layer V-VI neuron cluster (Appendix 2 g).

GENE CO-EXPRESSION NETWORKS ARE DYSREGULATED BY ALCOHOL EXPOSURE

Modularity in gene expression patterns has been previously documented in transcriptomic and functional studies [270] and disruption in gene expression modules was shown in

neurodevelopmental and neurodegenerative disorders [271-273]. Therefore, to investigate if PAE resulted in coregulated disruption of gene expression, we constructed gene co-expression networks using weighted gene co-expression network analysis (WGCNA), and identified modules of highly interconnected genes as well as hub genes for each module. Co-expression networks were first constructed independently for the top variant 20% genes in each control male and female cluster, and 'hub-genes' for each module were also identified. We next quantified the number of genes dysregulated by PAE in each of the modules, and ranked the modules according to the proportion of dysregulated genes (Figure 3-5a).

EEF1A1-CRITICAL GENE FOR TRANSLATIONAL CONTROL- IS A HUB GENE DYSREGULATED BY PAE

Our data show that for both male and female cells, the hub genes in 4 out of the top 5 dysregulated modules, were also dysregulated by PAE. For example, Eef1a1, the fetal neural-isoform of the translation elongation factor [274, 275], was both downregulated by PAE (Appendix 2 f), and identified as a hub gene for two highly dysregulated modules in both males and females (Figure 3-5-b). Gene members of those 'Eef1a1 hub modules' (female cluster 10 blue module, female cluster 12 blue module, male cluster 3 brown module, male cluster 10 blue module) share common gene members, but also include cluster- and module-specific genes (Figure 3-5-c), and are collectively associated with ribonucleoprotein complex assembly and translation (Table 3-5). The gene members of Eef1a1-containing modules exhibited a similar PAE-related pattern of downregulation in 2 VZ clusters (clusters 6 and 12), 1 SVZ cluster (cluster 3), a lateral layer V-VI and a rostral layer V-VI clusters in both males and females (Figure 3-5-d). In contrast, several clusters exhibited sex-specific responses to PAE. For example, genes in Eef1a1-containing modules were generally downregulated in one male

interneuron cluster, but few of those genes were down-regulated in the comparable female cluster. In females, Eef1a1-assocatied genes were downregulated in response to PAE in 3 neuronal clusters belonging lateral layer V-VI (Figure 3-5).

FEMALE-SPECIFIC RESPONSES TO PAE MEDIATED BY XIST DOWNREGULATION

Female clusters showed higher number of dysregulated genes compared to male clusters. We sought to investigate possible female-specific mechanisms that explained this vulnerability. As expected Xist lncRNA was only expressed in female but not male cells (Figure 3-6-a). Our data shows that Xist was significantly downregulated by alcohol exposure in 19 out of 33 clusters (Figure 3-4-b), interestingly the Xist antagonist, Tsix, was upregulated by PAE in 7 clusters (Appendix 2 f). WGCNA also identified Xist as a hub for modules in 1 VZ cluster, 1 VZ/SVZ cluster and one Layer I cluster (6_ yellow (Figure 3-6-b), 14_yellow, 25_brown), those modules exhibited 63, 13, and 8 downregulated genes out of 91, 54 and 207 genes respectively. Genes from the above three modules exhibited little overlap, compared to Eef1a1-associated modules (Figure 3-6-c) suggesting that Xist-associated modules serve divergent functions in different cell types. Due to the high percentage of dysregulated genes in co-expression module 6_yellow, identified as VZ cluster 6, we asked if Xist-associated genes were specifically downregulated in this cluster in females or if they were co-dysregulated by alcohol exposure in other clusters in males and females. Xist modules collectively contained 352 genes, 146 of which were downregulated in female cluster 6, 186 genes downregulated in female cluster 10, 154 genes downregulated in female cluster 12 and 107 genes downregulated in female cluster 17, however, this pattern of dysregulation was not evident in any of the male clusters (Figure 3-6-d).

Since Xist was frequently downregulated in female clusters, we asked if genes belonging to Xist module in cluster 6 were also frequently dysregulated genes in other clusters. A subset of the

Xist module genes were downregulated in 29 clusters in females, whereas they were downregulated in only 7 male clusters (Appendix 2 h-a). Gene members of Xist modules, including Celf2 (downregulated in 18 female clusters), Luc7l2 and Tia2 (dysregulated in 17 female clusters), Ddx17 (downregulated in 12 female clusters), Rbm25 (downregulated in 12 female clusters) are all regulators of spliceosome assembly and alternative mRNA splicing activity (Appendix 2 h-c).

RNA splicing was a significant and common functional annotation for genes that were enriched in the Xist modules (Table 3-6). We therefore further investigated the effect of PAE on the expression of spliceosomal transcripts in male and female fetal cells. We extracted the genes falling into each of the following gene ontologies: positive regulation of mRNA splicing via spliceosome, negative regulation of mRNA splicing via spliceosome, and spliceosomal complex assembly. Our results show that VZ-type cells (clusters 3, 6 and 12) and lateral layer V-VI-type cells (cluster 10) exhibited similar dysregulation of spliceosomal function genes in female cells, but this dysregulation was not observed in males (Appendix 2 h-a). Additionally, a group of mRNA splice site selection genes, Celf2, Luc7l2 and Luc7l3, were downregulated in more than 10 female clusters but not in any male clusters. Luc7l3 specifically was upregulated in 12 male clusters (Appendix 2 h-b).

Since X-chromosome inactivation in females is maintained by the expression of Xist lncRNA [111], we expected that the downregulation of Xist lncRNA will result in the loss of X-inactivation and the upregulation of genes harbored within the inactivation region. We surveyed the effect of alcohol exposure on the expression level of the X-chromosome genes in females and males, as expected we found a pattern of x-genes upregulation in the female clusters compared to male clusters (Appendix 2 i). We asked if the upregulated genes were specifically located within

X-inactivation regions [276, 277], and found that upregulated x-genes (specifically genes frequently upregulated across cluster) were indeed located within those regions (Figure 3-6-e). Tmsb4x, an example of an upregulated gene, is implicated in enhanced cell proliferation in cancer cells [278].

Cis-regulatory effects of Xist lncRNA on X-chromosome inactivation are well established. However, it is not known whether Xist expression may also be associated with autosomal gene repression. We therefore investigated the association between Xist transcript expression and autosomal repression by assessing autosomal genes that were significantly negatively correlated with Xist in each of the 33 graph-based clusters. In each cluster we identified the chromosomal location of the genes showing significant negative correlation with Xist, and quantified the number of significantly negatively correlated genes on each chromosome. Our results show a high frequency of autosomal genes (and high normalized value, relative to the number of genes expressed from a chromosome) significantly negatively correlated with Xist in Clusters 1,14 and at a lower frequency, in cluster 6. All three clusters were characterized by high percentage of cells in S and G2M phases. Our analyses identified chromosomes 2, 4, 5, 7 and 11 as harboring the highest frequency of genes whose expression was inversely correlated with Xist expression (Appendix 2 j). Autosomal genes frequently showing significant negative correlations with Xist were preferentially implicated in translation control.

3.5 DISCUSSION

In this study we used scRNAseq transcriptomics analysis, which facilitates characterizing different cell-types based on transcriptional profiles of individual cells [122, 245], to assess the impact of a single, temporally restricted episode of PAE at gestational day 12.5. In mice, this

exposure period corresponds to the emergence of the subventricular zone (SVZ) and the birthdate of the early neurons the cerebral cortical pre-plate (layer I and VI) and cortical plate Layer V neurons [241-243, 279]. We assessed cerebral cortical development after 48 hours, i.e., at GD14.5, corresponding to the completion of 50% of the founder-population cell cycles required for cortical neurogenesis [243] and the migration of Layer V neurons to their appropriate position within the cortical plate [241]. Our analysis showed that there is a high degree of cellular heterogeneity associated with each assessed cortical laminae and that this cell-population diversity was replicated across six separate pregnancies, and in both male and female fetal cortices. For example, we documented 15 sub-populations of immature neural progenitors with identities related to the VZ and SVZ. The major VZ cluster #1 could itself be sub-divided into 8 additional sub-clusters suggesting that at least 24 different neural precursor subtypes contribute to the neuroepithelium of the VZ and SVZ at this developmental stage. We also identified 10 distinct neuronal sub-populations with identities related to layer V, VI and I, as well as four sub-populations of cortical interneurons.

Pseudotime analysis ordered identified cell clusters along a temporal continuum, suggesting that these molecularly heterogenous clusters represented stages along a developmental continuum. Consistent with this finding, developmental stage-selective mRNA transcripts [261, 280]like radial glial markers Pax6 and Gas1 were preferentially expressed during early *pseudotime*, neurogenic markers Neurog2 and Eomes at mid-*pseudotime*, and post-mitotic neuronal markers Tbr1 and Plxna4 during late-*pseudotime*. At this developmental stage, there were relatively few cells identified as having endothelial or microglial identities, and none with mature astrocyte or oligodendrocyte identity. Cell-proliferative capacity was mainly relegated to the more immature cell-types, though residual proliferative capacity was observed among two interneuron

populations. VZ clusters differed in numbers of cells assigned to the mitotic stage suggesting possible differences in spatial localization to apical (clusters 1.6, 21 and 6 with cells predominantly in mitosis) compared to basal VZ (clusters 2 and 12 with cells predominantly in S-phase), and potentially, different VZ microenvironments. Interestingly, cluster 14 which was identified as an intermediate stage between VZ and SVZ had an equivalent number of mitotic-type cells as VZ clusters 21 and 6, and possibly represent an asymmetrically dividing VZ fraction, destined to exit the VZ towards the SVZ. Importantly, we did not observe any significant differences in cell population distribution due to fetal sex. Both male and female fetal cortices exhibited identical patterns of cellular heterogeneity.

PAE has been well documented to result in microcephaly in both human populations [72, 239, 281] and in animal studies [158, 240]. Consistent with our previous observations in ex vivo model systems [86, 162], we did not find evidence for extensive cell death due to ethanol exposure. For instance, there was no loss of cell populations as defined by cluster identity, and the numbers of cells in each cluster were not altered due to PAE. Surprisingly, no cells expressed caspase-7 and some cell clusters with SVZ/transient progenitor identity also expressed low to undetectable levels of Caspases 3 and 6 which have all been defined as identified as an final common effector caspases [282], suggesting overall developmental-stage related resistance to apoptosis induction. Moreover, there were no changes in mitochondrial gene reads and either no change or decreased expression of caspase gene transcripts due to PAE, suggesting that apoptosis and the potential for cell death remained unaffected by exposure. However, we found evidence that PAE resulted in sex-independent increases in the proportion of cells in S-phase in one SVZ subpopulation, consistent with previously published reports [86, 161, 283]. PAE also did result in shifts in *pseudotime* scores across multiple immature and mature cell type, suggesting an

impact on maturation. However, the effect of PAE on pseudotime shifts was also sex-dependent. In SVZ cluster #13 and lateral layer V/VI neurons for example, fetal male cells experienced a *pseudotime* delay whereas females exhibited a *pseudotime* advance, suggesting that ethanol inhibited maturation in males, but facilitated maturation in identical female cells. Moreover, an analysis of differentially expressed genes showed that PAE dysregulated more genes in female compared to male cells across developmental stages, and in general, a majority of genes in both sexes were down-regulated rather than upregulated by PAE.

Despite the discrepancy in numbers of dysregulated genes between males and females, pathway overrepresentation analysis showed male and female cells exhibited some common vulnerabilities in pathways related to inhibition of translation initiation suggesting that sex-differences in gene responses to PAE could nevertheless result in some common functional outcomes. WGCNA analysis further emphasized the importance of inhibition of translation as a key response to PAE, since translation initiation genes were hub genes in a number gene modules in both males and female. Our data identified Eef1a1, the eukaryotic translation elongation factor, as a downregulated hub gene in highly dysregulated modules in four cell clusters, whose members were associated with ribosomal biogenesis, translational processes and apoptosis. Eef1a1-associated genes exhibited a similar pattern of dysregulation in migrating cells and rostral layer V/VI neurons in both males and females, but sexually dimorphic pattern of dysregulation in SVZ/VZ progenitors and later layer V/VI neurons and interneurons. Additionally, differentially expressed genes due to PAE in both male and female cortices were significantly enriched for pathways related to Eukaryotic Initiation Factor 2 (eIF2) and mTor (mammalian target of rapamycin, an inhibitor of translation [284]) in most of the clusters. These data are consistent with previous reports showing that that ribosomal biogenesis and disruption

in translational processes are targets of PAE [285]. Perturbations in global protein synthesis and dysregulation of mTor have also been implicated in ASD and other neurodevelopmental disorders (reviewed in [286]) suggesting that perturbation of translation control is an important component of a number of disorders of development.

A key finding in our exploration of the basis of sex differences in the PAE response was our identification firstly, that the X-chromosome inactivation factor Xist was a hub gene in a number of gene co-expression modules and was itself down-regulated in a number of immature and mature cortical cell-types. Co-incident with loss of Xist, we found that the antisense ncRNA Tsix, which serves as an inhibitor of Xist [287, 288], was up-regulated, along with a number of X-chromosome genes within the X-inactivation region, indicating epigenetic silencing of Xist and loss of X-inactivation. These data suggest that in females, loss of X-chromosome inactivation represents an important and sex-specific response to PAE. We found that gene members of female, Xist hub-containing modules that were also downregulated by PAE were preferentially related to RNA splicing. Cortical development, lamination and cell fate are influenced by alternative splicing events, and important neural transcription factors exhibit maturation-stage specific isoforms (reviewed in [289]). Spliceosome deficits have also been identified in neurodevelopmental [290, 291] and craniofacial disorders ([292]. These outcomes are consistent with some previous transcriptomic studies in both mouse [150] and avian [293] models of prenatal alcohol exposure which have also implicated spliceosome function as an important developmental target of ethanol. The Xist locus appears to be particularly vulnerable to developmental pathology. Other environmental exposures during development, for example, exposure to bisphenol-A, have been shown to result in loss of X-chromosome inactivation, which may mediate the neurodevelopmental effects of these chemical teratogens [116]. X-

chromosome skewness (non-random X chromosome inactivation) in the absence of chromosomal mutations has been reported in females with autism [113], X-linked mental retardation [114], and Rett syndrome [115]. Therefore, loss of X-chromosome inactivation may contribute to etiology of a number of neurodevelopmental disorders. Moreover, while Xist is known to act *in cis*, to inactivate the X-chromosome, our data show that loss of Xist is also associated with the induction of autosomal-encoded genes across a number of chromosomes, suggesting that Xist also suppresses gene expression *in trans*. The mechanisms underlying *trans*-inhibition are unknown, but appear to be restricted to VZ and SVZ progenitors, suggesting that these developing cell stages may be particularly vulnerable to widespread dysregulation of gene expression following loss of Xist.

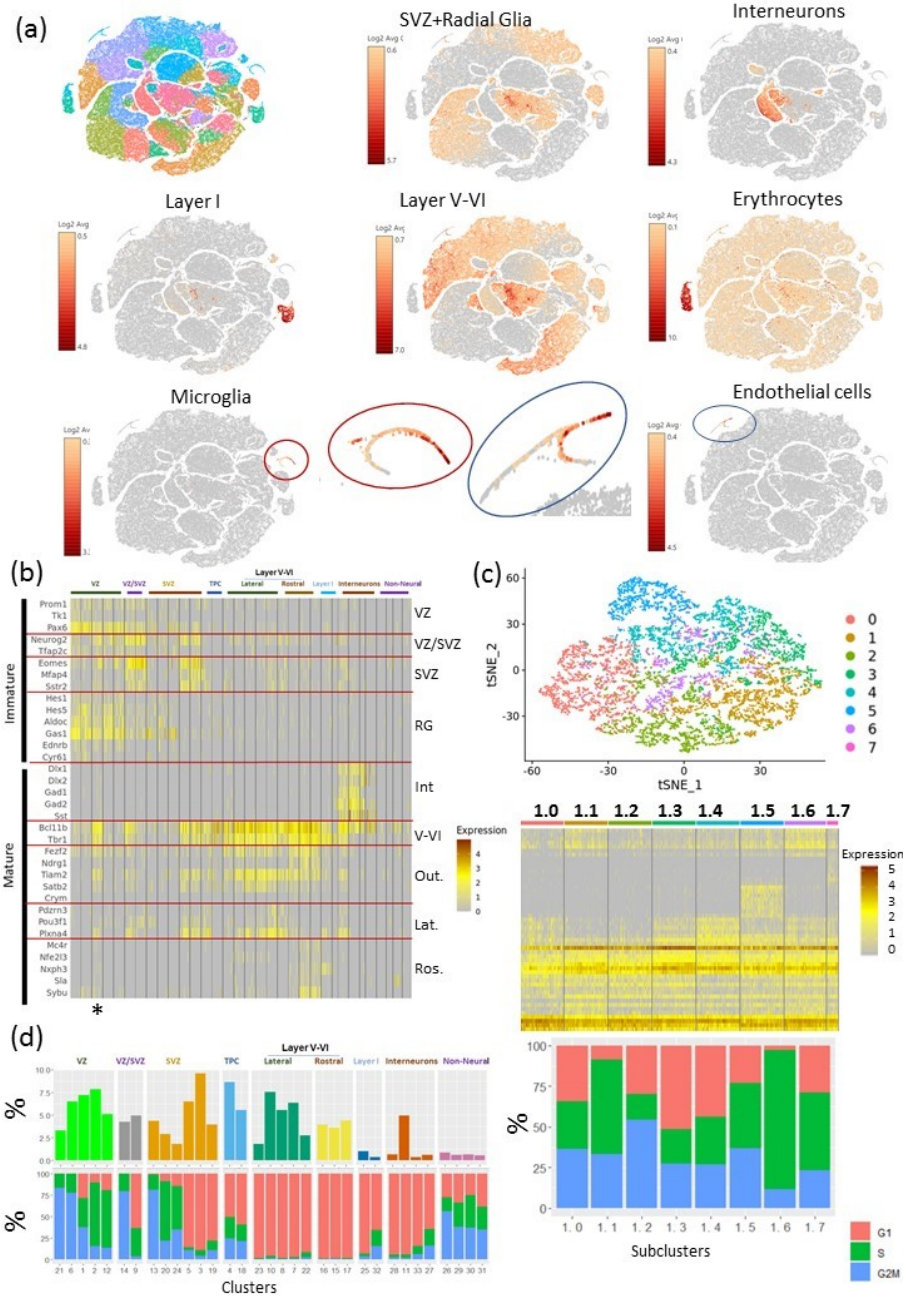
Another key finding was that hemoglobin genes, embryonic beta-like subunit (Hbb-y) in females and multiple alpha (Hba-a1, Hba-a2), beta (Hbb-bs), and embryonic beta-like (Hbb-y) subunits in males, were one of the few gene classes that were consistently and significantly upregulated by PAE at multiple stages of neural maturation. Hemoglobin gene expression has previously been documented in neurons [294]. Importantly, beta-hemoglobin subunits have been shown to be present within the nucleus of neural cells, associated with nuclear histones, and to increase histone H3 trimethylation (H3K4me3, [295]), an epigenetic activation mark, which along with histone methyltransferases which has also been documented to be increased following PAE in rodent models [296, 297]. This evidence raises the intriguing possibility that neural expression of hemoglobin genes, and beta globin genes in particular, may mediate some of the epigenetic effects of PAE. Moreover, BCL11a, a member of the Swi/Snf chromatin remodeling complex [298], and inhibitor of embryonic and fetal hemoglobin genes [299-301], was itself inhibited by PAE, in both males and females, in layer V/VI neurons as well as in other neural clusters.

BCL11a haploinsufficiency in mice results in microcephaly, impaired cognitive and social behaviors [302], and in humans, microdeletions within the BCL11a gene locus result in both retention of fetal hemoglobin and neurological symptoms like developmental delay and autism [303, 304]. These data suggest that diverse congenital neurodevelopmental disabilities may share, at least in part, a common etiology. This hypothesis is further supported by our observation that PAE resulted in persistent downregulation of up to 26 % of ASD-linked genes in females and up to 14% of ASD genes in males in both VZ cells and layer V/VI neurons.

In summary, our data shows that a common developmental perturbagen, alcohol, administered for a brief interval during the initiation of fetal cortical neurogenesis has profound and persistent consequences on gene expression patterns, particularly controlling RNA translation in progenitor cells and daughter neurons that were born during the exposure period. Moreover, the effects of PAE were sex-dependent, with female cortical cells exhibiting an exaggerated transcriptomic response compared to male cells, that was linked to loss X-chromosome inactivation. The loss of Xist was linked to *cis* de-repression of X-chromosome genes and was associated with repression of RNA splicing machinery. However, in a select group of VZ and SVZ progenitors, we also observed a de-repression *in trans*, related to activation of genes located on somatic chromosomes, suggesting that for progenitor-type cells, loss of X-inactivation may have more wide-ranging consequences of gene regulation. Finally, we identified an elevated hemoglobin gene response including elevated embryonic hemoglobin gene expression related to loss of the epigenetic repressive factor, BCL11a. These data suggest that the fetal response to PAE shares similarities with other congenital disorders of neural development.

3.6 FIGURES

FIGURE 3-1: SCRNA SEQUENCING REVEALS THE HETEROGENICITY OF ED 14.5 MURINE DEVELOPING CORTEX

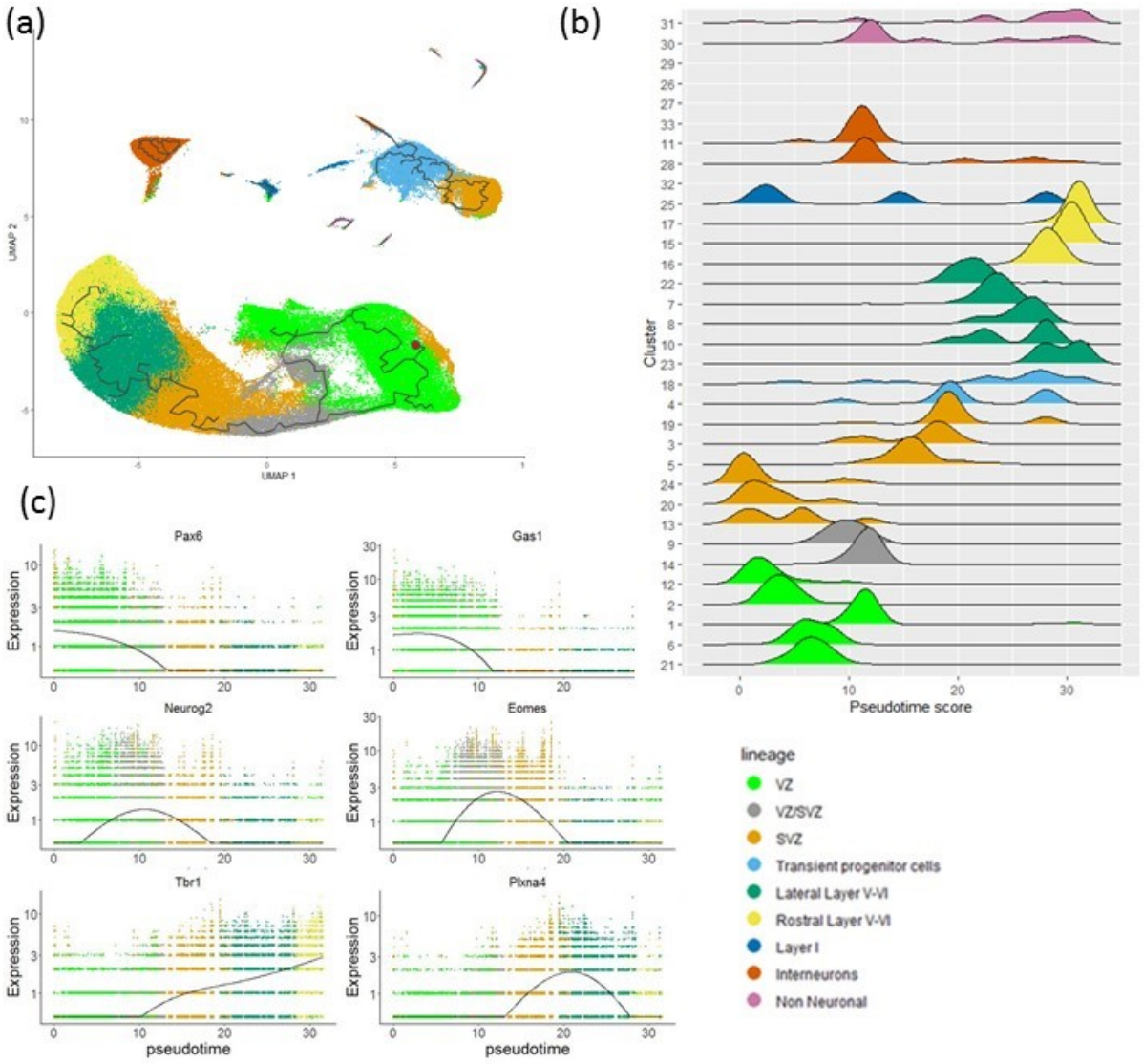


a) t-SNE plots of the sequenced cells showing the breakdown of the cells to 33 clusters and the mean expression levels of markers of different cell types, red and blue eclipses show magnified microglia and endothelial cell populations respectively (b) Heatmap showing expression of markers of mature and immature cell types in various cell types, black lines separate clusters,

Figure 3.1 continued:

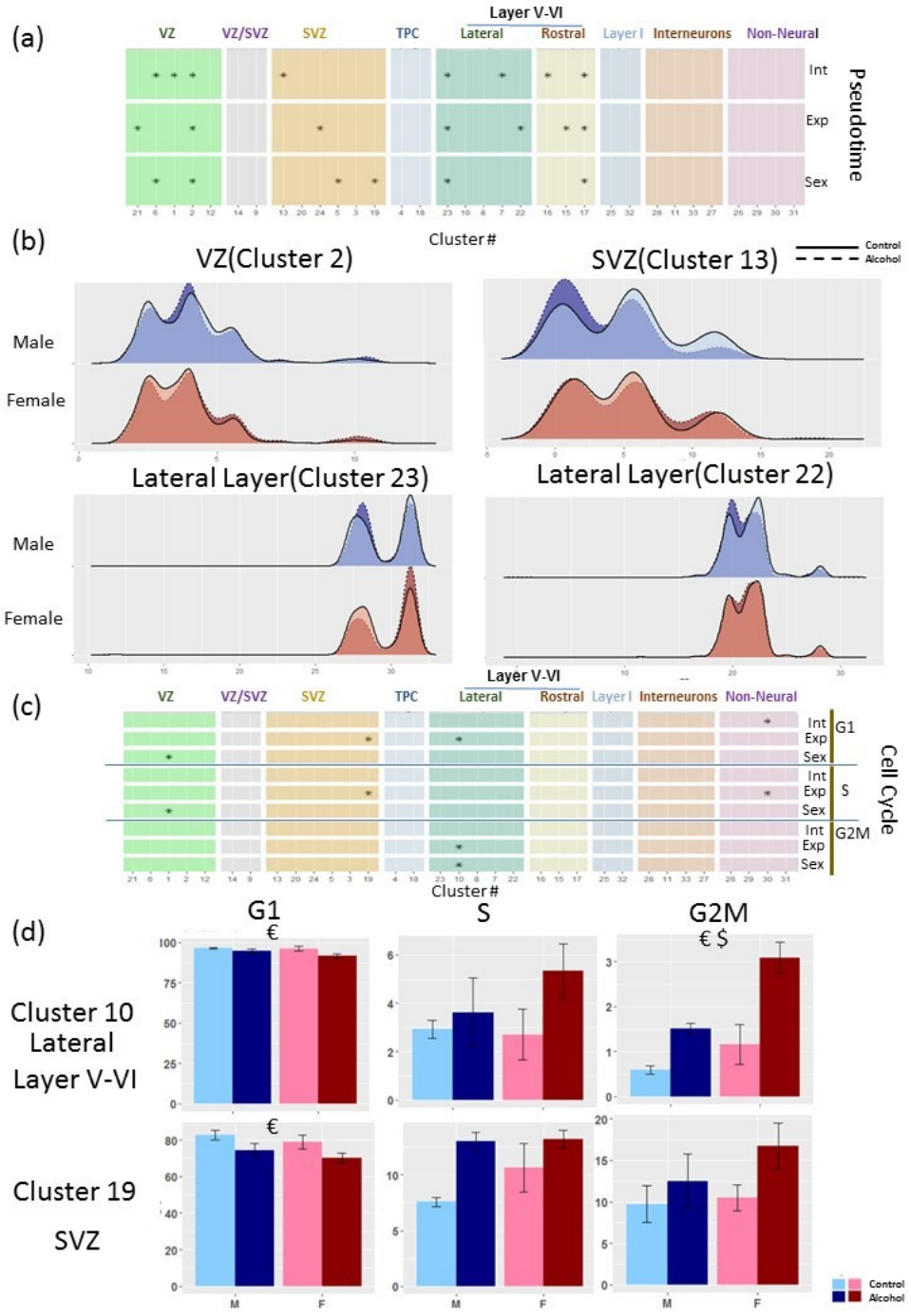
horizontal lines separate major cell subtypes, red lines separate markers of each major subtype. Each vertical line (yellow shades) represents an individual cell and the color represents the z score of the individual cell expression level. * marks cluster 1. VZ: Ventricular Zone, SVZ: Subventricular zone, RG: Radial Glia, Int: Interneurons, V-VI: Layer V-VI, Out.: Outer layer V-VI, Lat.: Lateral Layer V-VI, Ros.: Rostral Layer V-VI. (c) t-SNE plot showing the breakdown of cluster 1 into 8 subclusters, heatmap showing the expression (z-scored) of main subclusters markers in each subcluster (complete heat map presented in Appendix 2) and bar chart represents the breakdown of each subcluster into the three phases of cell cycle (d) Barchart shows the percentage of cells in each cluster (top panel), and the percentage of cells in each of the three cycle phases in each cluster (bottom panel)

FIGURE 3-2: PSEUDOTIME ANALYSIS ORDERS THE CELLS BASED ON SEQUENCE OF GENE EXPRESSION CHANGES THROUGH NEURONAL MATURATION PROGRESSION



(a) U-MAP of sequenced cells, color coded by clusters' subtypes. Grey line represents the structure of the trajectory graph, red circle represents the root of the trajectory (b) Ridge plot of pseudotime scores distribution in each of the clusters (c) Expression level of each of Pax6 (VZ marker), Gas1 (RG marker), NeuroG2 (VZ/SVZ marker), Eomes (SVZ marker), Tbr1(neuronal marker) and Plxna4 (axonal marker) across the pseudotime score continuum

FIGURE 3-3: EFFECTS OF PRENATAL ALCOHOL EXPOSURE ON CELLS TRAJECTORIES AND PHASES

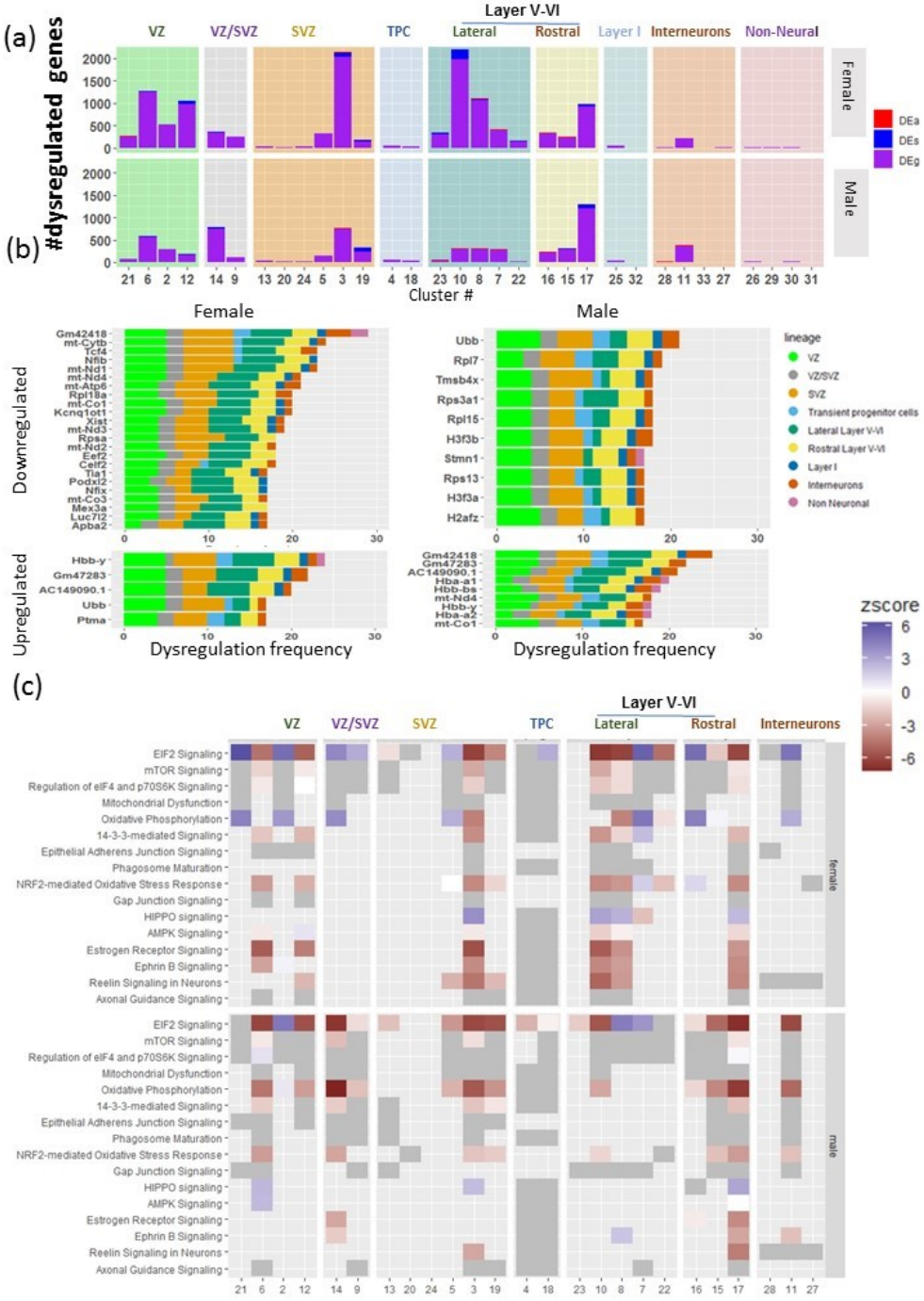


a) figure shows clusters with significant main or interaction effects of sex and alcohol exposure (p values<0.05 are noted by *) on pseudotime scores.(b) density plots showing the distribution of pseudotime scores in select male (blue) and female (pink) control (solid line) and alcohol

Figure 3.3 continued:

exposed (dotted) clusters.(c) figure shows clusters with significant main or interaction effects of sex and alcohol exposure (p values<0.05 are noted by *) on cell cycle phases.(d) Percentage of cells in each of the three cycle phases in select control (light color) and alcohol (dark color) males (blue) and females (red) clusters (€ : Indicates significant main effect of exposure , $ indicates significant main effect of sex)

FIGURE 3-4: PRENATAL ALCOHOL EXPOSURE RESULTS IN GENE DYSREGULATION IN MURINE DEVELOPING CORTEX IN LINEAGE AND SEX SPECIFIC MANNER



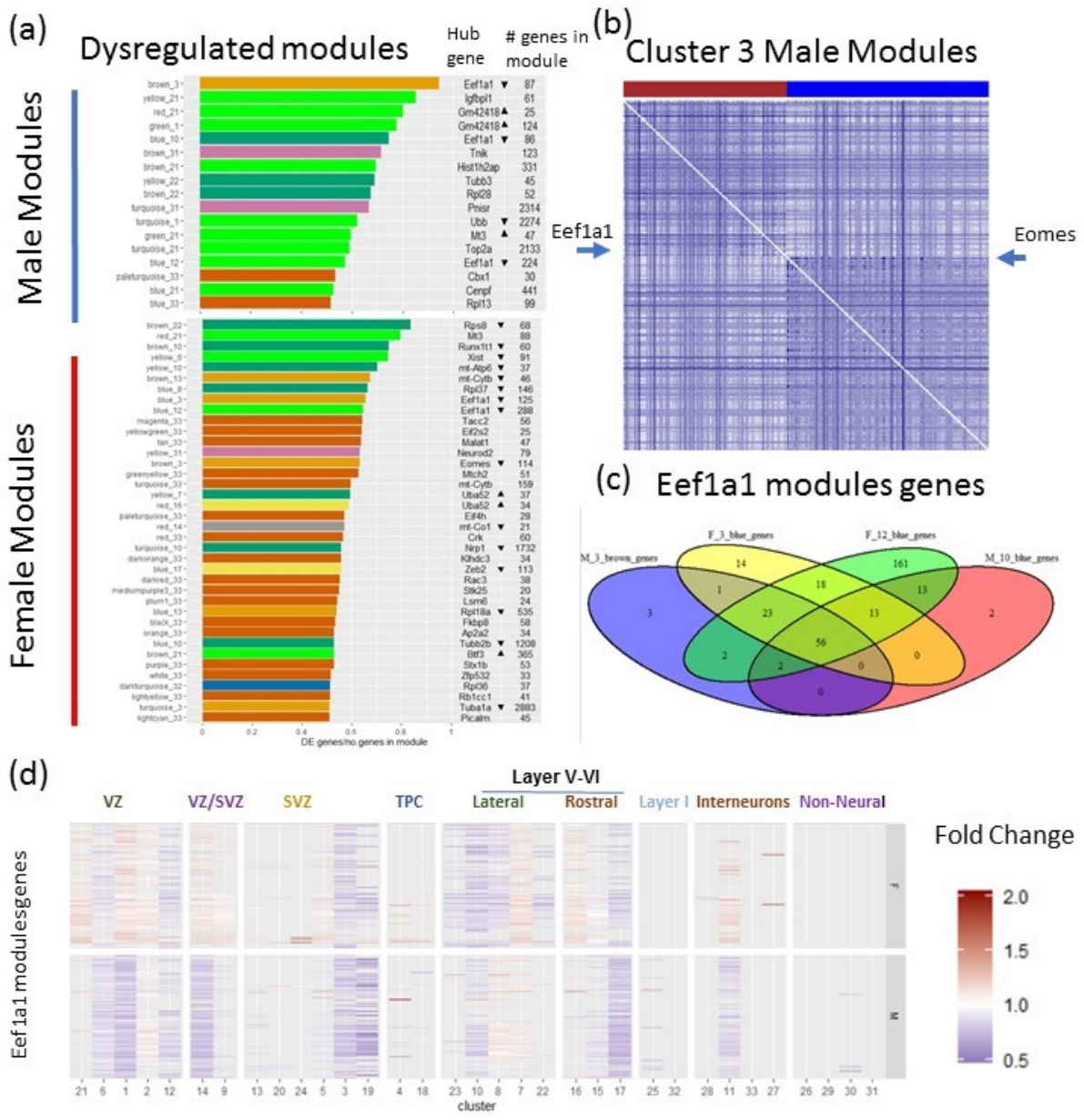
(a) bar charts showing the number of dysregulated genes by alcohol exposure in each female (top panel) and male (bottom panel) cluster, and the breakdown of the dysregulated genes into three

Figure 3.4 continued:

dysregulation subtypes. (b) bar charts showing the number of female (left) and male (right) clusters(x-axis) in which each of the genes (on y-axis) was downregulated (top panels) or upregulated (bottom panels), and the bars are color coded by major cell lineages breakdown.

c)Figure showing pathways enriched in the dysregulated genes in each female (top panel) and male (bottom panel) clusters. Significant pathways (B-H p-value <0.05) are marked by boxes, grey color indicates no known direction of activation, color scale indicates the z-score of activation

FIGURE 3-5: PRENATAL ALCOHOL EXPOSURE DISRUPTS GENE EXPRESSION MODULARITY

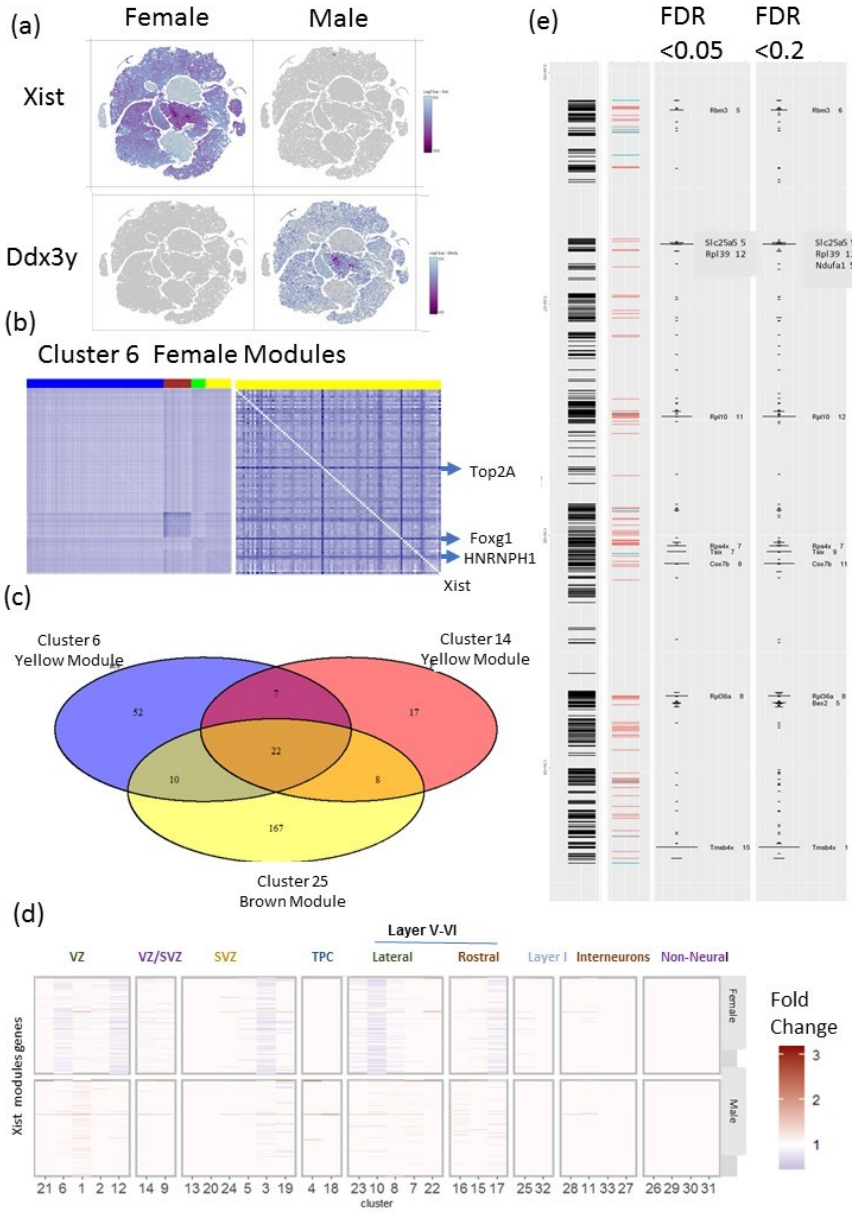


(a) Bar charts show top dysregulated gene modules(y-axis) by alcohol exposure in male (top panel) and female (bottom panel) clusters, color coded by cluster lineage, hub gene and number of genes constituting each module is shown, arrows indicate that hub gene is significantly downregulated (downwards arrow) or upregulated (upwards arrow) by alcohol exposure in the respective cluster. Ratio of module dysregulated genes relative to total number of genes is shown

Figure 3.5 continued:

on x-axis. (b) Topological overlap matrix of gene members of brown and blue modules of cluster 3, arrows show Eef1a1 and Eomes genes, the hub genes of brown and blue modules respectively. Light colors represent low topological overlap and progressively dark blue represents higher topological overlap. (c) Venn diagram shows the overlap between gene members of Eef1a1 module genes in female and male cluster 2, female cluster 12 and male cluster 10 (d) Heatmap showing Eef1a1 module genes (individual vertical lines) regulation by alcohol exposure in female (top panel) and male (bottom panel) clusters. Color legend indicates the fold change of dysregulation.

FIGURE 3-6: FEMALE SPECIFIC RESPONSES TO PRENATAL ALCOHOL EXPOSURE ARE MEDIATED BY ALCOHOL INDUCED DISRUPTION OF THE FEMALE SPECIFIC XIST GENE



(a) t-SNE plot of sequenced female (left) and male (right) cells color coded by the expression level of the female specific Xist (top) and the male specific Ddx3y (bottom)(b) Topological overlap matrix of gene members of blue, brown, green and yellow modules

Figure 3.6 continued:
of cluster 6 (left), and zoomed matrix showing yellow matrix genes dissimilarity matrix, Xist hub gene and arrows showing key highly connected genes. Light colors represent low topological overlap and progressively dark blue represents higher topological overlap.(c) Venn diagram shows the overlap between gene members of Xist module genes in female cluster 6 and 14 yellow modules and cluster 25 brown module(d) Heatmap showing Xist module genes (individual vertical lines) regulation by alcohol exposure in female (top panel) and male (bottom panel) clusters. Color legend indicates the fold change of dysregulation.(e) Figure showing the location of X-chromosome expressed genes in ED 14.5 murine developing cortices (first vertical panel) , the location of X-inactivated genes (orange) , genes that escape x inactivation (blue) data derived from (Berletch et al., 2010; Carrel and Willard, 2005) (second vertical panel) , and the location of X chromosome-genes upregulated by alcohol exposure in female clusters(third and fourth vertical panels , showing FDR threshold of 0.05 and 0.2 respectively) , width of the line indicate the number of clusters in which the respective gene is upregulated , genes upregulated in more than 5 clusters are labelled by gene name and frequency of upregulation.

3.7 TABLES

TABLE 3-1: DAM AND SAMPLE IDS AND MATERNAL BLOOD ALCOHOL LEVEL OF EACH SAMPLE

Dam ID	**Maternal Blood Alcohol level**	**Sample IDs**
9	260 mg/dl	PAE_male_1 PAE_female_1
10	0 mg/dl	control_male_1 control_female_1
12	250mg/dl	PAE_male_2 PAE_female_2
2	0 mg/dl	control_male_3 control_female_3
1	210 mg/dl	PAE_male_3 PAE_female_3
8	0 mg/dl	control_male_4 control_female_4

TABLE 3-2: PERCENTAGES OF CELLS IN EACH CELL LINEAGE

Lineage	**Percentage of Cells**
VZ	24.91
VZ/SVZ	8.08
SVZ	22.1
Transient progenitor cells	8.57
Lateral Layer V-VI	17.87
Rostral Layer V-VI	9.66
Layer I	1.29
Interneurons	5.73
Non-Neural	1.79

TABLE 3-3 GENE ONTOLOGY ENRICHED GENES SUBCLUSTERS OF CLUSTER 1

Subcluster	Key GO biological processes for subcluster marker genes	Percentage of cluster 1 cells
1.0	Positive regulation of cell morphogenesis involved in differentiation; Synapse Organization	19.16
1.1	Mitotic spindle assembly; Ribonucleoprotein complex subunit organization	17.89
1.2	Regulation of axon extension involved in axon guidance; Chemotaxis	15.98
1.3	microtubule cytoskeleton organization; Apoptosis	13.09
1.4	regulation of synaptic transmission, glutamatergic; positive regulation of excitatory postsynaptic potential; positive regulation of filopodium assembly	12.36
1.5	Regulation of membrane potential; cerebral cortex GABAergic interneuron differentiation; Synapse organization	10.68
1.6	Positive regulation of cytokinesis; Regulation of mitotic cell cycle	8.52
1.7	Regulation of axon extension/Axon guidance; Regulation of migration/localization	2.33

TABLE 3-4: SIGNIFICANT UPSTREAM REGULATOR ACROSS MULTIPLE CLUSTERS

Female		**Male**	
Transcriptional regulator	# of Clusters	Transcriptional regulator	# of Clusters
YAP1	23	BCL11A	21
CTNNB1	22	HTT	21
MYC	22	MYC	20
HTT	21	MYCN	20
NFKBIA	21	YAP1	20
YY1	21	CTNNB1	19
MLXIPL	20	MECP2	19
MYCN	20	PML	19
TP53	20	TP53	19
BCL11A	19		

TABLE 3-5:IPA FUNCTION ANNOTATION FOR EEF1A1 COMBINED MODULES GENES (586 GENES)

Function annotation	p-value	# Molecules
Metabolism of protein	**3.02E-54**	**121**
Synthesis of protein	**2.25E-68**	**98**
Expression of RNA	**1.70E-11**	**95**
Translation	**3.10E-82**	**94**
Apoptosis	**1.35E-07**	**94**

TABLE 3-6:IPA FUNCTION ANNOTATION FOR XIST COMBINED MODULES GENES (352 GENES)

Function annotation	p-value	# Molecules
Transcription	**2.65E-19**	**101**
Expression of RNA	**6.87E-19**	**104**
Transcription of RNA	**3.97E-17**	**87**
Transcription of DNA	**1.88E-16**	**77**
Splicing of RNA	**2.09E-16**	**29**
Organismal death	**1.70E-15**	**101**
Activation of DNA endogenous promoter	**1.35E-14**	**64**
Processing of RNA	**2.73E-14**	**33**
Processing of mRNA	**1.04E-11**	**24**
Perinatal death	**1.40E-11**	**36**
Splicing of mRNA	**6.04E-11**	**21**

4 CONCLUSIONS

Prenatal alcohol exposure results in the disruption of the orchestrated process of proliferation, differentiation, and migrating of neurons [160, 161], resulting in cortical dysgenesis [305], heterotopias, persistent brain growth deficits [73], microcephaly [72], and other deficits collectively termed fetal alcohol spectrum disorders (FASD).

In chapters 1 and 2, I presented mechanisms of alcohol exposure mechanisms on brain development with particular emphasis on long non-coding RNAs, and special emphasis on sex-specific mechanisms in chapter 2. In the first chapter, I utilized neural stem cell culture model, in the second chapter I utilized an animal model of prenatal alcohol exposure. Our results not only explored the effects of alcohol but also explored the basic mechanisms of RNA biology.

In the first chapter, we presented an often-under-investigated class of long non-coding RNAs-pseudogenes-. Studying pseudogenes present a challenge due to the high sequence homology to a protein-coding gene. Previously reported Oct4/Pou5f1 expression in tissues might have been a false signal of pseudogenes [306]. The importance of pseudogene studying exceeds the need to avoid a false signal; previous data and our data show that they indeed exert an essential role in the regulation of genes expression through various mechanisms[307], and here we reported an Oct4 pseudogene is a target of prenatal alcohol exposure and mediates some of the effects of alcohol on neural stem cells. We showed that it associates with specific miRNAs and modulate their availability. We show that Oct4 pseudogenes are tissue-specific. We studied the NSCs specific Oct4 pseudogene mechanisms of mediating alcohol exposure; however, there is an unanswered question of whether alcohol affects the expression of Oct4 pseudogenes in other tissues, i.e., more developed cell types in the brain and whether their expression in mature tissues is the repressing signal for Oct4 repression. Our analysis of genes affected by the manipulation

of Oct4 pseudogene in the neural stem cell culture model was based on known gene markers for stemness and differentiation. However, a more comprehensive analysis using RNA sequencing after pseudogene manipulation would have allowed for the construction of gene networks affected by the pseudogene manipulation and created a list of gene targets whose promoters can further be explored for gene-specific methylation and hydroxy-methylation changes due to pseudogene manipulation.

Mining of publicly available data is an approach adopted in this chapter and was crucial in enriching the data. I started the project by designing primers for potential Oct4 pseudogenes loci and testing for their expression on RNA isolated for neurosphere cultures, which lead to the identification of a previously unannotated pseudogene, with no evidence of expression of other Oct4 pseudogenes in the neurosphere culture model. The following question to answer was if there was evidence of other Oct4 pseudogenes' expression in other cell types and tissues. To answer this question, I used publicly available RNA sequencing data from the ENCODE project [308, 309]for select tissues from C57BL/6NJ mice. I applied a pipeline to identify novel un-annotated RNA transcripts in which I mapped reads to the mouse genome (mm10), followed by transcript assembly and assessment for alignment with genomic sequences for Oct4 pseudogenes. I used a similar pipeline to mine data from my single-cell RNA sequencing of neurospheres for unannotated transcripts of Oct4 pseudogenes.

In chapter 2, I moved from the neurosphere cell culture model to in vivo animal model of alcohol exposure. As mentioned before, the developing brain cells are heterogeneous, and many orchestrated processes take place during its development. I sought to utilize the rapidly developing field of single-cell RNA sequencing to study the effects of prenatal alcohol exposure on the developing cortex on a cell type-specific basis. Additionally, I designed the experiment to

study sex differences in those effects while eliminating the maternal environment's influence by utilizing tissues from male and female fetuses originating from the same dam for each sample. This study provides a comprehensive catalog of the cell type-specific transcriptomes of the developing brain at embryonic day 14.5; this catalog provides the opportunity to study baseline (no alcohol exposure) sex differences in the developing brain's transcriptome. Another output of the study is the ability to use each cell's transcriptome to build gene co-expression networks and identify hub genes that are highly connected with gene members of this network. In this study, we identified gene networks in each cluster in each of the control males and females, and using the differential expression data; we were able to identify modules highly dysregulated by prenatal alcohol exposure and identified critical dysregulated hub genes. A future direction for these results will be to validate that manipulating the identified dysregulated hub gene in a specific cell type results in the dysregulation of the co-expression networks and identifying the phenotypes resulting from those manipulations. A Database containing transcriptional signatures from thousands of perturbagens (L1000) [310] have been used to identify potential therapeutics to reverse transcriptomic signatures resulting from substances of abuse [311]; similar approach can be applied to the cell type-specific PAE transcriptomic signatures and identify potential drugs that can rescue those changes. This approach presents more challenges in the developing brain than the adult brain due to the continuous developmental changes occurring, which raises the need for tight selection of the treatment window of the selected therapeutics.

In chapter 2, I identify a female-specific mechanism of alcohol-induced transcriptional dysregulation in the mouse developing cortex mediated by Xist lncRNA. I show that it is downregulated by alcohol exposure in most cell types, and it acts as a hub for a co-expression network mainly implicated in the spliceosomal activity. I further investigated transcriptional

outcomes of Xist downregulation; I hypothesized that Xist downregulation would result in the upregulation of genes on the X-chromosome that are ordinarily subject to X-inactivation by Xist lncRNA, the differential expression analysis supported this hypothesis, and I identified a group of X-linked genes on regions subject to X-inactivation upregulated by alcohol exposure. We identified an additional female-specific mechanism of alcohol-induced dysregulation of transcription in which the loss of Xist lncRNA results in the upregulation of autosomal genes in three mitotic VZ/SVZ clusters. Xist lncRNA's action to repress autosomal genes was previously shown [312]; however, teratogen induced dysregulation in this mechanism has not been shown.

The power of single-cell RNA sequencing is determining cell-type-specific transcriptional profiles and identification of cell type-specific responses to a perturbagens (prenatal alcohol exposure in this case). However, cell types do not act in isolation of other cell types; an effect on a cell type can result in a response in different cell types through cell-cell communication; single-cell RNA sequencing does not capture the regionality and localization of cell types. A future direction in which the project can be expanded is to utilize the developing technique of spatial transcriptomics, which uses single-cell RNA sequencing advancements but retains the cellular location information. The resulting data from using this technique will help determine waves of transcriptional responses that extend beyond a cell type but spread to neighboring cells.

The dataset presented in chapter 2 provides opportunities for a lot of future analyses. The sequencing output retains information about the splicing status of mRNA; unspliced mRNA predicts the fate of a cell using RNA velocity [313]. While I did not apply this approach in my dissertation, it is a future direction of this dataset; it will answer how prenatal alcohol exposure alters developing cortex cell fate beyond the ED14.5 time point.

Tissue from ED14.5 developing cortices analyzed in chapter 2 was retained for future single-cell ATAC-sequencing, allowing for the identification of active transcription factor mediating cell fate of each cell type [314] mediating the transcriptional responses to prenatal alcohol exposure.

REFERENCES

1. Savchev, S., et al., *Neurodevelopmental outcome of full-term small-for-gestational-age infants with normal placental function.* Ultrasound Obstet Gynecol, 2013. **42**(2): p. 201-6.

2. Mallard, C., et al., *Reduced number of neurons in the hippocampus and the cerebellum in the postnatal guinea-pig following intrauterine growth-restriction.* Neuroscience, 2000. **100**(2): p. 327-333.

3. Padilla, N., et al., *Differential effects of intrauterine growth restriction on brain structure and development in preterm infants: a magnetic resonance imaging study.* Brain Res, 2011. **1382**: p. 98-108.

4. Ke, X., et al., *Uteroplacental insufficiency affects epigenetic determinants of chromatin structure in brains of neonatal and juvenile IUGR rats.* Physiol Genomics, 2006. **25**(1): p. 16-28.

5. Vazquez-Gomez, M., et al., *Sex and intrauterine growth restriction modify brain neurotransmitters profile of newborn piglets.* Int J Dev Neurosci, 2016. **55**: p. 9-14.

6. Yamoto, K., et al., *De novo IGF2 mutation on the paternal allele in a patient with Silver-Russell syndrome and ectrodactyly.* Hum Mutat, 2017. **38**(8): p. 953-958.

7. Sachwitz, J., et al., *NSD1 duplication in Silver-Russell syndrome (SRS): molecular karyotyping in patients with SRS features.* Clin Genet, 2017. **91**(1): p. 73-78.

8. Muurinen, M., et al., *Hypomethylation of HOXA4 promoter is common in Silver-Russell syndrome and growth restriction and associates with stature in healthy children.* Scientific Reports, 2017. **7**(1): p. 15693.

9. Noeker, M. and H.A. Wollmann, *Cognitive development in Silver-Russell syndrome: a sibling-controlled study.* Dev Med Child Neurol, 2004. **46**(5): p. 340-6.

10. Falkert, A., K. Dittmann, and B. Seelbach-Gobel, *Silver-Russell syndrome as a cause for early intrauterine growth restriction.* Prenat Diagn, 2005. **25**(6): p. 497-501.

11. Varendi, K., et al., *miR-1, miR-10b, miR-155, and miR-191 are novel regulators of BDNF.* Cellular and Molecular Life Sciences, 2014. **71**(22): p. 4443-4456.

12. Kadoch, C., R.A. Copeland, and H. Keilhack, *PRC2 and SWI/SNF Chromatin Remodeling Complexes in Health and Disease.* Biochemistry, 2016. **55**(11): p. 1600-1614.

13. Kosho, T., et al., *Clinical correlations of mutations affecting six components of the SWI/SNF complex: Detailed description of 21 patients and a review of the literature.* American Journal of Medical Genetics Part A, 2013. **161**(6): p. 1221-1237.

14. Coffin, G.S. and E. Siris, *Mental retardation with absent fifth fingernail and terminal phalanx.* Am J Dis Child, 1970. **119**(5): p. 433-9.

15. Van Houdt, J.K., et al., *Heterozygous missense mutations in SMARCA2 cause Nicolaides-Baraitser syndrome.* Nature genetics, 2012. **44**(4): p. 445-449.

16. Bultman, S., et al., *A Brg1 null mutation in the mouse reveals functional differences among mammalian SWI/SNF complexes.* Molecular cell, 2000. **6**(6): p. 1287-1295.

17. Long, S.W., et al., *A brain-derived MeCP2 complex supports a role for MeCP2 in RNA processing.* Bioscience reports, 2011. **31**(5): p. 333-343.

18. Li, R., et al., *Misregulation of Alternative Splicing in a Mouse Model of Rett Syndrome.* PLOS Genetics, 2016. **12**(6): p. e1006129.

19. Li, W., X. Xu, and L. Pozzo-Miller, *Excitatory synapses are stronger in the hippocampus of Rett syndrome mice due to altered synaptic trafficking of AMPA-type glutamate receptors.* Proceedings of the National Academy of Sciences, 2016. **113**(11): p. E1575-E1584.

20. Li, W., et al., *A small-molecule TrkB ligand restores hippocampal synaptic plasticity and object location memory in Rett syndrome mice.* Disease Models & Mechanisms, 2017. **10**(7): p. 837-845.

21. Ramocki, M.B., et al., *Autism and other neuropsychiatric symptoms are prevalent in individuals with MeCP2 duplication syndrome.* Annals of Neurology, 2009. **66**(6): p. 771-782.

22. Shimada, S., et al., *MECP2 duplication syndrome in both genders.* Brain and Development, 2013. **35**(5): p. 411-419.

23. Jiang, M., et al., *Dendritic Arborization and Spine Dynamics Are Abnormal in the Mouse Model of MECP2 Duplication Syndrome.* The Journal of Neuroscience, 2013. **33**(50): p. 19518.

24. Na, E.S., et al., *A Mouse Model for MeCP2 Duplication Syndrome: MeCP2 Overexpression Impairs Learning and Memory and Synaptic Transmission.* The Journal of Neuroscience, 2012. **32**(9): p. 3109.

25. Sztainberg, Y., et al., *Reversal of phenotypes in MECP2 duplication mice using genetic rescue or antisense oligos.* Nature, 2015. **528**(7580): p. 123-126.

26. Willemsen, R., J. Levenga, and B.A. Oostra, *CGG repeat in the FMR1 gene: size matters.* Clinical genetics, 2011. **80**(3): p. 214-225.

27. Riddle, J.E., et al., *Phenotypic Involvement in Females With the FMR1 Gene Mutation.* American Journal on Mental Retardation, 1997. **102**(6): p. 590-601.

28. Hagerman, R.J., et al., *Intention tremor, parkinsonism, and generalized brain atrophy in male carriers of fragile X.* Neurology, 2001. **57**(1): p. 127-30.

29. Hunter, J.E., J.K. Rohr, and S.L. Sherman, *Co-occurring diagnoses among FMR1 premutation allele carriers.* Clin Genet, 2010. **77**(4): p. 374-81.

30. Primerano, B., et al., *Reduced FMR1 mRNA translation efficiency in Fragile X patients with premutations.* RNA, 2003. **8**(12): p. 1482-1488.

31. Kenneson, A., et al., *Reduced FMRP and increased FMR1 transcription is proportionally associated with CGG repeat number in intermediate-length and premutation carriers.* Human Molecular Genetics, 2001. **10**(14): p. 1449-1454.

32. Darnell, Jennifer C., et al., *FMRP Stalls Ribosomal Translocation on mRNAs Linked to Synaptic Function and Autism.* Cell, 2011. **146**(2): p. 247-261.

33. Edbauer, D., et al., *Regulation of Synaptic Structure and Function by FMRP-Associated MicroRNAs miR-125b and miR-132.* Neuron, 2010. **65**(3): p. 373-384.

34. Antar, L.N., et al., *Local functions for FMRP in axon growth cone motility and activity-dependent regulation of filopodia and spine synapses.* Molecular and Cellular Neuroscience, 2006. **32**(1): p. 37-48.

35. Butler, N.R., H. Goldstein, and E.M. Ross, *Cigarette Smoking in Pregnancy: Its Influence on Birth Weight and Perinatal Mortality.* British Medical Journal, 1972. **2**(5806): p. 127-130.

36. Delpisheh, A., et al., *Prenatal smoking exposure and asymmetric fetal growth restriction.* Annals of Human Biology, 2008. **35**(6): p. 573-583.

37. Chen, M., et al., *Nicotine-induced prenatal overexposure to maternal glucocorticoid and intrauterine growth retardation in rat.* Exp Toxicol Pathol, 2007. **59**(3-4): p. 245-51.

38. He, X., et al., *Prenatal nicotine exposure induces HPA axis-hypersensitivity in offspring rats via the intrauterine programming of up-regulation of hippocampal GAD67.* Arch Toxicol, 2017. **91**(12): p. 3927-3943.

39. McCormick, C.M., et al., *Sex-specific effects of prenatal stress on hypothalamic-pituitary-adrenal responses to stress and brain glucocorticoid receptor density in adult rats.* Developmental Brain Research, 1995. **84**(1): p. 55-61.

40. Davis, E.P., et al., *Fetal Glucocorticoid Exposure Is Associated with Preadolescent Brain Development.* Biological Psychiatry, 2013. **74**(9): p. 647-655.

41. Lee, B.E., et al., *Secondhand smoke exposure during pregnancy and infantile neurodevelopment.* Environ Res, 2011. **111**(4): p. 539-44.

42. Maccani, M.A. and V.S. Knopik, *Cigarette Smoke Exposure-Associated Alterations to Non-Coding RNA.* Frontiers in Genetics, 2012. **3**: p. 53.

43. Balaraman, S., U.H. Winzer-Serhan, and R.C. Miranda, *Opposing Actions of Ethanol and Nicotine on MicroRNAs are Mediated by Nicotinic Acetylcholine Receptors in Fetal Cerebral Cortical–Derived Neural Progenitor Cells.* Alcoholism: Clinical and Experimental Research, 2012. **36**: p. 1669-1677.

44. Coolen, M., S. Katz, and L. Bally-Cuif, *miR-9: a versatile regulator of neurogenesis.* Frontiers in Cellular Neuroscience, 2013. **7**: p. 220.

45. Põlajeva, J., et al., *miRNA-21 is developmentally regulated in mouse brain and is co-expressed with SOX2 in glioma.* BMC Cancer, 2012. **12**(1): p. 378.

46. Wei, C., et al., *miR-153 regulates SNAP-25, synaptic transmission, and neuronal development.* PloS one, 2013. **8**(2): p. e57080.

47. Cheng, Q., et al., *PAX6, a novel target of miR-335, inhibits cell proliferation and invasion in glioma cells.* Molecular medicine reports, 2014. **10**(1): p. 399-404.

48. Murphy, S.K., et al., *Gender-specific methylation differences in relation to prenatal exposure to cigarette smoke.* Gene, 2012. **494**(1): p. 36-43.

49. Slotkin, T.A., et al., *Effects of tobacco smoke on PC12 cell neurodifferentiation are distinct from those of nicotine or benzo[a]pyrene.* Neurotoxicol Teratol, 2014. **43**: p. 19-24.

50. Kaiserman, M.J. and W.S. Rickert, *Carcinogens in tobacco smoke: benzo[a]pyrene from Canadian cigarettes and cigarette tobacco.* American Journal of Public Health, 1992. **82**(7): p. 1023-1026.

51. McCallister, M.M., et al., *Prenatal exposure to benzo(a)pyrene impairs later-life cortical neuronal function.* NeuroToxicology, 2008. **29**(5): p. 846-854.

52. Peterson, B.S., et al., *Effects of prenatal exposure to air pollutants (polycyclic aromatic hydrocarbons) on the development of brain white matter, cognition, and behavior in later childhood.* JAMA Psychiatry, 2015. **72**(6): p. 531-540.

53. Slotkin, T.A. and F.J. Seidler, *Benzo[a]pyrene impairs neurodifferentiation in PC12 cells.* Brain Research Bulletin, 2009. **80**(1): p. 17-21.

54. Andersen, H.R., et al., *Occupational pesticide exposure in early pregnancy associated with sex-specific neurobehavioral deficits in the children at school age.* Neurotoxicology and Teratology, 2015. **47**: p. 1-9.

55. Rauh, V., et al., *Seven-Year Neurodevelopmental Scores and Prenatal Exposure to Chlorpyrifos, a Common Agricultural Pesticide.* Environmental Health Perspectives, 2011. **119**(8): p. 1196-1201.

56. Rauh, V.A., et al., *Impact of Prenatal Chlorpyrifos Exposure on Neurodevelopment in the First 3 Years of Life Among Inner-City Children.* Pediatrics, 2006. **118**(6): p. e1845-e1859.

57. Rauh, V.A., et al., *Brain anomalies in children exposed prenatally to a common organophosphate pesticide.* Proceedings of the National Academy of Sciences, 2012. **109**(20): p. 7871-7876.

58. Laugeray, A., et al., *In utero and lactational exposure to low-doses of the pyrethroid insecticide cypermethrin leads to neurodevelopmental defects in male mice—An ethological and transcriptomic study.* PLOS ONE, 2017. **12**(10): p. e0184475.

59. LaGasse, L.L., et al., *Prenatal Methamphetamine Exposure and Childhood Behavior Problems at 3 and 5 Years of Age.* Pediatrics, 2012.

60. Diaz, S.D., et al., *Effects of Prenatal Methamphetamine Exposure on Behavioral and Cognitive Findings at 7.5 Years of Age.* The Journal of Pediatrics, 2014. **164**(6): p. 1333-1338.

61. Cernerud, L., et al., *Amphetamine addiction during pregnancy: 14-year follow-up of growth and school performance.* Acta Pædiatrica, 1996. **85**(2): p. 204-208.

62. Chang, L., et al., *Smaller subcortical volumes and cognitive deficits in children with prenatal methamphetamine exposure.* Psychiatry Research: Neuroimaging, 2004. **132**(2): p. 95-106.

63. Marroun, H.E., et al., *Intrauterine cannabis exposure leads to more aggressive behavior and attention problems in 18-month-old girls.* Drug and Alcohol Dependence, 2011. **118**(2): p. 470-474.

64. El Marroun, H., et al., *Prenatal cannabis and tobacco exposure in relation to brain morphology: a prospective neuroimaging study in young children.* Biological psychiatry, 2016. **79**(12): p. 971-979.

65. Goldschmidt, L., N.L. Day, and G.A. Richardson, *Effects of prenatal marijuana exposure on child behavior problems at age 10.* Neurotoxicology and Teratology, 2000. **22**(3): p. 325-336.

66. Wang, X., et al., *In utero marijuana exposure associated with abnormal amygdala dopamine D2 gene expression in the human fetus.* Biological Psychiatry, 2004. **56**(12): p. 909-915.

67. Smith, A.M., et al., *Effects of prenatal marijuana on response inhibition: an fMRI study of young adults.* Neurotoxicology and Teratology, 2004. **26**(4): p. 533-542.

68. Lange, S., et al., *Global Prevalence of Fetal Alcohol Spectrum Disorder Among Children and Youth: A Systematic Review and Meta-analysis.* JAMA Pediatr, 2017. **171**(10): p. 948-956.

69. Grant, B.F., et al., *Epidemiology of DSM-5 Alcohol Use Disorder: Results From the National Epidemiologic Survey on Alcohol and Related Conditions III.* JAMA Psychiatry, 2015. **72**(8): p. 757-66.

70. Finer, L.B. and M.R. Zolna, *Declines in Unintended Pregnancy in the United States, 2008-2011.* N Engl J Med, 2016. **374**(9): p. 843-52.

71. Bakhireva, L.N., et al., *Prevalence of Prenatal Alcohol Exposure in the State of Texas as Assessed by Phosphatidylethanol in Newborn Dried Blood Spot Specimens.* Alcohol Clin Exp Res, 2017. **41**(5): p. 1004-1011.

72. Archibald, S.L., et al., *Brain dysmorphology in individuals with severe prenatal alcohol exposure.* Dev Med Child Neurol, 2001. **43**(3): p. 148-54.

73. Riley, E.P., C.L. McGee, and E.R. Sowell, *Teratogenic effects of alcohol: a decade of brain imaging.* Am J Med Genet C Semin Med Genet, 2004. **127C**(1): p. 35-41.

74. Coles, C.D., et al., *A comparison of children affected by prenatal alcohol exposure and attention deficit, hyperactivity disorder.* Alcohol Clin Exp Res, 1997. **21** (1): p. 150-161.

75. May, P.A., et al., *Prevalence of Fetal Alcohol Spectrum Disorders in 4 US Communities.* JAMA, 2018. **319**(5): p. 474-482.

76. Riley, E.P., M.A. Infante, and K.R. Warren, *Fetal alcohol spectrum disorders: an overview.* Neuropsychol Rev, 2011. **21**(2): p. 73-80.

77. Jones, K.L., *The effects of alcohol on fetal development.* Birth Defects Research Part C: Embryo Today: Reviews, 2011. **93**: p. 3-11.

78. Jones, K.L., et al., *Fetal alcohol spectrum disorders: Extending the range of structural defects.* American Journal of Medical Genetics Part A, 2010. **152A**: p. 2731-2735.

79. Hoyme, H.E., et al., *A practical clinical approach to diagnosis of fetal alcohol spectrum disorders: clarification of the 1996 institute of medicine criteria.* Pediatrics, 2005. **115**: p. 39-47.

80. Riley, E.P. and C.L. McGee, *Fetal alcohol spectrum disorders: an overview with emphasis on changes in brain and behavior.* Exp Biol Med (Maywood), 2005. **230**(6): p. 357-65.

81. Roebuck, T.M., et al., *Prenatal exposure to alcohol affects the ability to maintain postural balance.* Alcohol Clin Exp Res, 1998. **22** (1): p. 252-258.

82. Zhou, D., et al., *Developmental cortical thinning in fetal alcohol spectrum disorders.* NeuroImage, 2011. **58**: p. 16-25.

83. Miller, M.W., *Effects of Alcohol on the Generation and Migration of Cerebral Cortical Neurons.* Science, 1986. **233**(4770): p. 1308-1311.

84. Miranda, R.C., et al., *Modeling the impact of alcohol on cortical development in a dish: strategies from mapping neural stem cell fate.* Methods Mol Biol, 2008. **447**: p. 151-68.

85. Nash, R., et al., *Human embryonic stem cell model of ethanol-mediated early developmental toxicity.* Experimental Neurology, 2012. **234**: p. 127-135.

86. Santillano, D.R., et al., *Ethanol induces cell-cycle activity and reduces stem cell diversity to alter both regenerative capacity and differentiation potential of cerebral cortical neuroepithelial precursors.* BMC Neurosci, 2005. **6**: p. 59.

87. Camarillo, C. and R.C. Miranda, *Ethanol exposure during neurogenesis induces persistent effects on neural maturation: evidence from an ex vivo model of fetal cerebral cortical neuroepithelial progenitor maturation.* Gene Expr, 2008. **14**(3): p. 159-71.

88. Patten, A.R., C.J. Fontaine, and B.R. Christie, *A Comparison of the Different Animal Models of Fetal Alcohol Spectrum Disorders and Their Use in Studying Complex Behaviors.* Frontiers in Pediatrics, 2014. **2**(93).

89. Cudd, T.A., *Animal Model Systems for the Study of Alcohol Teratology.* Experimental Biology and Medicine, 2005. **230**(6): p. 389-393.

90. Valenzuela, C.F., et al., *Does moderate drinking harm the fetal brain? Insights from animal models.* Trends in Neurosciences, 2012. **35**(5): p. 284-292.

91. Seguin, D. and R. Gerlai, *Fetal alcohol spectrum disorders: Zebrafish in the analysis of the milder and more prevalent form of the disease.* Behavioural brain research, 2018. **352**: p. 125-132.

92. Bilotta, J., et al., *Ethanol exposure alters zebrafish development: A novel model of fetal alcohol syndrome.* Neurotoxicology and Teratology, 2004. **26**(6): p. 737-743.

93. Cleal, M. and M.O. Parker, *Moderate developmental alcohol exposure reduces repetitive alternation in a zebrafish model of fetal alcohol spectrum disorders.* Neurotoxicol Teratol, 2018. **70**: p. 1-9.

94. Muralidharan, P., S. Sarmah, and J.A. Marrs, *Retinal Wnt signaling defect in a zebrafish fetal alcohol spectrum disorder model.* PLOS ONE, 2018. **13**(8): p. e0201659.

95. Watari, H., D.E. Born, and C.A. Gleason, *Effects of First Trimester Binge Alcohol Exposure on Developing White Matter in Fetal Sheep.* Pediatric Research, 2006. **59**(4): p. 560-564.

96. Ramadoss, J., et al., *Temporal Vulnerability of Fetal Cerebellar Purkinje Cells to Chronic Binge Alcohol Exposure: Ovine Model.* Alcoholism: Clinical and Experimental Research, 2007. **31**(10): p. 1738-1745.

97. Ramadoss, J., et al., *All Three Trimester Binge Alcohol Exposure Causes Fetal Cerebellar Purkinje Cell Loss in the Presence of Maternal Hypercapnea, Acidemia, and Normoxemia: Ovine Model.* Alcoholism: Clinical and Experimental Research, 2007. **31**(7): p. 1252-1258.

98. Sawant, O.B., et al., *Maternal choline supplementation mitigates alcohol-induced fetal cranio-facial abnormalities detected using an ultrasonographic examination in a sheep model.* Alcohol, 2019. **81**: p. 31-38.

99. Birch, S.M., et al., *Maternal choline supplementation in a sheep model of first trimester binge alcohol fails to protect against brain volume reductions in peripubertal lambs.* Alcohol (Fayetteville, N.Y.), 2016. **55**: p. 1-8.

100. Balaraman, S., et al., *Maternal and Neonatal Plasma MicroRNA Biomarkers for Fetal Alcohol Exposure in an Ovine Model.* Alcoholism: Clinical and Experimental Research, 2014. **38**(5): p. 1390-1400.

101. Sulik, K., M. Johnston, and M. Webb, *Fetal alcohol syndrome: embryogenesis in a mouse model.* Science, 1981. **214**(4523): p. 936-938.

102. Marquardt, K. and J.L. Brigman, *The impact of prenatal alcohol exposure on social, cognitive and affective behavioral domains: Insights from rodent models.* Alcohol, 2016. **51**: p. 1-15.

103. Mahnke, A.H., et al., *Nonprotein-coding RNAs in Fetal Alcohol Spectrum Disorders*, in *Progress in Molecular Biology and Translational Science*, D.R. Grayson, Editor. 2018, Academic Press. p. 299-342.

104. Gudenas, B.L., A.K. Srivastava, and L. Wang, *Integrative genomic analyses for identification and prioritization of long non-coding RNAs associated with autism.* PloS one, 2017. **12**(5): p. e0178532.

105. Robledo, R.F., et al., *The Dlx5 and Dlx6 homeobox genes are essential for craniofacial, axial, and appendicular skeletal development.* Genes & development, 2002. **16**(9): p. 1089-1101.

106. Burrowes, S.G., et al., *The BAF (BRG1/BRM-Associated Factor) chromatin-remodeling complex exhibits ethanol sensitivity in fetal neural progenitor cells and regulates transcription at the miR-9-2 encoding gene locus.* Alcohol, 2017. **60**: p. 149-158.

107. Rogic, S., A. Wong, and P. Pavlidis, *Meta-Analysis of Gene Expression Patterns in Animal Models of Prenatal Alcohol Exposure Suggests Role for Protein Synthesis Inhibition and Chromatin Remodeling.* Alcoholism: Clinical and Experimental Research, 2016. **40**(4): p. 717-727.

108. Veazey, K.J., et al., *Alcohol-Induced Epigenetic Alterations to Developmentally Crucial Genes Regulating Neural Stemness and Differentiation.* Alcoholism: Clinical and Experimental Research, 2013. **37**(7): p. 1111-1122.

109. Caldwell, K.K., et al., *Arsenic exposure during embryonic development alters the expression of the long noncoding RNA growth arrest specific-5 (Gas5) in a sex-dependent manner.* Neurotoxicology and Teratology, 2018. **66**: p. 102-112.

110. Ang, C.E., et al., *The novel lncRNA lnc-NR2F1 is pro-neurogenic and mutated in human neurodevelopmental disorders.* Elife, 2019. **8**: p. e41770.

111. Marahrens, Y., J. Loring, and R. Jaenisch, *Role of the Xist Gene in X Chromosome Choosing.* Cell, 1998. **92**(5): p. 657-664.

112. Lee, J.T., L.S. Davidow, and D. Warshawsky, *Tsix, a gene antisense to Xist at the X-inactivation centre.* Nat Genet, 1999. **21**(4): p. 400-4.

113. Talebizadeh, Z., et al., *Brief report: non-random X chromosome inactivation in females with autism.* Journal of autism and developmental disorders, 2005. **35**(5): p. 675-681.

114. Plenge, R.M., et al., *Skewed X-chromosome inactivation is a common feature of X-linked mental retardation disorders.* American journal of human genetics, 2002. **71**(1): p. 168-173.

115. Knudsen, G.P., et al., *Increased skewing of X chromosome inactivation in Rett syndrome patients and their mothers.* Eur J Hum Genet, 2006. **14**(11): p. 1189-94.

116. Kumamoto, T. and S. Oshio, *Effect of fetal exposure to bisphenol A on brain mediated by X-chromosome inactivation.* J Toxicol Sci, 2013. **38**(3): p. 485-94.

117. Ji, B., et al., *Over-expression of XIST, the Master Gene for X Chromosome Inactivation, in Females With Major Affective Disorders.* EBioMedicine, 2015. **2**(8): p. 909-918.

118. Jaitin, D.A., et al., *Massively Parallel Single-Cell RNA-Seq for Marker-Free Decomposition of Tissues into Cell Types.* Science, 2014. **343**(6172): p. 776-779.

119. Yuan, J. and P.A. Sims, *An Automated Microwell Platform for Large-Scale Single Cell RNA-Seq.* Scientific Reports, 2016. **6**(1): p. 33883.

120. Fan, H.C., G.K. Fu, and S.P.A. Fodor, *Combinatorial labeling of single cells for gene expression cytometry.* Science, 2015. **347**(6222): p. 1258367.

121. Shalek, A.K., et al., *Single-cell RNA-seq reveals dynamic paracrine control of cellular variation.* Nature, 2014. **510**(7505): p. 363-369.

122. Macosko, Evan Z., et al., *Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets.* Cell, 2015. **161**(5): p. 1202-1214.

123. See, P., et al., *A Single-Cell Sequencing Guide for Immunologists.* Frontiers in Immunology, 2018. **9**(2425).

124. Ståhl, P.L., et al., *Visualization and analysis of gene expression in tissue sections by spatial transcriptomics.* Science, 2016. **353**(6294): p. 78-82.

125. Baccin, C., et al., *Combined single-cell and spatial transcriptomics reveal the molecular, cellular and spatial bone marrow niche organization.* Nature cell biology, 2019: p. 1-11.

126. Navarro, J.F., et al., *Spatial Transcriptomics Reveals Genes Associated with Dysregulated Mitochondrial Functions and Stress Signaling in Alzheimer Disease.* iScience, 2020. **23**(10): p. 101556.

127. Chen, W.-T., et al., *Spatial and temporal transcriptomics reveal microglia-astroglia crosstalk in the amyloid-β plaque cell niche of Alzheimer's disease.* bioRxiv, 2019: p. 719930.

128. Gregory, J.M., et al., *Spatial transcriptomics identifies spatially dysregulated expression of GRM3 and USP47 in amyotrophic lateral sclerosis.* Neuropathology and Applied Neurobiology, 2020. **46**(5): p. 441-457.

129. Chen, W.-T., et al., *Spatial Transcriptomics and In Situ Sequencing to Study Alzheimer's Disease.* Cell, 2020. **182**(4): p. 976-991.e19.

130. Zhong, S., et al., *A single-cell RNA-seq survey of the developmental landscape of the human prefrontal cortex.* Nature, 2018. **555**(7697): p. 524-528.

131. Weng, Q., et al., *Single-Cell Transcriptomics Uncovers Glial Progenitor Diversity and Cell Fate Determinants during Development and Gliomagenesis.* Cell stem cell, 2019. **24**(5): p. 707-723.e8.

132. Mayer, C., et al., *Developmental diversification of cortical inhibitory interneurons.* Nature, 2018. **555**(7697): p. 457-462.

133. Xiang, Y., et al., *Fusion of Regionally Specified hPSC-Derived Organoids Models Human Brain Development and Interneuron Migration.* Cell Stem Cell, 2017. **21**(3): p. 383-398.e7.

134. Anderson, A.G., et al., *Single-Cell Analysis of Foxp1-Driven Mechanisms Essential for Striatal Development.* Cell Reports, 2020. **30**(9): p. 3051-3066.e7.

135. Co, M., et al., *Cortical Foxp2 Supports Behavioral Flexibility and Developmental Dopamine D1 Receptor Expression.* Cerebral Cortex, 2019. **30**(3): p. 1855-1870.

136. Fazel Darbandi, S., et al., *Enhancing WNT Signaling Restores Cortical Neuronal Spine Maturation and Synaptogenesis in Tbr1 Mutants.* Cell Reports, 2020. **31**(2): p. 107495.

137. Guo, H., et al., *Single-Cell RNA Sequencing of Human Embryonic Stem Cell Differentiation Delineates Adverse Effects of Nicotine on Embryonic Development.* Stem Cell Reports, 2019. **12**(4): p. 772-786.

138. Lee, C.-M., et al., *Single-cell RNA-seq analysis revealed long-lasting adverse effects of tamoxifen on neurogenesis in prenatal and adult brains.* Proceedings of the National Academy of Sciences, 2020. **117**(32): p. 19578-19589.

139. Hsu, K.S., et al., *Single-cell RNA-seq Analysis Reveals That Prenatal Arsenic Exposure Results in Long-term, Adverse Effects on Immune Gene Expression in Response to Influenza A Infection.* Toxicological Sciences, 2020. **176**(2): p. 312-328.

140. Mohammad, S., et al., *Kcnn2 blockade reverses learning deficits in a mouse model of fetal alcohol spectrum disorders.* Nature Neuroscience, 2020. **23**(4): p. 533-543.

141. Xue, Z., et al., *Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing.* Nature, 2013. **500**(7464): p. 593-597.

142. Oldham, M.C., et al., *Functional organization of the transcriptome in human brain.* Nature neuroscience, 2008. **11**(11): p. 1271.

143. Johnson, M.R., et al., *Systems genetics identifies a convergent gene network for cognition and neurodevelopmental disease.* Nature Neuroscience, 2016. **19**(2): p. 223-232.

144. Voineagu, I., et al., *Transcriptomic analysis of autistic brain reveals convergent molecular pathology.* Nature, 2011. **474**(7351): p. 380-384.

145. Vogel Ciernia, A., et al., *Epigenomic Convergence of Neural-Immune Risk Factors in Neurodevelopmental Disorder Cortex.* Cerebral Cortex, 2019. **30**(2): p. 640-655.

146. Zhong, S., et al., *A single-cell RNA-seq survey of the developmental landscape of the human prefrontal cortex.* Nature, 2018. **555**(7697): p. 524-528.

147. Hanamsagar, R., et al., *Generation of a microglial developmental index in mice and in humans reveals a sex difference in maturation and immune reactivity.* Glia, 2017. **65**(9): p. 1504-1520.

148. Khalid, O., et al., *Gene expression signatures affected by alcohol-induced DNA methylomic deregulation in human embryonic stem cells.* Stem Cell Research, 2014. **12**(3): p. 791-806.

149. Kim, Y.Y., et al., *Alcohol-Induced Molecular Dysregulation in Human Embryonic Stem Cell-Derived Neural Precursor Cells.* PLOS ONE, 2016. **11**(9): p. e0163812.

150. Alberry, B.L., C.A. Castellani, and S.M. Singh, *Hippocampal transcriptome analysis following maternal separation implicates altered RNA processing in a mouse model of fetal alcohol spectrum disorder.* Journal of Neurodevelopmental Disorders, 2020. **12**: p. 1-16.

151. Miranda, R., *MicroRNAs and Fetal Brain Development: Implications for Ethanol Teratology during the Second Trimester Period of Neurogenesis.* Frontiers in Genetics, 2012. **3**(77).

152. Roozen, S., et al., *Worldwide Prevalence of Fetal Alcohol Spectrum Disorders: A Systematic Literature Review Including Meta-Analysis.* Alcohol Clin Exp Res, 2016. **40**(1): p. 18-32.

153. SAMHSA *The NSDUH Report: Substance Use among Women During Pregnancy and Following Childbirth*. 2009. **NSDUH09-0521**.

154. SAMHSA *The NSDUH Report: 18 percent of pregnant women drink alcohol during early pregnancy*. NSDUH Report, 2013.

155. Ethen, M.K., et al., *Alcohol consumption by women before and during pregnancy.* Matern Child Health J, 2009. **13**(2): p. 274-85.

156. Bayer, S., et al., *Timetables of neurogenesis in the human brain based on experimentally determined patterns in the rat.* Neuro Toxicology, 1993. **14**(1): p. 83-144.

157. Bystron, I., C. Blakemore, and P. Rakic, *Development of the human cerebral cortex: Boulder Committee revisited.* Nat Rev Neurosci, 2008. **9**(2): p. 110-22.

158. Maier, S.E., J.A. Miller, and J.R. West, *Prenatal binge-like alcohol exposure in the rat results in region-specific deficits in brain growth.* Neurotoxicol Teratol, 1999. **21**(3): p. 285-91.

159. Maier, S.E., et al., *Alcohol exposure during the first two trimesters equivalent alters granule cell number and neurotrophin expression in the developing rat olfactory bulb.* J Neurobiol, 1999. **41**(3): p. 414-23.

160. Miller, M., *Effects of prenatal exposure to ethanol on neocortical development: II. Cell proliferation in the ventricular and subventricular zones of the rat.* The Journal of Comparative Neurology, 1989. **287** p. 326-338.

161. Miller, M.W. and R.S. Nowakowski, *Effect of prenatal exposure to ethanol on the cell cycle kinetics and growth fraction in the proliferative zones of fetal rat cerebral cortex.* Alcohol Clin Exp Res, 1991. **15**(2): p. 229-32.

162. Prock, T.L. and R.C. Miranda, *Embryonic cerebral cortical progenitors are resistant to apoptosis, but increase expression of suicide receptor DISC-complex genes and suppress autophagy following ethanol exposure.* Alcoholism: Clinical and Experimental Research, 2007. **31**(4): p. 694-703.

163. Tingling, J.D., et al., *CD24 expression identifies teratogen-sensitive fetal neural stem cell subpopulations: evidence from developmental ethanol exposure and orthotopic cell transfer models.* PLoS One, 2013. **8**(7): p. e69560.

164. Tsai, P.C., et al., *MiR-153 targets the nuclear factor-1 family and protects against teratogenic effects of ethanol exposure in fetal neural stem cells.* Biol Open, 2014. **3**(8): p. 741-58.

165. Veazey, K.J., et al., *Dose-dependent alcohol-induced alterations in chromatin structure persist beyond the window of exposure and correlate with fetal alcohol syndrome birth defects.* Epigenetics Chromatin, 2015. **8**: p. 39.

166. Pappalardo-Carter, D.L., et al., *Suppression and epigenetic regulation of MiR-9 contributes to ethanol teratology: evidence from zebrafish and murine fetal neural stem cell models.* Alcohol Clin Exp Res, 2013. **37**(10): p. 1657-67.

167. Sathyan, P., H.B. Golden, and R.C. Miranda, *Competing interactions between micro-RNAs determine neural progenitor survival and proliferation after ethanol exposure: evidence from an ex vivo model of the fetal cerebral cortical neuroepithelium.* J Neurosci, 2007. **27**(32): p. 8546-57.

168. Tseng, A.M., et al., *Ethanol Exposure Increases miR-140 in Extracellular Vesicles: Implications for Fetal Neural Stem Cell Proliferation and Maturation.* Alcohol Clin Exp Res, 2019. **43**(7): p. 1414-1426.

169. Fernandez-Tresguerres, B., et al., *Evolution of the mammalian embryonic pluripotency gene regulatory network.* Proc Natl Acad Sci U S A, 2010. **107**(46): p. 19955-60.

170. Nichols, J., et al., *Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4.* Cell, 1998. **95**(3): p. 379-91.

171. Takahashi, K., et al., *Induction of pluripotent stem cells from adult human fibroblasts by defined factors.* Cell, 2007. **131**(5): p. 861-72.

172. Archer, T.C., J. Jin, and E.S. Casey, *Interaction of Sox1, Sox2, Sox3 and Oct4 during primary neurogenesis.* Dev Biol, 2011. **350**(2): p. 429-40.

173. Okuda, T., et al., *Oct-3/4 repression accelerates differentiation of neural progenitor cells in vitro and in vivo.* Brain Res Mol Brain Res, 2004. **132**(1): p. 18-30.

174. Sabnis, N.G., et al., *The Efflux Transporter ABCG2 Maintains Prostate Stem Cells.* Mol Cancer Res, 2017. **15**(2): p. 128-140.

175. Wang, X.Q., et al., *Octamer 4 (Oct4) mediates chemotherapeutic drug resistance in liver cancer cells through a potential Oct4-AKT-ATP-binding cassette G2 pathway.* Hepatology, 2010. **52**(2): p. 528-39.

176. Poursani, E.M., B. Mohammad Soltani, and S.J. Mowla, *Differential Expression of OCT4 Pseudogenes in Pluripotent and Tumor Cell Lines.* Cell J, 2016. **18**(1): p. 28-36.

177. Lin, H., et al., *Stem cell regulatory function mediated by expression of a novel mouse Oct4 pseudogene.* Biochemical and Biophysical Research Communications, 2007. **355**(1): p. 111-116.

178. Liedtke, S., et al., *Oct4 and its pseudogenes confuse stem cell research.* Cell Stem Cell, 2007. **1**(4): p. 364-6.

179. An, Y., K.L. Furber, and S. Ji, *Pseudogenes regulate parental gene expression via ceRNA network.* Journal of Cellular and Molecular Medicine, 2017. **21**(1): p. 185-192.

180. Glenfield, C. and A. McLysaght, *Pseudogenes Provide Evolutionary Evidence for the Competitive Endogenous RNA Hypothesis.* Mol Biol Evol, 2018. **35**(12): p. 2886-2899.

181. Tam, O.H., et al., *Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes.* Nature, 2008. **453**: p. 534.

182. Hawkins, P.G. and K.V. Morris, *Transcriptional regulation of Oct4 by a long non-coding RNA antisense to Oct4-pseudogene 5.* Transcription, 2010. **1**(3): p. 165-175.

183. Bai, M., et al., *OCT4 pseudogene 5 upregulates OCT4 expression to promote proliferation by competing with miR-145 in endometrial carcinoma.* Oncol Rep, 2015. **33**(4): p. 1745-52.

184. Wang, L., et al., *Pseudogene OCT4-pg4 functions as a natural micro RNA sponge to regulate OCT4 expression by competing for miR-145 in hepatocellular carcinoma.* Carcinogenesis, 2013. **34**(8): p. 1773-81.

185. Scarola, M., et al., *Epigenetic silencing of Oct4 by a complex containing SUV39H1 and Oct4 pseudogene lncRNA.* Nat Commun, 2015. **6**: p. 7631.

186. Adachi, J., et al., *Degrees of alcohol intoxication in 117 hospitalized cases.* J Stud Alcohol, 1991. **52**(5): p. 448-453.

187. Wheeler, D.L., et al., *Database resources of the National Center for Biotechnology Information.* Nucleic Acids Research, 2006. **35**(suppl_1): p. D5-D12.

188. Langmead, B., et al., *Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.* Genome Biology, 2009. **10**(3): p. R25.

189. Trapnell, C., et al., *Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks.* Nature Protocols, 2012. **7**(3): p. 562-578.

190. Li, Y., et al., *Transcriptome analysis reveals determinant stages controlling human embryonic stem cell commitment to neuronal cells.* J Biol Chem, 2017. **292**(48): p. 19590-19604.

191. Lachmann, A., et al., *Massive mining of publicly available RNA-seq data from human and mouse.* Nat Commun, 2018. **9**(1): p. 1366.

192. Vance, K.W., *Mapping Long Noncoding RNA Chromatin Occupancy Using Capture Hybridization Analysis of RNA Targets (CHART).* Methods Mol Biol, 2017. **1468**: p. 39-50.

193. Hedges, L.V., *Distribution Theory for Glass's Estimator of Effect size and Related Estimators.* Journal of Educational Statistics, 1981. **6**(2): p. 107-128.

194. Lin, S., et al., *Comparison of the transcriptional landscapes between human and mouse tissues.* Proc Natl Acad Sci U S A, 2014. **111**(48): p. 17224-9.

195. Pervouchine, D.D., et al., *Enhanced transcriptome maps from multiple mouse tissues reveal evolutionary constraint in gene expression.* Nat Commun, 2015. **6**: p. 5903.

196. Schaum, N., et al., *Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris.* Nature, 2018. **562**(7727): p. 367-372.

197. Abrajano, J.J., et al., *REST and CoREST modulate neuronal subtype specification, maturation and maintenance.* PLoS One, 2009. **4**(12): p. e7936.

198. Conaco, C., et al., *Reciprocal actions of REST and a microRNA promote neuronal identity.* PNAS, 2006. **103**(7): p. 2422-2427.

199. Singh, S.K., et al., *REST maintains self-renewal and pluripotency of embryonic stem cells.* Nature, 2008. **453**(7192): p. 223-7.

200. Santiago, M., et al., *TET enzymes and DNA hydroxymethylation in neural development and function - how critical are they?* Genomics, 2014. **104**(5): p. 334-40.

201. Sun, A.Y., et al., *Ethanol-induced cell death by lipid peroxidation in PC12 cells.* Neurochem Res, 1997. **22**(10): p. 1187-92.

202. Zenki, K.C., et al., *Effects of ethanol and acetaldehyde in zebrafish brain structures: An in vitro approach on glutamate uptake and on toxicity-related parameters.* Toxicology in Vitro, 2014. **28**(5): p. 822-828.

203. Prock, T.L. and R.C. Miranda, *Embryonic cerebral cortical progenitors are resistant to apoptosis, but increase expression of suicide receptor DISC-complex genes and suppress autophagy following ethanol exposure.* Alcohol Clin Exp Res, 2007. **31**(4): p. 694-703.

204. Cheetham, S.W., G.J. Faulkner, and M.E. Dinger, *Overcoming challenges and dogmas to understand the functions of pseudogenes.* Nat Rev Genet, 2020. **21**(3): p. 191-201.

205. Li, L., et al., *Targeted disruption of Hotair leads to homeotic transformation and gene derepression.* Cell Rep, 2013. **5**(1): p. 3-12.

206. Hayashi, H., et al., *The OCT4 pseudogene POU5F1B is amplified and promotes an aggressive phenotype in gastric cancer.* Oncogene, 2015. **34**(2): p. 199-208.

207. Yu, J., et al., *The Octamer-Binding Transcription Factor 4 (OCT4) Pseudogene, POU Domain Class 5 Transcription Factor 1B (POU5F1B), is Upregulated in Cervical Cancer and Down-Regulation Inhibits Cell Proliferation and Migration and Induces Apoptosis in Cervical Cancer Cell Lines.* Med Sci Monit, 2019. **25**: p. 1204-1213.

208. Pan, Y., et al., *POU5F1B promotes hepatocellular carcinoma proliferation by activating AKT.* Biomed Pharmacother, 2018. **100**: p. 374-380.

209. Saha, S.K., et al., *Systematic expression alteration analysis of master reprogramming factor OCT4 and its three pseudogenes in human cancer and their prognostic outcomes.* Scientific reports, 2018. **8**(1): p. 14806-14806.

210. Suo, G., et al., *Oct4 pseudogenes are transcribed in cancers.* Biochemical and Biophysical Research Communications, 2005. **337**(4): p. 1047-1051.

211. Poursani, E.M., B. Mohammad Soltani, and S.J. Mowla, *Differential Expression of OCT4 Pseudogenes in Pluripotent and Tumor Cell Lines.* Cell journal, 2016. **18**(1): p. 28-36.

212. Orford, K.W. and D.T. Scadden, *Deconstructing stem cell self-renewal: genetic insights into cell-cycle regulation.* Nature Reviews Genetics, 2008. **9**(2): p. 115-128.

213. Noctor, S.C., V. Martinez-Cerdeno, and A.R. Kriegstein, *Distinct behaviors of neural stem and progenitor cells underlie cortical neurogenesis.* J Comp Neurol, 2008. **508**(1): p. 28-44.

214. Lui, J.H., D.V. Hansen, and A.R. Kriegstein, *Development and evolution of the human neocortex.* Cell, 2011. **146**(1): p. 18-36.

215. Lee, S., et al., *Polo Kinase Phosphorylates Miro to Control ER-Mitochondria Contact Sites and Mitochondrial Ca(2+) Homeostasis in Neural Stem Cell Development.* Dev Cell, 2016. **37**(2): p. 174-189.

216. Cabianca, D.S., et al., *A long ncRNA links copy number variation to a polycomb/trithorax epigenetic switch in FSHD muscular dystrophy.* Cell, 2012. **149**(4): p. 819-31.

217. Yi, W., et al., *Bioengineered miR-328-3p modulates GLUT1-mediated glucose uptake and metabolism to exert synergistic antiproliferative effects with chemotherapeutics.* Acta Pharm Sin B, 2020. **10**(1): p. 159-170.

218. Delic, S., et al., *MiR-328 promotes glioma cell invasion via SFRP1-dependent Wnt-signaling activation.* Neuro-Oncology, 2013. **16**(2): p. 179-190.

219. Mulligan, K.A. and B.N. Cheyette, *Wnt signaling in vertebrate neural development and function.* J Neuroimmune Pharmacol, 2012. **7**(4): p. 774-87.

220. Cesana, M., et al., *A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA.* Cell, 2011. **147**(2): p. 358-69.

221. Hansen, T.B., et al., *Natural RNA circles function as efficient microRNA sponges.* Nature, 2013. **495**(7441): p. 384-8.

222. Sandberg, R., et al., *Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites.* Science, 2008. **320**(5883): p. 1643-7.

223. Lan, Y., et al., *Long noncoding RNA OCC-1 suppresses cell growth through destabilizing HuR protein in colorectal cancer.* Nucleic Acids Res, 2018. **46**(11): p. 5809-5821.

224. Doganer, B.A., L.K.Q. Yan, and H. Youk, *Autocrine Signaling and Quorum Sensing: Extreme Ends of a Common Spectrum.* Trends Cell Biol, 2016. **26**(4): p. 262-271.

225. Wang, M. and P.J. Casey, *Protein prenylation: unique fats make their mark on biology.* Nature Reviews Molecular Cell Biology, 2016. **17**(2): p. 110-122.

226. Sidarala, V. and A. Kowluru, *Exposure to chronic hyperglycemic conditions results in Ras-related C3 botulinum toxin substrate 1 (Rac1)-mediated activation of p53 and ATM kinase in pancreatic beta-cells.* Apoptosis, 2017. **22**(5): p. 597-607.

227. Kim, J. and P.K. Wong, *Loss of ATM impairs proliferation of neural stem cells through oxidative stress-mediated p38 MAPK signaling.* Stem Cells, 2009. **27**(8): p. 1987-98.

228. Salem, N.A., et al., *A novel Oct4/Pou5f1-like non-coding RNA controls neural maturation and mediates developmental effects of ethanol.* Neurotoxicology and Teratology, 2021. **83**: p. 106943.

229. Zablotsky, B., et al., *Prevalence and trends of developmental disabilities among children in the United States: 2009–2017.* Pediatrics, 2019. **144**(4): p. e20190811.

230. Modabbernia, A., E. Velthorst, and A. Reichenberg, *Environmental risk factors for autism: an evidence-based review of systematic reviews and meta-analyses.* Mol Autism, 2017. **8**: p. 13.

231. Umer, A., et al., *Prevalence of alcohol use in late pregnancy.* Pediatr Res, 2020. **88**(2): p. 312-319.

232. Stevens, S.A., et al., *Autism characteristics in children with fetal alcohol spectrum disorders.* Child Neuropsychol, 2013. **19**(6): p. 579-87.

233. Bishop, S., S. Gahagan, and C. Lord, *Re-examining the core features of autism: a comparison of autism spectrum disorder and fetal alcohol spectrum disorder.* Journal of Child Psychology and Psychiatry, 2007. **48**(11): p. 1111-1121.

234. Mukherjee, R., et al., *Autism and autistic traits in people exposed to heavy prenatal alcohol: data from a clinical series of 21 individuals and nested case control study.* Advances in Mental Health and Intellectual Disabilities, 2011. **5**(1): p. 42-49.

235. Infante, M.A., et al., *Objective assessment of ADHD core symptoms in children with heavy prenatal alcohol exposure.* Physiol Behav, 2015. **148**: p. 45-50.

236. Riikonen, R.S., et al., *Deep serotonergic and dopaminergic structures in fetal alcoholic syndrome: a study with nor-beta-CIT-single-photon emission computed tomography and magnetic resonance imaging volumetry.* Biol Psychiatry, 2005. **57**(12): p. 1565-72.

237. Lange, S., et al., *Prevalence of externalizing disorders and Autism Spectrum Disorders among children with Fetal Alcohol Spectrum Disorder: systematic review and meta-analysis.* Biochem Cell Biol, 2017: p. 1-11.

238. Popova, S., et al., *Comorbidity of fetal alcohol spectrum disorder: a systematic review and meta-analysis.* Lancet, 2016. **387**(10022): p. 978-987.

239. Norman, A.L., et al., *Neuroimaging and fetal alcohol spectrum disorders.* Dev Disabil Res Rev, 2009. **15**(3): p. 209-17.

240. O'Leary-Moore, S.K., et al., *Magnetic resonance microscopy-based analyses of the brains of normal and ethanol-exposed fetal mice.* Birth Defects Res A Clin Mol Teratol, 2010. **88**(11): p. 953-64.

241. Workman, A.D., et al., *Modeling transformations of neurodevelopmental sequences across mammalian species.* J Neurosci, 2013. **33**(17): p. 7368-83.

242. Takahashi, T., et al., *Sequence of neuron origin and neocortical laminar fate: relation to cell cycle of origin in the developing murine cerebral wall.* J Neurosci, 1999. **19**(23): p. 10357-71.

243. Takahashi, T., R.S. Nowakowski, and V.S. Caviness, Jr., *The cell cycle of the pseudostratified ventricular epithelium of the embryonic murine cerebral wall.* J Neurosci, 1995. **15**(9): p. 6046-57.

244. Sudheendran, N., et al., *Comparative assessments of the effects of alcohol exposure on fetal brain development using optical coherence tomography and ultrasound imaging.* J Biomed Opt, 2013. **18**(2): p. 20506.

245. Loo, L., et al., *Single-cell transcriptomic analysis of mouse neocortical development.* Nat Commun, 2019. **10**(1): p. 134.

246. Renthal, W., et al., *Characterization of human mosaic Rett syndrome brain tissue by single-nucleus RNA sequencing.* Nature Neuroscience, 2018. **21**(12): p. 1670-1679.

247. Varlinskaya, E.I. and S.M. Mooney, *Acute exposure to ethanol on gestational day 15 affects social motivation of female offspring.* Behav Brain Res, 2014. **261**: p. 106-9.

248. Raineki, C., et al., *Short- and long-term effects of stress during adolescence on emotionality and HPA function of animals exposed to alcohol prenatally.* Psychoneuroendocrinology, 2016. **74**: p. 13-23.

249. Loo, L., et al., *Single-cell transcriptomic analysis of mouse neocortical development.* Nature Communications, 2019. **10**(1): p. 134.

250. Butler, A., et al., *Integrating single-cell transcriptomic data across different conditions, technologies, and species.* Nature Biotechnology, 2018. **36**(5): p. 411-420.

251. Stuart, T., et al., *Comprehensive integration of single-cell data.* Cell, 2019. **177**(7): p. 1888-1902. e21.

252. Cole, T., et al., *Lennon Niall J, Livak Kenneth J, Mikkelsen Tarjei S, Rinn John L. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells.* Nature Biotechnology, 2014. **32**(4): p. 381-386.

253. Qiu, X., et al., *Reversed graph embedding resolves complex single-cell trajectories.* Nature methods, 2017. **14**(10): p. 979.

254. Cao, J., et al., *The single-cell transcriptional landscape of mammalian organogenesis.* Nature, 2019. **566**(7745): p. 496-502.

255. McInnes, L., J. Healy, and J. Melville, *Umap: Uniform manifold approximation and projection for dimension reduction.* arXiv preprint arXiv:1802.03426, 2018.

256. Traag, V.A., L. Waltman, and N.J. van Eck, *From Louvain to Leiden: guaranteeing well-connected communities.* Scientific reports, 2019. **9**(1): p. 1-12.

257. Miao, Z., et al., *DEsingle for detecting three types of differential expression in single-cell RNA-seq data.* Bioinformatics, 2018. **34**(18): p. 3223-3224.

258. Wang, T., et al., *Comparative analysis of differential gene expression analysis tools for single-cell RNA sequencing data.* BMC Bioinformatics, 2019. **20**(1): p. 40.

259. Langfelder, P. and S. Horvath, *WGCNA: an R package for weighted correlation network analysis.* BMC Bioinformatics, 2008. **9**(1): p. 559.

260. Holmberg Olausson, K., et al., *Prominin-1 (CD133) defines both stem and non-stem cell populations in CNS development and gliomas.* PloS one, 2014. **9**(9): p. e106694-e106694.

261. Englund, C., et al., *Pax6, Tbr2, and Tbr1 Are Expressed Sequentially by Radial Glia, Intermediate Progenitor Cells, and Postmitotic Neurons in Developing Neocortex.* The Journal of Neuroscience, 2005. **25**(1): p. 247-251.

262. Johnson, M.B., et al., *Single-cell analysis reveals transcriptional heterogeneity of neural progenitors in human cortex.* Nature neuroscience, 2015. **18**(5): p. 637-646.

263. Vinci, L., et al., *Immunohistochemical markers of neural progenitor cells in the early embryonic human cerebral cortex.* European journal of histochemistry : EJH, 2016. **60**(1): p. 2563-2563.

264. Reithmeier, R.A.F., et al., *Band 3, the human red cell chloride/bicarbonate anion exchanger (AE1, SLC4A1), in a structural context.* Biochimica et Biophysica Acta (BBA) - Biomembranes, 2016. **1858**(7, Part A): p. 1507-1532.

265. Chiabrando, D., S. Mercurio, and E. Tolosano, *Heme and erythropoieis: more than a structural role.* Haematologica, 2014. **99**(6): p. 973-983.

266. Kingsley, P.D., et al., *"Maturational" globin switching in primary primitive erythroid cells.* Blood, 2006. **107**(4): p. 1665-1672.

267. Telley, L., et al., *Sequential transcriptional waves direct the differentiation of newborn neurons in the mouse neocortex.* Science, 2016. **351**(6280): p. 1443-1446.

268. Chen, G., et al., *Semaphorin-3A guides radial migration of cortical neurons during development.* Nature Neuroscience, 2008. **11**(1): p. 36-44.

269. Vigneswara, V. and Z. Ahmed, *The Role of Caspase-2 in Regulating Cell Fate.* Cells, 2020. **9**(5): p. 1259.

270. Kimbrough, A., et al., *Brain-wide functional architecture remodeling by alcohol dependence and abstinence.* Proceedings of the National Academy of Sciences, 2020. **117**(4): p. 2149-2159.

271. Parikshak, N.N., M.J. Gandal, and D.H. Geschwind, *Systems biology and gene networks in neurodevelopmental and neurodegenerative disorders.* Nature Reviews Genetics, 2015. **16**(8): p. 441-458.

272. Tebbenkamp, A.T.N., et al., *The developmental transcriptome of the human brain: implications for neurodevelopmental disorders.* Current opinion in neurology, 2014. **27**(2): p. 149-156.

273. Mariani, J., et al., *FOXG1-Dependent Dysregulation of GABA/Glutamate Neuron Differentiation in Autism Spectrum Disorders.* Cell, 2015. **162**(2): p. 375-390.

274. Pan, J., et al., *Immuno-characterization of the switch of peptide elongation factors eEF1A-1/EF-1alpha and eEF1A-2/S1 in the central nervous system during mouse development.* Brain Res Dev Brain Res, 2004. **149**(1): p. 1-8.

275. Cho, S.J., et al., *Translation elongation factor-1A1 (eEF1A1) localizes to the spine by domain III.* BMB Rep, 2012. **45**(4): p. 227-32.

276. Berletch, J.B., F. Yang, and C.M. Disteche, *Escape from X inactivation in mice and humans.* Genome biology, 2010. **11**(6): p. 213-213.

277. Carrel, L. and H.F. Willard, *X-inactivation profile reveals extensive variability in X-linked gene expression in females.* Nature, 2005. **434**(7031): p. 400-404.

278. Chi, L.-H., et al., *Global Proteomics-based Identification and Validation of Thymosin Beta-4 X-Linked as a Prognostic Marker for Head and Neck Squamous Cell Carcinoma.* Scientific Reports, 2017. **7**(1): p. 9031.

279. Caviness, V.S., Jr., *Neocortical histogenesis in normal and reeler mice: a developmental study based upon [3H]thymidine autoradiography.* 1982. **256**(3): p. 293-302.

280. Telley, L., et al., *Sequential transcriptional waves direct the differentiation of newborn neurons in the mouse neocortex.* Science, 2016. **351**(6280): p. 1443-6.

281. Roussotte, F.F., et al., *Regional brain volume reductions relate to facial dysmorphology and neurocognitive function in fetal alcohol spectrum disorders.* Hum Brain Mapp, 2012. **33**(4): p. 920-37.

282. Thornberry, N.A. and Y. Lazebnik, *Caspases: enemies within.* Science, 1998. **281**(5381): p. 1312-6.

283. Miller, M.W., *Effects of prenatal exposure to ethanol on neocortical development: II. Cell proliferation in the ventricular and subventricular zones of the rat.* Journal of Comparative Neurology, 1989. **287**(3): p. 326-338.

284. Fournier, M.J., et al., *Inactivation of the mTORC1-eukaryotic translation initiation factor 4E pathway alters stress granule formation.* Mol Cell Biol, 2013. **33**(11): p. 2285-301.

285. Berres, M.E., et al., *Transcriptome Profiling Identifies Ribosome Biogenesis as a Target of Alcohol Teratogenicity and Vulnerability during Early Embryogenesis.* PLOS ONE, 2017. **12**(1): p. e0169351.

286. Mohammad, L., et al., *Protein synthesis and translational control in neural stem cell development and neurogenesis*, in *The Oxford Handbook of Neuronal Protein Synthesis*. 2019.

287. Sado, T., Y. Hoki, and H. Sasaki, *Tsix silences Xist through modification of chromatin structure.* Dev Cell, 2005. **9**(1): p. 159-65.

288. Ohhata, T., et al., *Crucial role of antisense transcription across the Xist promoter in Tsix-mediated Xist chromatin modification.* Development, 2008. **135**(2): p. 227-35.

289. Porter, R.S., F. Jaamour, and S. Iwase, *Neuron-specific alternative splicing of transcriptional machineries: Implications for neurodevelopmental disorders.* Molecular and Cellular Neuroscience, 2018. **87**: p. 35-45.

290. Sanders, S.J., G.B. Schwartz, and K.K.-H. Farh, *Clinical impact of splicing in neurodevelopmental disorders.* Genome Medicine, 2020. **12**(1): p. 36.

291. Smith, R. and W. Sadee, *Synaptic Signaling and Aberrant RNA Splicing in Autism Spectrum Disorders.* Frontiers in Synaptic Neuroscience, 2011. **3**(1).

292. Lehalle, D., et al., *A review of craniofacial disorders caused by spliceosomal defects.* Clinical Genetics, 2015. **88**(5): p. 405-415.

293. Garic, A., M.E. Berres, and S.M. Smith, *High-throughput transcriptome sequencing identifies candidate genetic modifiers of vulnerability to fetal alcohol spectrum disorders.* Alcoholism, clinical and experimental research, 2014. **38**(7): p. 1874-1882.

294. Richter, F., et al., *Neurons express hemoglobin alpha- and beta-chains in rat and human brains.* J Comp Neurol, 2009. **515**(5): p. 538-47.

295. Brown, N., et al., *Neuronal Hemoglobin Expression and Its Relevance to Multiple Sclerosis Neuropathology.* J Mol Neurosci, 2016. **59**(1): p. 1-17.

296. Schaffner, S.L., et al., *Neonatal Alcohol Exposure in Mice Induces Select Differentiation- and Apoptosis-Related Chromatin Changes Both Independent of and Dependent on Sex.* Front Genet, 2020. **11**: p. 35.

297. Chater-Diehl, E.J., B.I. Laufer, and S.M. Singh, *Changes to histone modifications following prenatal alcohol exposure: An emerging picture.* Alcohol, 2017. **60**: p. 41-52.

298. Kadoch, C., et al., *Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy.* Nature genetics, 2013. **45**(6): p. 592-601.

299. Guda, S., et al., *miRNA-embedded shRNAs for Lineage-specific BCL11A Knockdown and Hemoglobin F Induction.* Mol Ther, 2015. **23**(9): p. 1465-74.

300. Sankaran, V.G., et al., *Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A.* Science, 2008. **322**(5909): p. 1839-42.

301. Xu, J., et al., *Corepressor-dependent silencing of fetal hemoglobin expression by BCL11A.* Proc Natl Acad Sci U S A, 2013. **110**(16): p. 6518-23.

302. Dias, C., et al., *BCL11A Haploinsufficiency Causes an Intellectual Disability Syndrome and Dysregulates Transcription.* American journal of human genetics, 2016. **99**(2): p. 253-274.

303. Basak, A., et al., *BCL11A deletions result in fetal hemoglobin persistence and neurodevelopmental alterations.* J Clin Invest, 2015. **125**(6): p. 2363-8.

304. Funnell, A.P., et al., *2p15-p16.1 microdeletions encompassing and proximal to BCL11A are associated with elevated HbF in addition to neurologic impairment.* Blood, 2015. **126**(1): p. 89-93.

305. Kotkoskie, L.A. and S. Norton, *Prenatal Brain Malformations following Acute Ethanol Exposure in the Rat.* Alcoholism: Clinical and Experimental Research, 1988. **12**(6): p. 831-836.

306. Liedtke, S., et al., *Oct4 and Its Pseudogenes Confuse Stem Cell Research.* Cell Stem Cell, 2007. **1**(4): p. 364-366.

307. Milligan, M.J. and L. Lipovich, *Pseudogene-derived lncRNAs: emerging regulators of gene expression.* Frontiers in Genetics, 2015. **5**(476).

308. *An integrated encyclopedia of DNA elements in the human genome.* Nature, 2012. **489**(7414): p. 57-74.

309. Davis, C.A., et al., *The Encyclopedia of DNA elements (ENCODE): data portal update.* Nucleic Acids Res, 2018. **46**(D1): p. D794-d801.

310. Subramanian, A., et al., *A next generation connectivity map: L1000 platform and the first 1,000,000 profiles.* Cell, 2017. **171**(6): p. 1437-1452. e17.

311. Ferguson, L.B., et al., *Genome-Wide Expression Profiles Drive Discovery of Novel Compounds that Reduce Binge Drinking in Mice.* Neuropsychopharmacology, 2018. **43**(6): p. 1257-1266.

312. Loda, A., et al., *Genetic and epigenetic features direct differential efficiency of Xist-mediated silencing at X-chromosomal and autosomal locations.* Nature Communications, 2017. **8**(1): p. 690.

313. La Manno, G., et al., *RNA velocity of single cells.* Nature, 2018. **560**(7719): p. 494-498.

314. Pott, S. and J.D. Lieb, *Single-cell ATAC-seq: strength in numbers.* Genome Biology, 2015. **16**(1): p. 172.

APPENDIX 1

Appendix 1 a: Sequences of RT-PCR primers for mouse Oct4/Pou5f1 and presumptive mouse Oct4 pseudogenes and PCR product size:

Target	Primer	Sequence (5'-3')	Product size
Oct4/Pou5f1	Forward	AGCTGCTGAAGCAGAAGAGGATCA	85
	Reverse	AACACCTTTCCAAAGAGAACGCCC	
Oct4pg9	Forward	TGCAGGAGCTAGAATGGGGA	291
	Reverse	AATGGCCTGGGGTCAAATGT	
Oct4pg14	Forward	TTCTGAAGTGCCCGAAGTCC	327
	Reverse	GCCAGGGTAGTGACAGGAAC	
Oct4pg6	Forward	GAGTGTGGTGAAGAGGGAGC	208
	Reverse	CTGGAGAAGGATGTGGCTCG	
Oct4pgX (primer 1)	Forward	TGGAGCATGAGTGGAGAGGA	546
	Reverse	CCTCCTGGTGCAGGTTACAG	
Oct4pgX (primer2)	Forward	TGGAGAAGTGGGTGGAGGAA	119
	Reverse	AGGTTCCCCTTCTCAGGGTT	

Appendix 1 b: Smartpool® siRNAs' sequences (Dharmacon) targeting m*Oct4pg9*

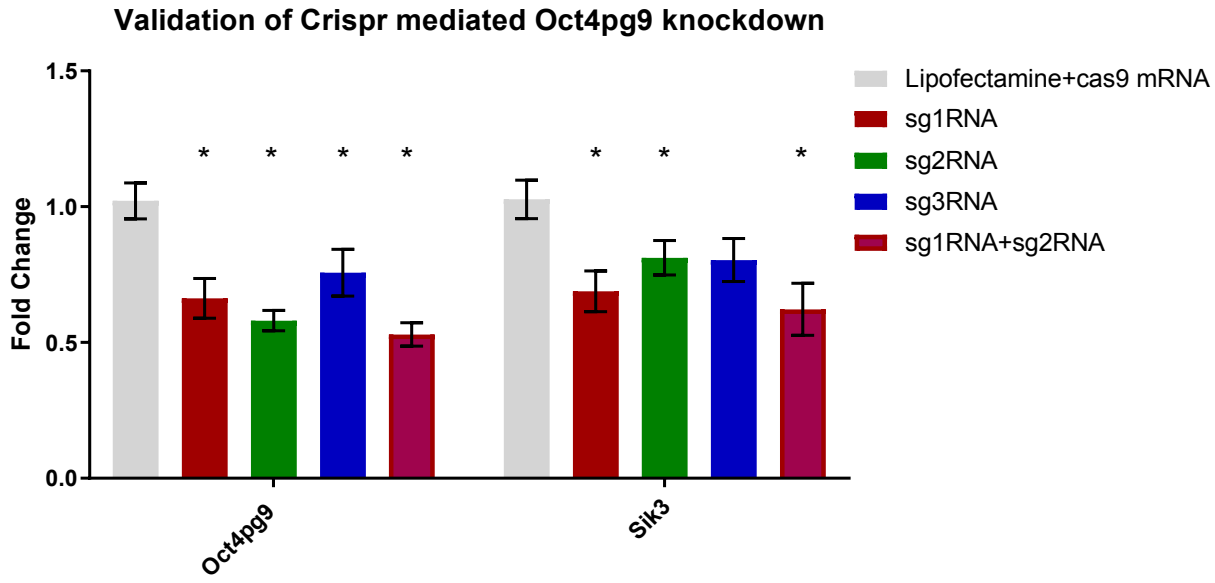
UUUAGCUGAAUACCUUCCCUU
AUCACCUUGUGCUCCUAUCUU
UCACAAUAUGUUAUCCUUCUU
CAACUUAGGGUAUUGGUCCUU

Appendix 1 c: CRISPR guide RNAs' sequences and double strand break genomic location of each of the sgRNAs.

	Sequence	Break location
sgRNA-1	AAGCCAGGUGUCUGACCAUG	Chr 9: 46161028
sgRNA-2	GAAAGAACACAGACCGCCCC	Chr 9: 46161792
sgRNA-3	GCACCAAAATGGGGGCCTTG	Chr 9: 46161819

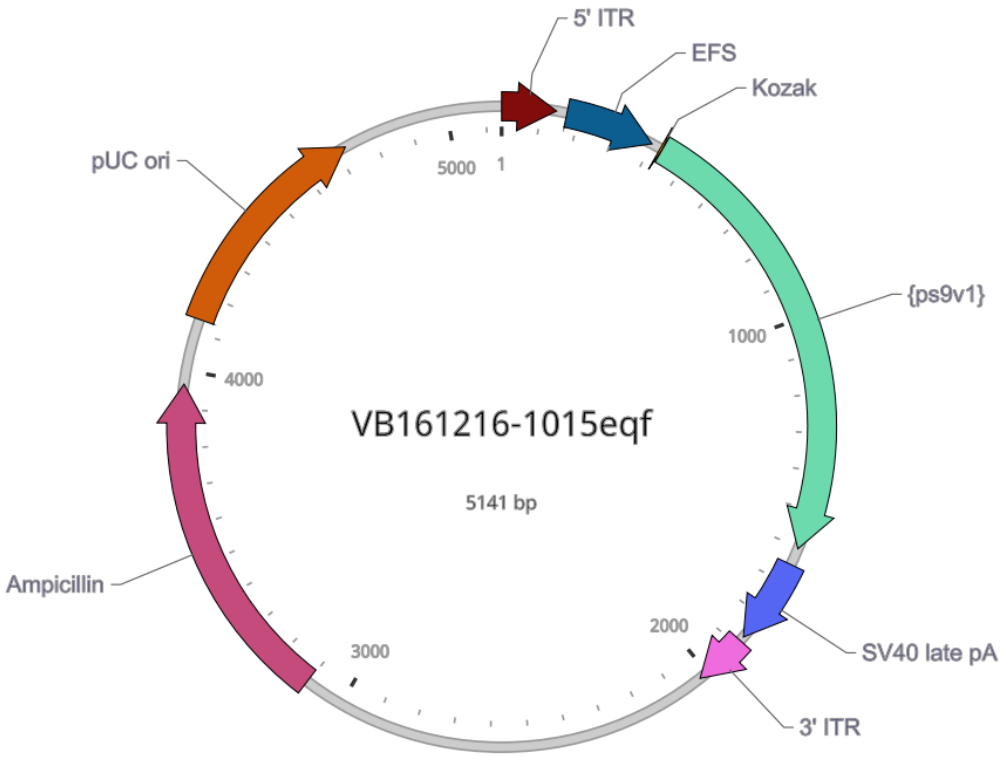
Appendix 1 d: Validation of CRISPR mediated knockdown

fold change of m*Oct4pg9 lncRNA and Sik3 expression after CRISPR mediated* m*Oct4pg9 lncRNA knockdown using three different single guide RNA designs*



	mOct4pg9 ncRNA (Mean decrease)	**P-value**	**Sik3 mRNA (Mean decrease)**	**P-value**
Sg1RNA	0.359±0.098	$t_{(21)}$=3.657, p=0.09813	0.339±0.103	$t_{(21)}$=3.288, p=0.0035
Sg2RNA	0.441±0.0759	$t_{(22)}$=5.812, p=0.000008	0.215±0.095	$t_{(22)}$=2.26, p=0.034
Sg3RNA	0.264±0.1163	$t_{(18)}$=2.467, p=0.023	0.223±0.108	$t_{(18)}$=2.068, p=0.053

Appendix 1 e: m*Oct4pg9 overexpression plasmid vector map:*



Appendix 1 f: m*Oct4pg9 lncRNA biotinylated capture probes' sequences:*

Scrambled probe (control)	GCTGGTCGTTACTCAGTGCC
Probe 1	CTGCACTAGGGTCTCTGCTG
Probe 2	CTACAGCCTGAGGCACCAAA
Probe 3	GAACTTCATCCCGAAGCCCT

Appendix 1 g: Primary and secondary antibodies used in Western immunoblot assays

Primary Antibodies:

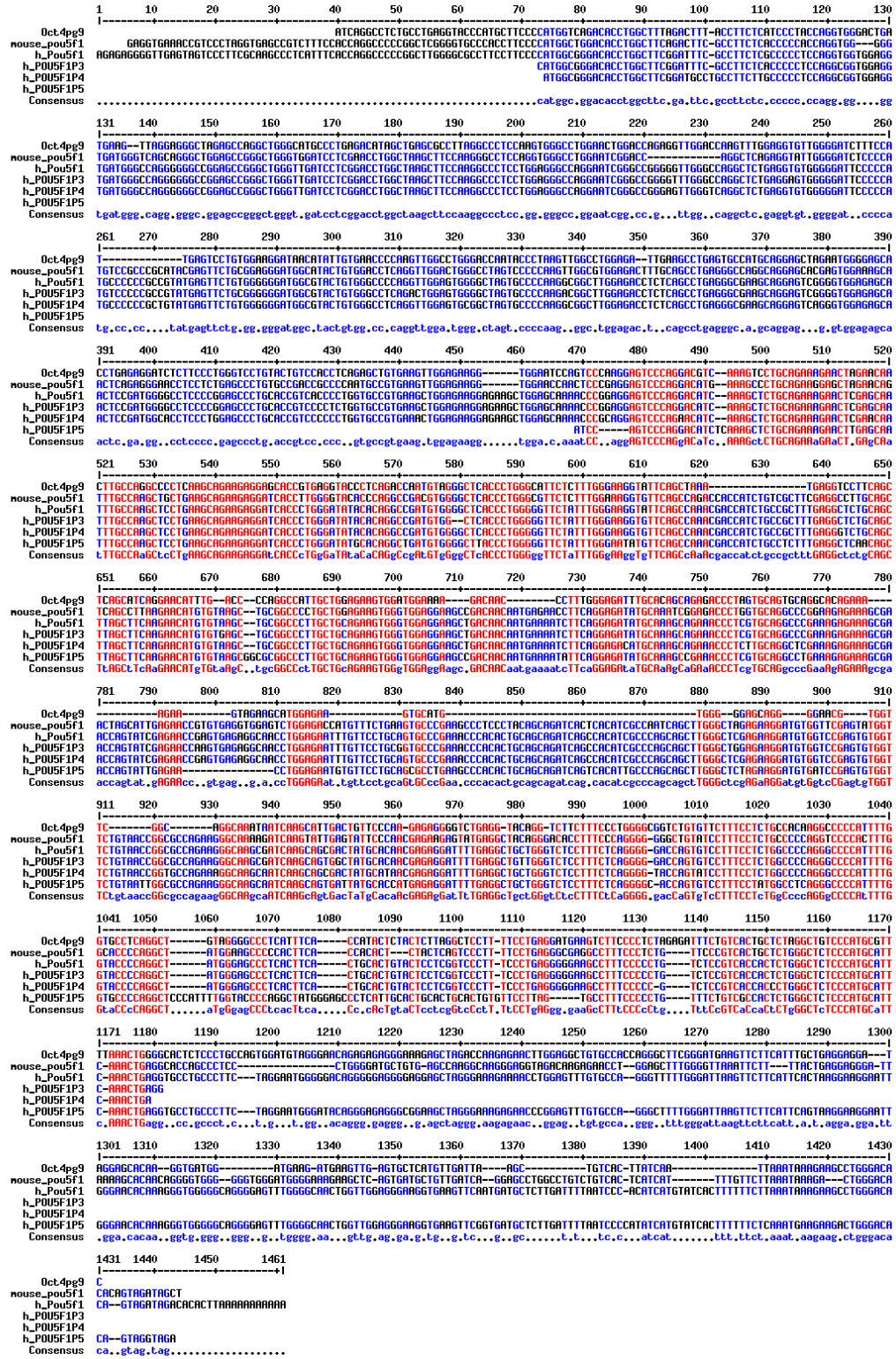
Target	**Host**	**Antibody (Supplier)**	**Concentration**	**Incubation time**
Oct4	Rabbit	Ab19857 (Abcam)	1:500	Overnight
NRSF/Rest	Rabbit	Bs-2590R (Bioss)	1:1,000	Overnight
GFAP	Mouse	MA5-12023 (Thermofisher)	1:1,000	Overnight
Olig2	Rabbit	P21954 (Thermofisher)	1:500	Overnight
Tubulin	Mouse	T5168 (Sigma)	1:10,000	1 hour

Secondary Antibodies:

Reactivity	**Host**	**Tag**	**Supplier**	**Dilution**
Rabbit	Goat	IRDye 680	Licor	1:15,000
Mouse	Goat	IRDye 800	Licor	1:15,000
Rabbit	Goat	HRP	Thermofisher	1:1,000
Mouse	Goat	HRP	Thermofisher	1:1,000

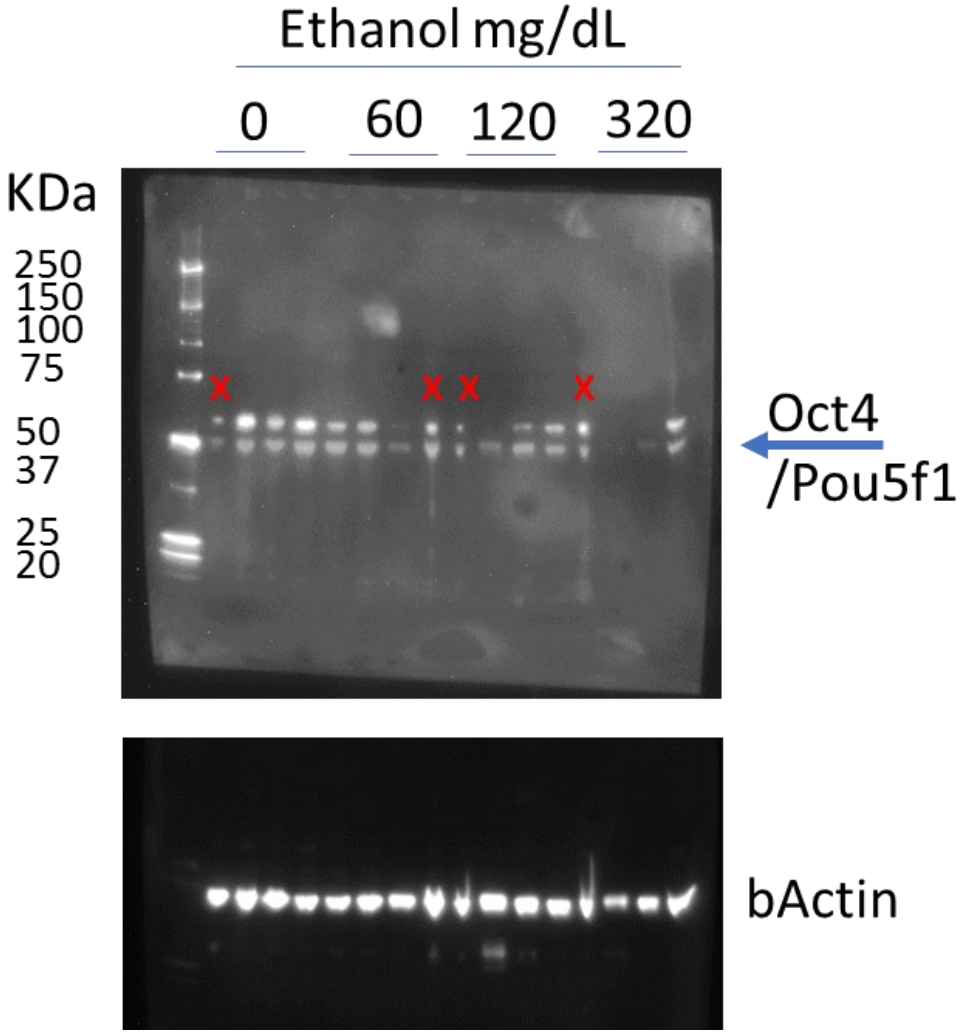
Appendix 1 h: Sequence homology between mouse and human OCT4/POU5F1, m*Oct4pg9, and human Oct4 pseudogenes.*

Figure shows the sequence alignment (MultiAlin software), from top to bottom, between m*Oct4pg9 ncRNA, mouse Oct4/Pou5f1 mRNA, human OCT4/POU5F1 mRNA, and human OCT4/POU5F1 pseudogenes, POU5F1P3,*

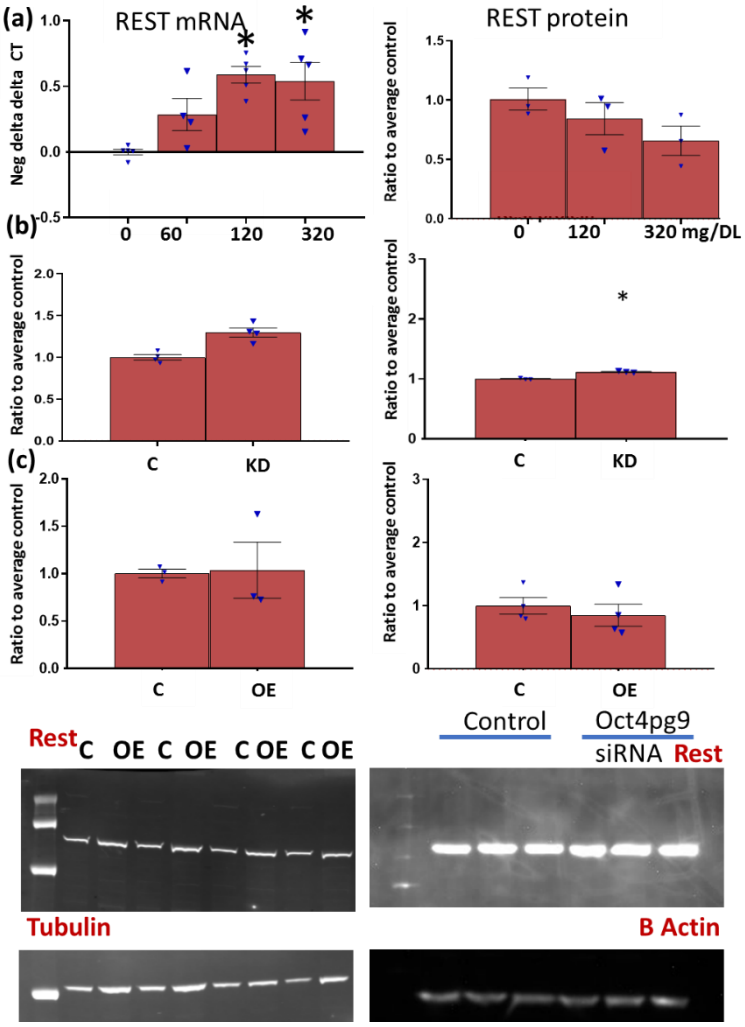


POU5F1P4, and POU5F1P5 with the consensus sequence shown below. Red and blue colors indicate high and low consensus respectively.

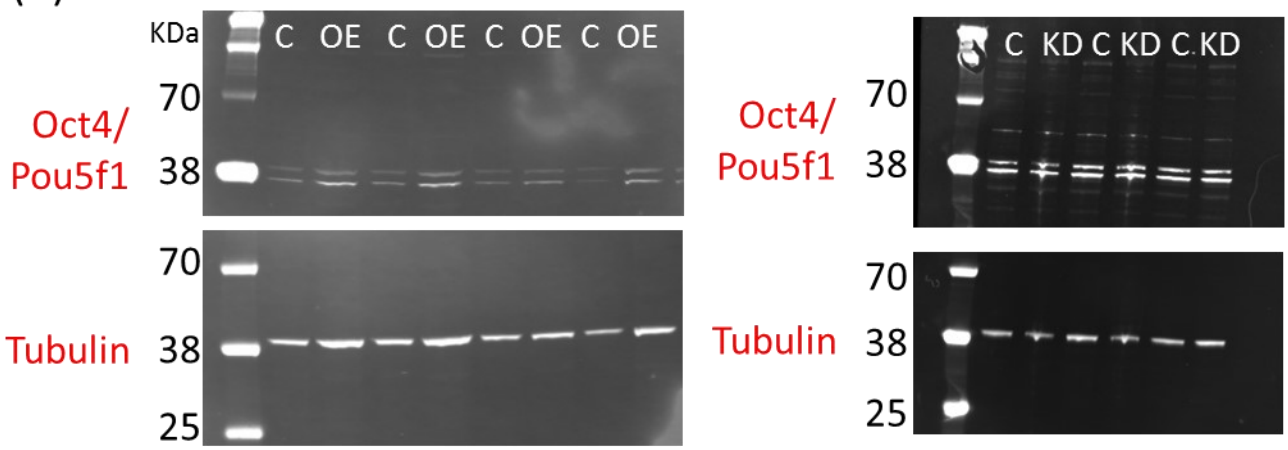
Appendix 1 i: Uncropped Western immunoblot showing the effect of ethanol treatment on OCT4/POU5F1 protein. Lanes marked with 'X' were not included in the analysis.



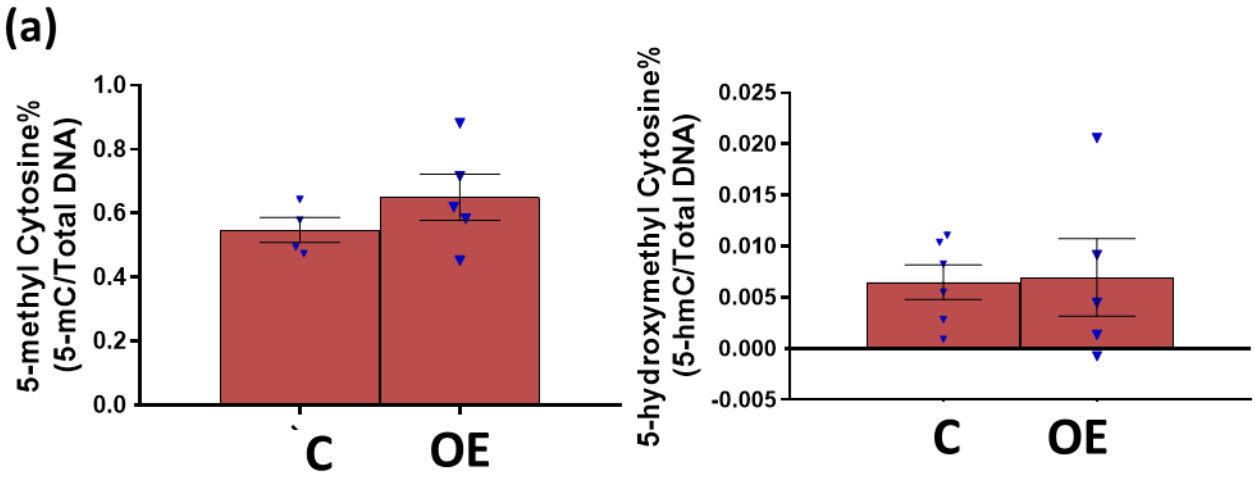
Appendix 1 j: Effects of ethanol treatment and m*Oct4pg9 manipulation on expression of Rest mRNA, protein and Oct4/Pou5f1 protein in neurosphere cultures: Expression of Rest mRNA and protein in neurospheres treated with (a) 0, 60, 120 or 320 mg/dl ethanol. Expression of REST protein in neurosphere cultures (b) treated with siRNA against* m*Oct4pg9 (KD) or control (C) siRNA, and (c) overexpressing* m*Oct4pg9 (OE) compared to treated with control plasmid (C). (d) Western blot immunostaining of REST protein and tubulin/β-actin housekeeping protein in cell lysates from* m*Oct4pg9 overexpressing neurospheres (or neurospheres treated with control plasmid) or neurospheres treated with siRNA against* m*Oct4pg9 (or control siRNA). Bars represent mean ± SEM, individual points represent independent biological replicates, * indicates Fisher LSD post hoc analysis p-value<0.05(a) or unpaired two-tailed t-test p-value<0.05 (b, c). (d) Uncropped Western immunoblots showing the effect of mOct4pg9 overexpression (OE) and knockdown (KD) on OCT4/POU5F1 protein.*



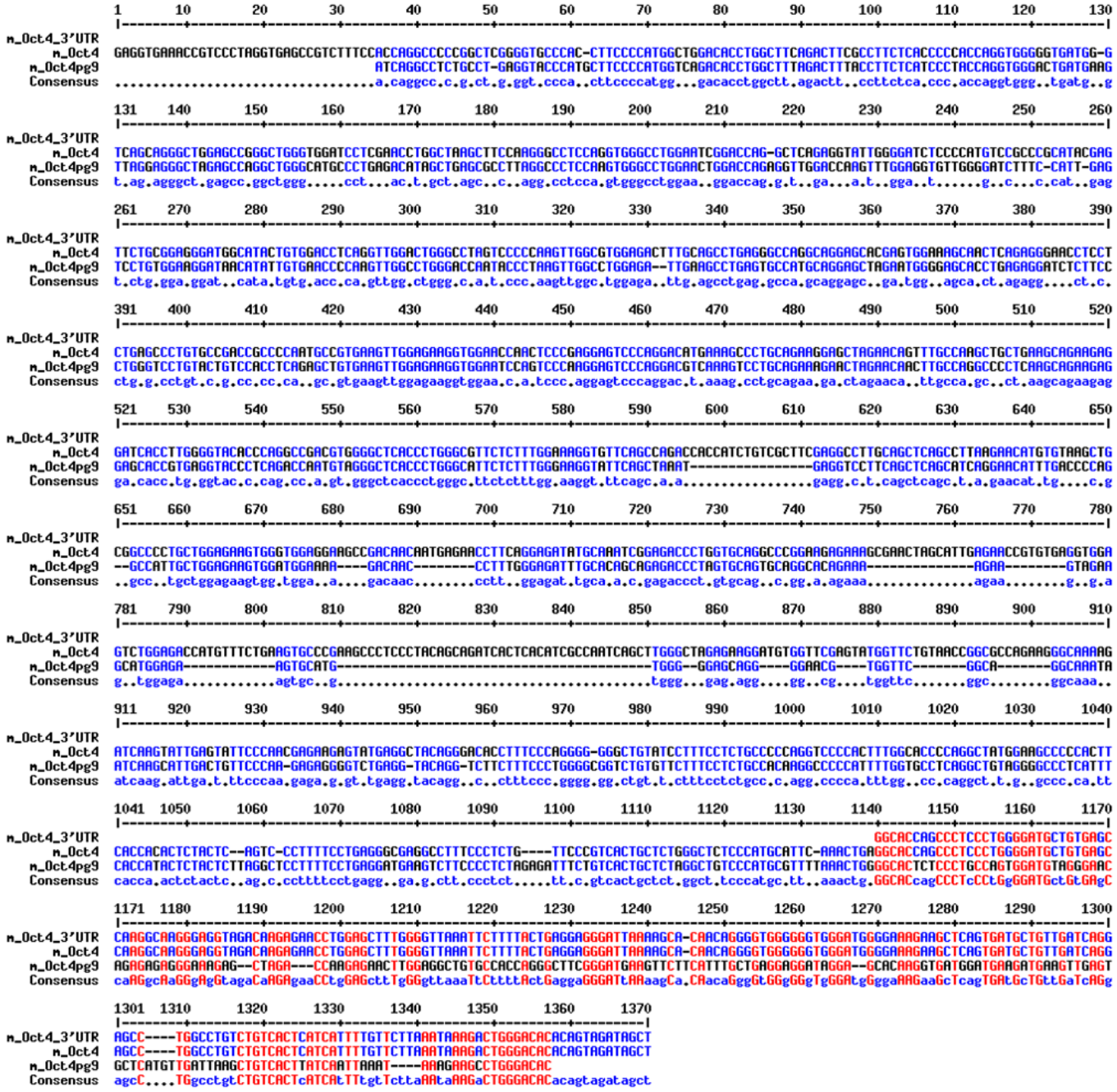
(d)


KDa
C OE C OE C OE C OE
Oct4/
Pou5f1
70
38
Tubulin
70
38
25
C KD C KD C KD
Oct4/
Pou5f1
70
38
Tubulin
70
38
25

Appendix 1 k: Effect of mOct4pg9 *overexpression (OE) on total cellular 5-methyl cytosine and 5-hydroxymethyl cytosine in neurospheres compared to controls (C). Bars represent mean ± SEM, individual points represent independent replicates, * indicates p<0.05, unpaired two-tailed t-test.*



Appendix 1 I: Sequence homology between mouse Oct4/Pou5f1 mRNA 3'UTR and m*Oct4pg9 ncRNA: Figure shows the alignment (MultinAlin software), top to bottom, between mouse Oct4/Pou5f1 3'UTR, Oct4/Pou5f1 mRNA, and* m*Oct4pg9 ncRNA with the consensus sequence shown below. Red and blue colors indicate high and low consensus respectively.*

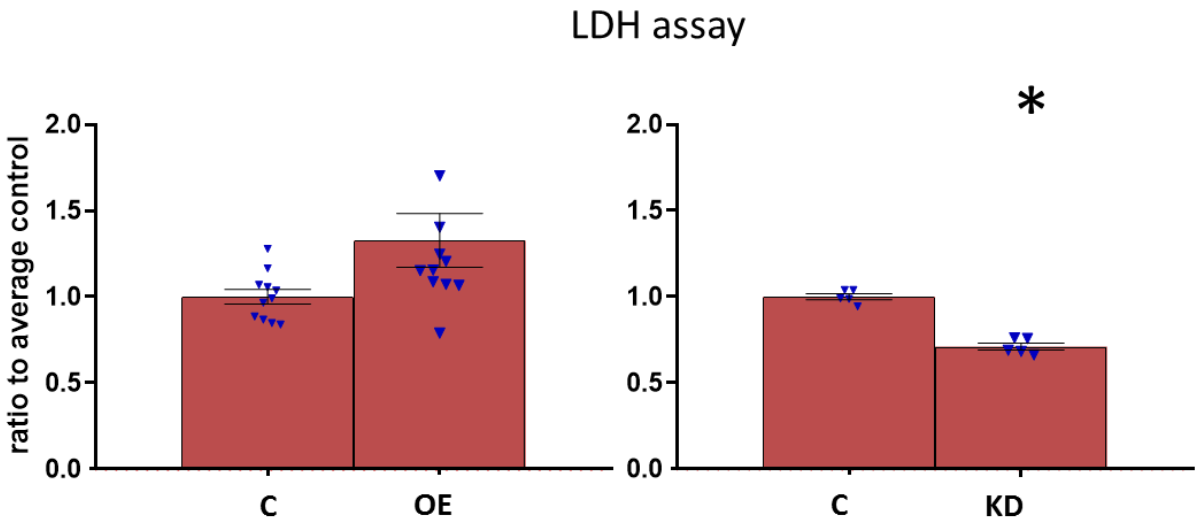


Appendix 1 m: Effect of m*Oct4pg9 overexpression on miRNAs expression. Table shows unpaired t-test p-values and Hedges' 'g' effect sizes of miRNAs with significant expression changes after* m*Oct4pg9 overexpression. Positive Hedges' g effect size indicate that the miRNA is downregulated by* m*Oct4pg9 overexpression and negative effect size indicates upregulation by mOct4pg9 overexpression.*

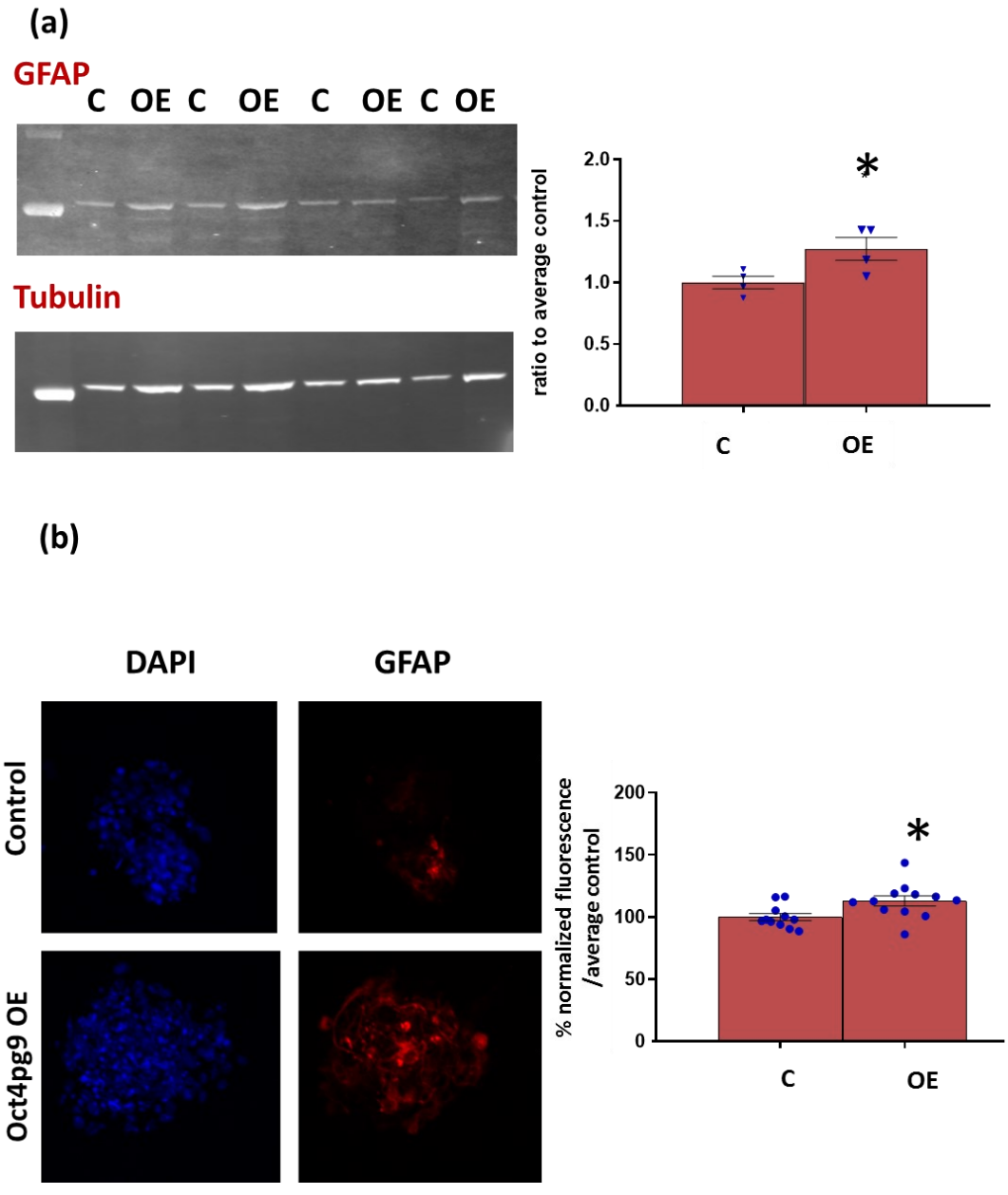
miRNA	Hedges' g effect size (delta CT (OE-ctrl))	p-value (unpaired t-test)
miR-378a-3p	5.100262	0.001452
miR-92a-3p	4.572092	0.002198
miR-320a	3.666798	0.00495
miR-328-3p	3.111464	0.008892
let-7d-3p	2.504039	0.018573
miR-181a-5p	2.492043	0.018872
miR-326	2.312549	0.024007
miR-92b-3p	2.295063	0.024589
let-7c-5p	2.237553	0.02663
miR-423-3p	2.147156	0.030321
let-7b-5p	2.144341	0.030414
miR-331-3p	2.129003	0.031079
miR-652-3p	1.938965	0.041196
miR-149-5p	1.902707	0.043542
miR-107	1.870883	0.045736
let-7i-5p	1.867439	0.045981
miR-98-5p	-1.82322	0.049252

miR-760	-3.1114	0.008883

Appendix 1 n: Effect of m*Oct4pg9 manipulation on neurosphere cell death: Quantification of the effects of* m*Oct4pg9 overexpression (OE, left) and siRNA mediated knockdown (KD, right) on lactate dehydrogenase (LDH) activity, measured by the colorimetric reaction product resulting from LDH activity that was quantified spectrophotometrically at 490 nm and normalized to absorbance at 680 nm for background correction. Normalized absorbance was then compared to the average normalized absorbance from the control group. Bars represent mean ± SEM, individual points represent independent biological replicates, * indicates unpaired two-tailed t-test p-value<0.05.*

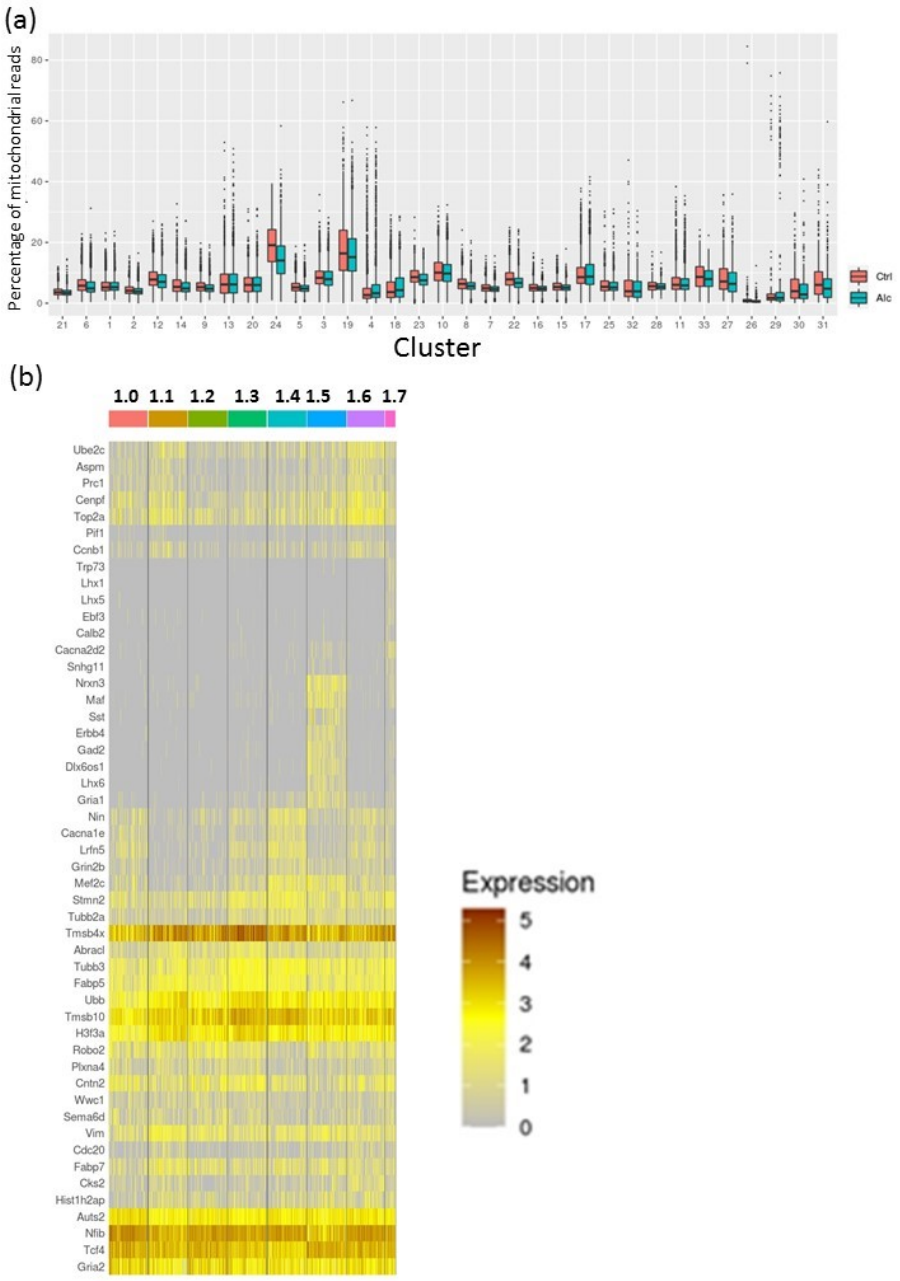


Appendix 1 o: GFAP expression in mOct4pg9 *overexpressing neurospheres: (a) Expression of GFAP protein in neurospheres overexpressing* mOct4pg9 *(OE) compared to neurospheres treated with control plasmid as immunoblot (left) and quantified (right) as GFAP fluorescence normalized to tubulin fluorescence and individual values were compared to average normalized values from the control group. (b) Representative images showing GFAP immunofluorescent staining in control and* mOct4pg9 *overexpressing neurospheres (with DAPI as a nuclear counterstain) (left) and quantification of GFAP immunofluorescence (right) in which GFAP fluorescence was normalized to DAPI stain fluorescence, and percent fluorescence was calculated by dividing normalized values from each neurosphere by the average normalized values of control neurospheres. Bars represent mean ± SEM, individual points represent independent biological replicates, * indicates unpaired two-tailed t-test p-value<0.05.*



APPENDIX 2

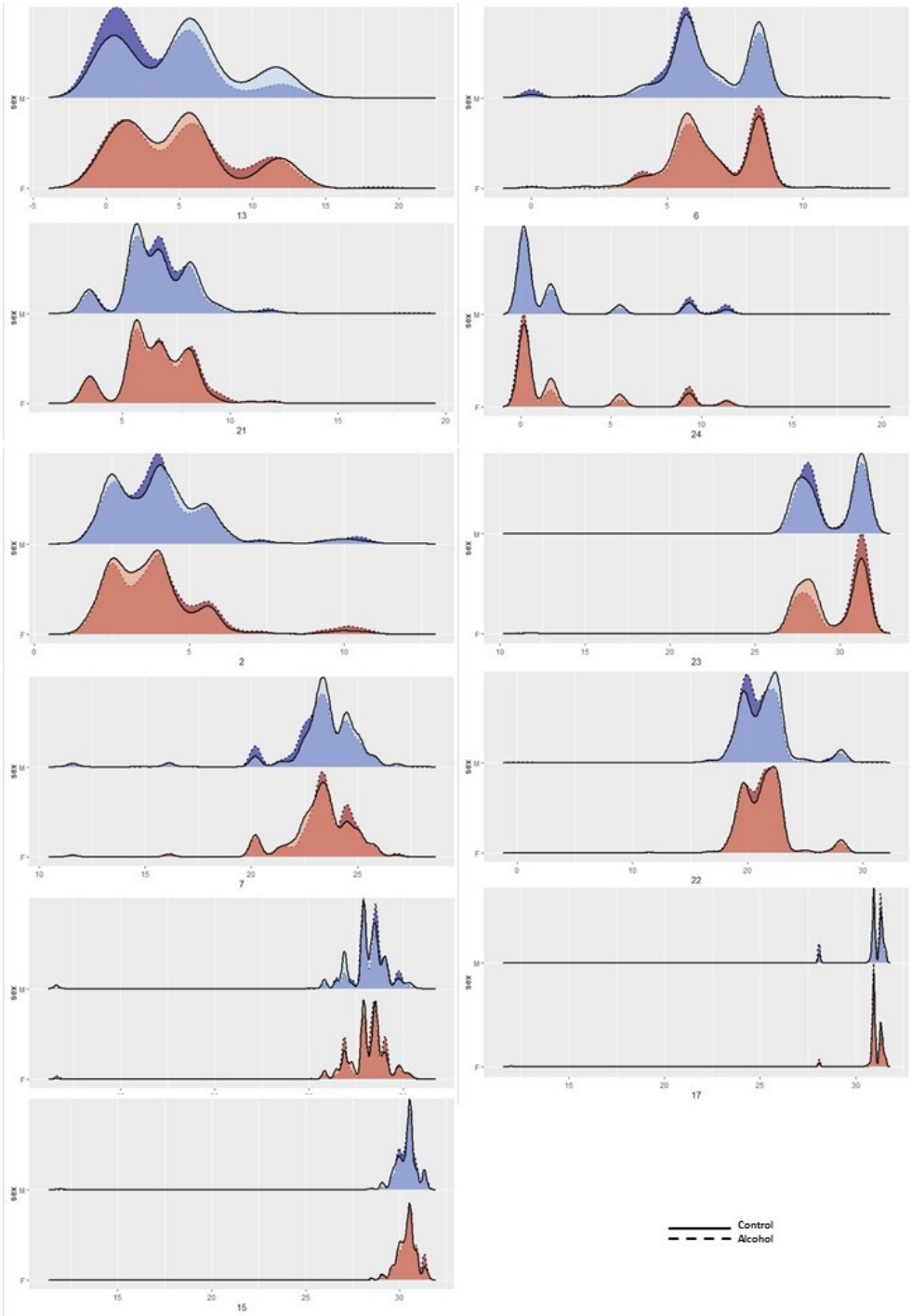
Appendix 2 a:



a) Box plots showing the median and interquartile range of cellular percentage of mitochondrial reads in each cluster.

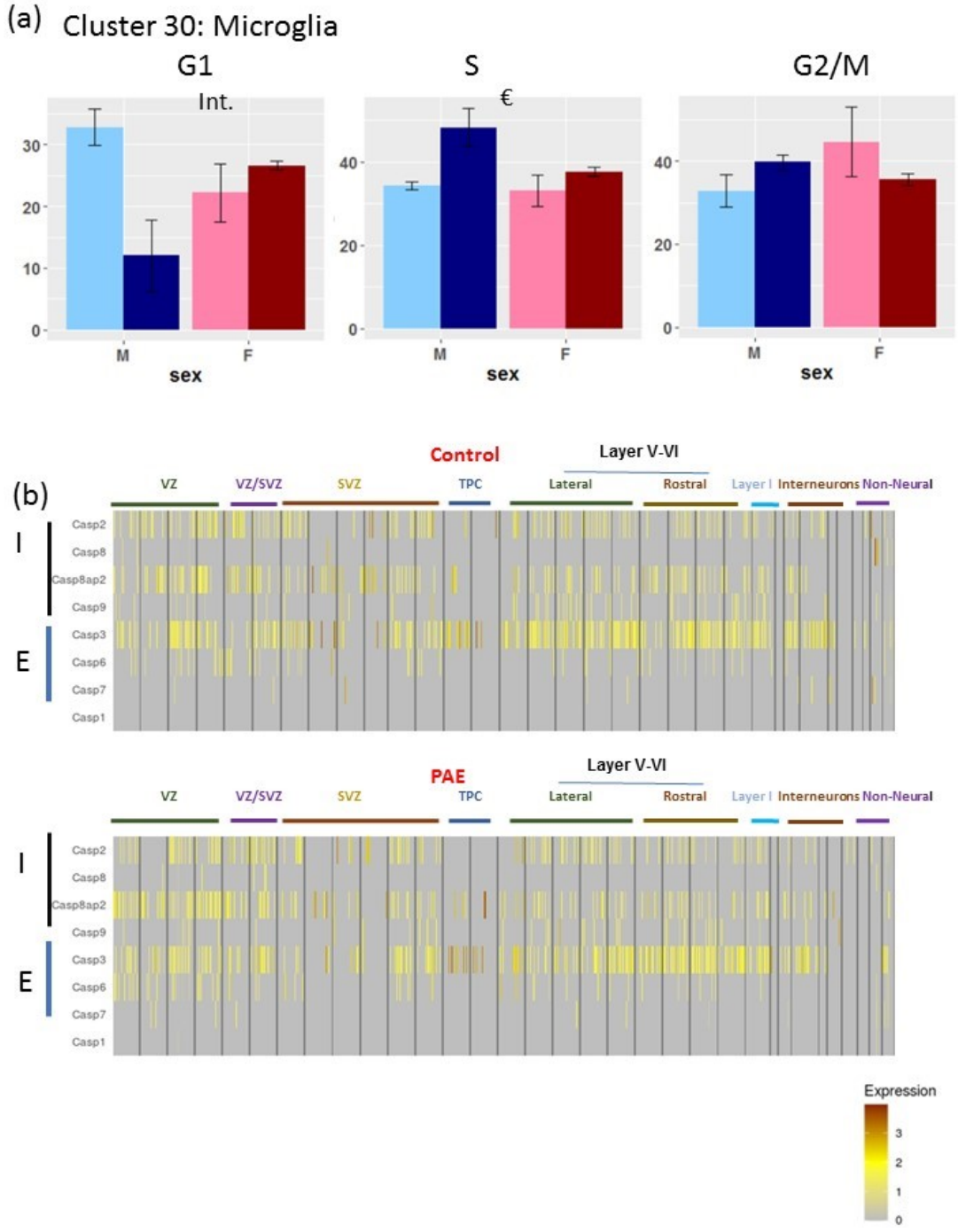
b) Heatmap showing the expression (z-scored) of main subclusters markers in each subcluster.

Appendix 2 b



Ridge plot representing the distribution of pseudotime scores in select clusters (with significant two-way ANOVA main or interaction effect), blue and red color represent male and female

scores respectively, transparent and dark colors represent control and alcohol exposure groups respectively.

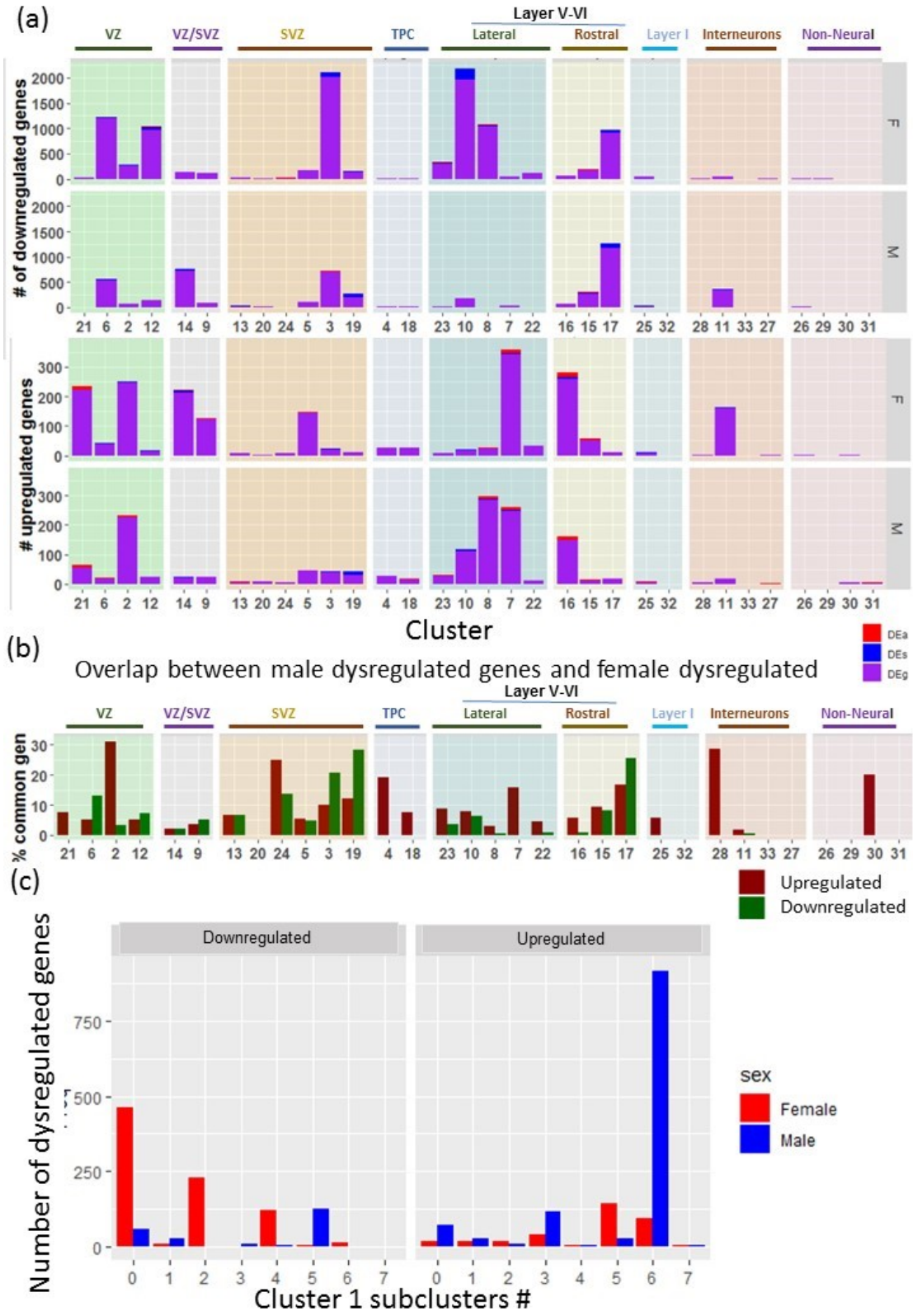


a) Percentage of cells in each of the three cycle phases in microglia cluster (light color) and alcohol (dark color) males (blue) and females (red) clusters. Int. indicates significant interaction effect, € indicates significant main effect of exposure.

b) Heatmap showing expression of cell death markers (Caspase 2, Caspase 3 and Caspase 6) in various cell types, black lines separate clusters, blue lines separate major cell subtypes. Each vertical line (yellow shades) represents an individual cell and the color represents the z score of

the individual cell expression level. TP: Transient progenitor cells, Lat.: Lateral Layer, Rost.: Rostral layer, Intn: Interneurons.

Appendix 2 d

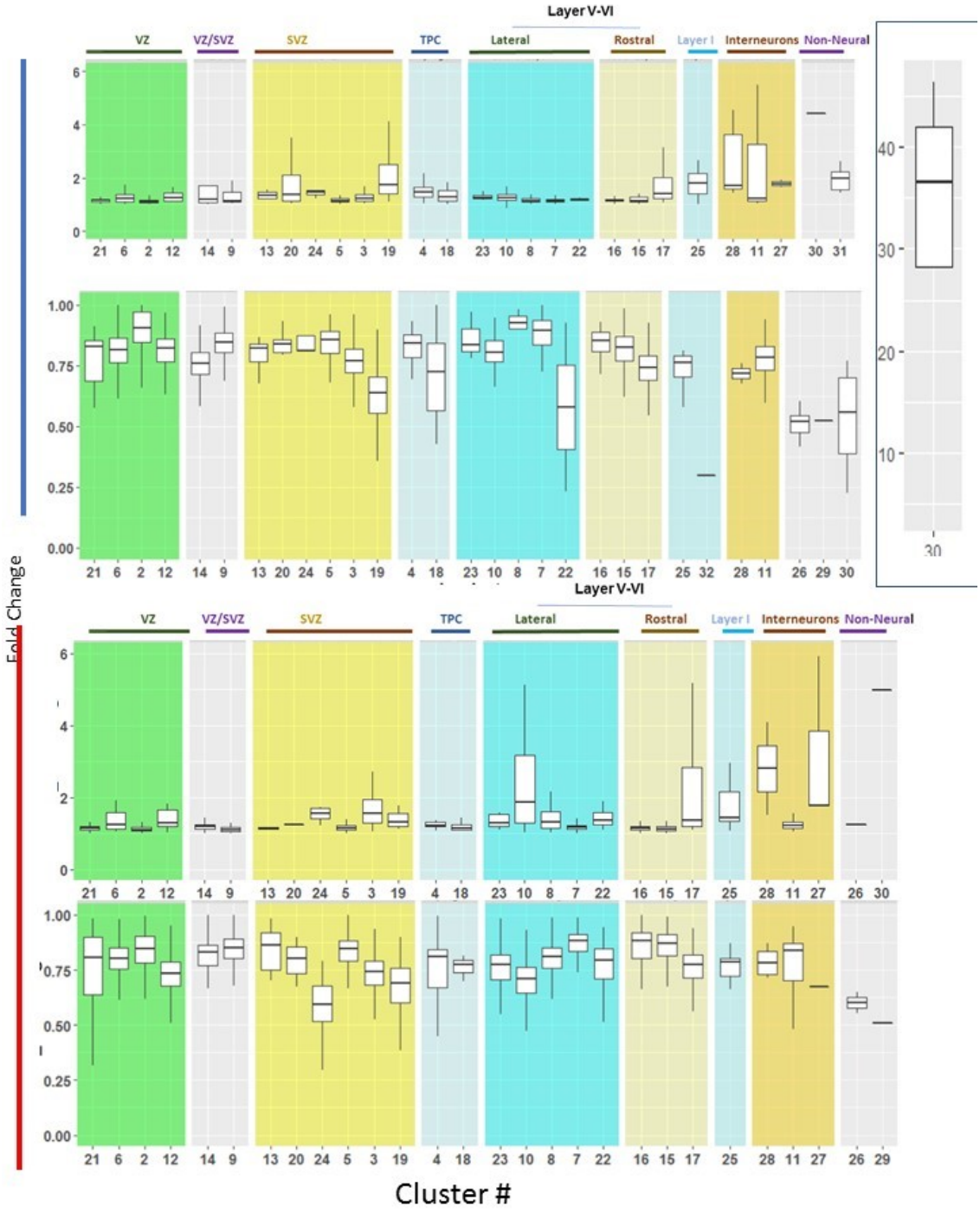

(a)
VZ
VZ/SVZ
SVZ
TPC
Layer V-VI
Lateral
Rostral
Layer I
Interneurons
Non-Neural
# of downregulated genes
upregulated genes
F
M
21 6 2 12 14 9 13 20 24 5 3 19 4 18 23 10 8 7 22 16 15 17 25 32 28 11 33 27 26 29 30 31
Cluster
(b)
Overlap between male dysregulated genes and female dysregulated
DEa
DEs
DEg
% common gen
(c)
Upregulated
Downregulated
Number of dysregulated genes
Cluster 1 subclusters #
sex
Female
Male

a) bar charts showing the number of downregulated (first panel) and upregulated (second panel) genes by alcohol exposure in each female and male cluster, and the breakdown of the dysregulated genes into three dysregulation subtypes.

b) bar chart showing the percentage of common dysregulated genes between males and females in each cluster

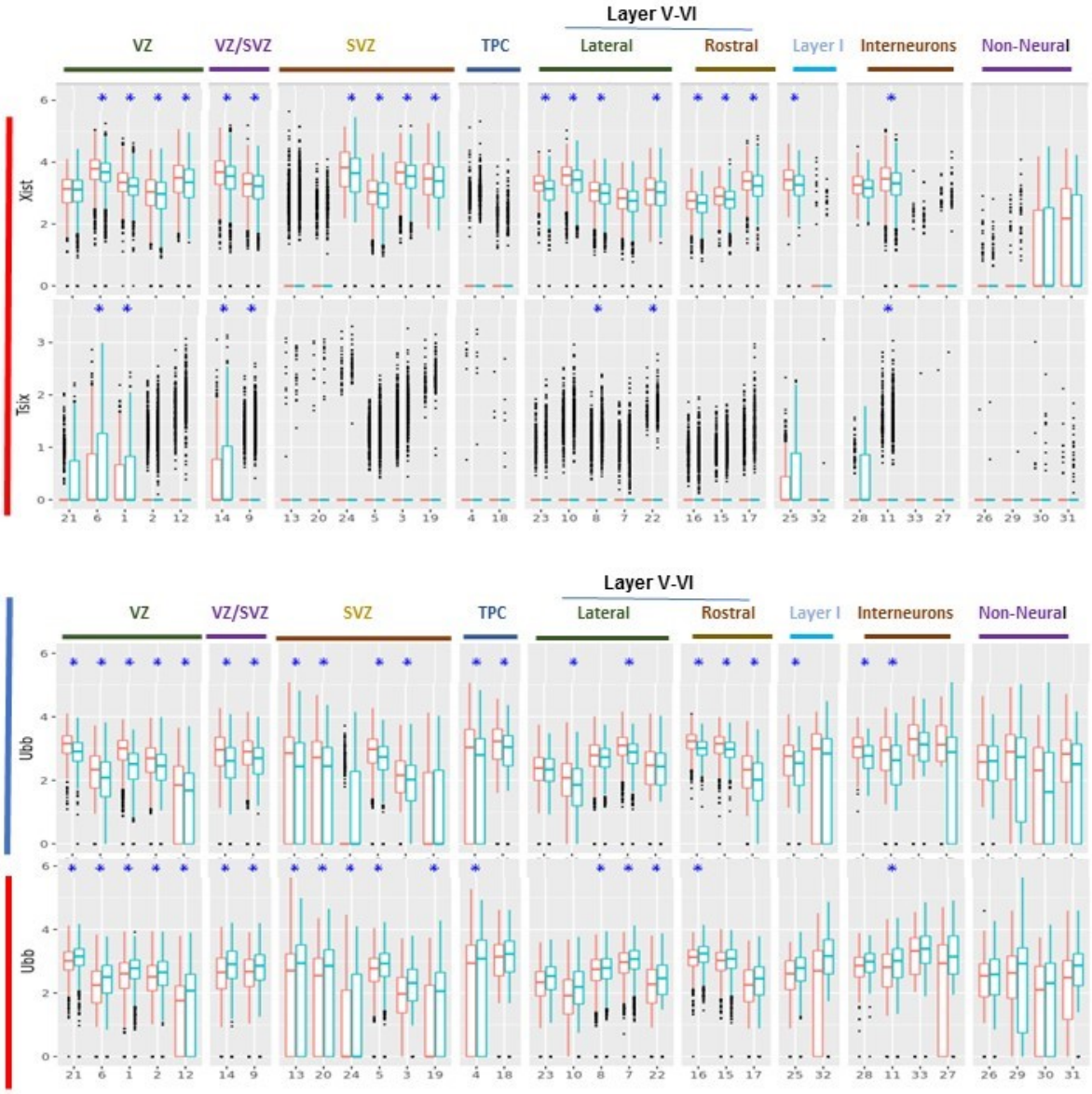
c) bar chart showing the number of downregulated (left panel) and upregulated (right panel) genes in each of male (blue) and female (red) cluster 1 subclusters

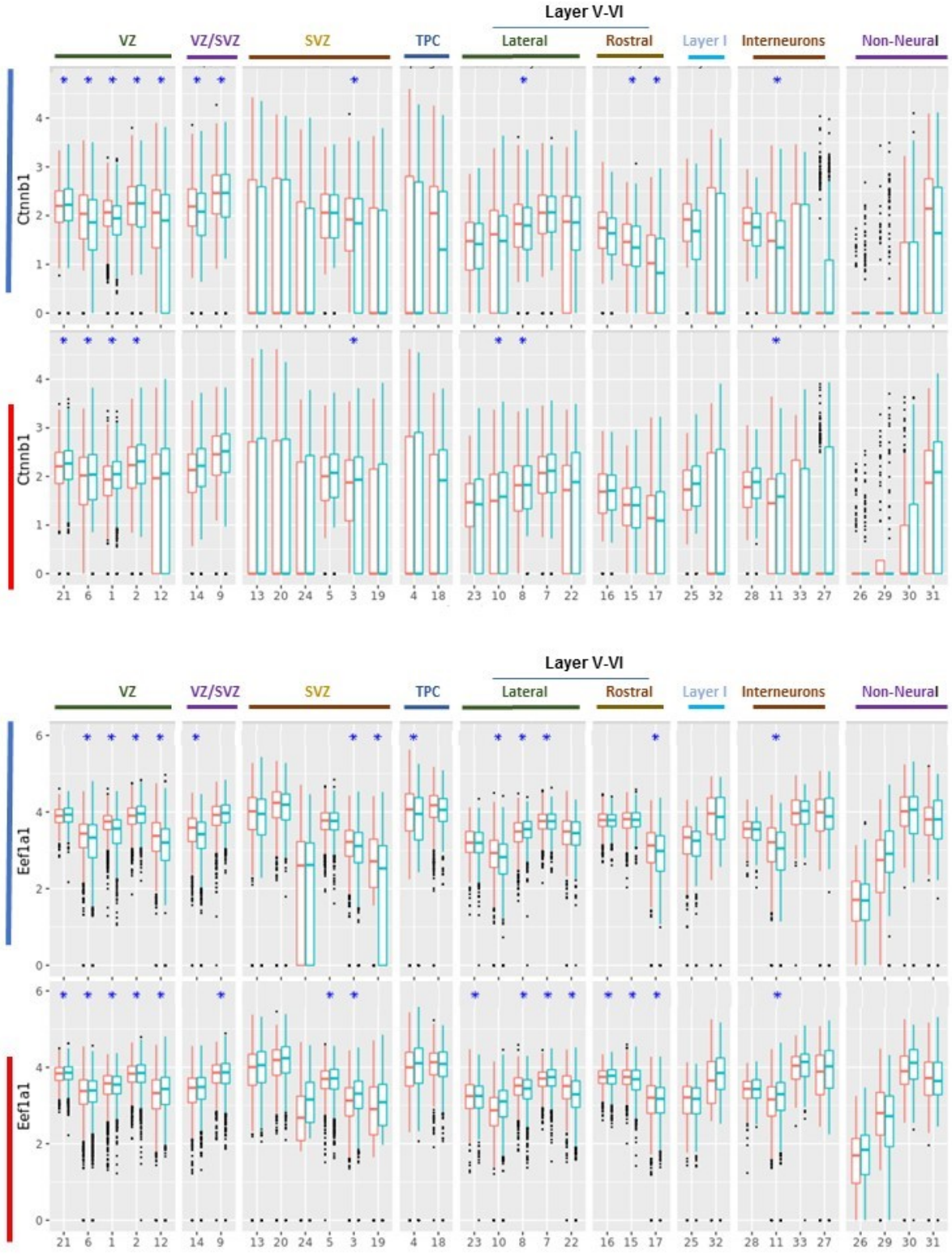
Appendix 2 e

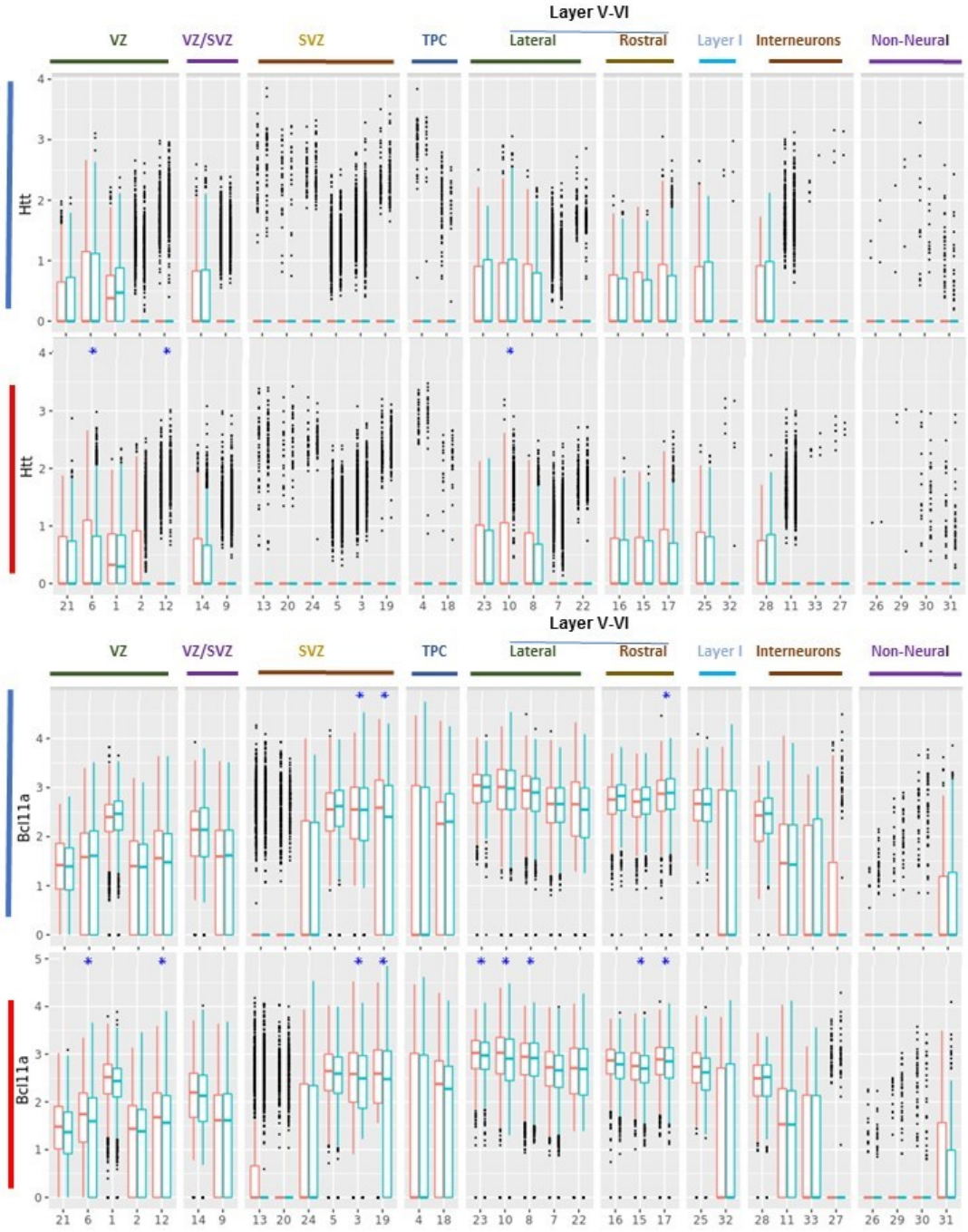


Box plots showing the median and interquartile range of fold change values of upregulated and downregulated genes in males (upper two panels) and females (lower two panels) clusters, side panel (blue box) showing distribution of fold change in male cluster 30.

Appendix 2 f

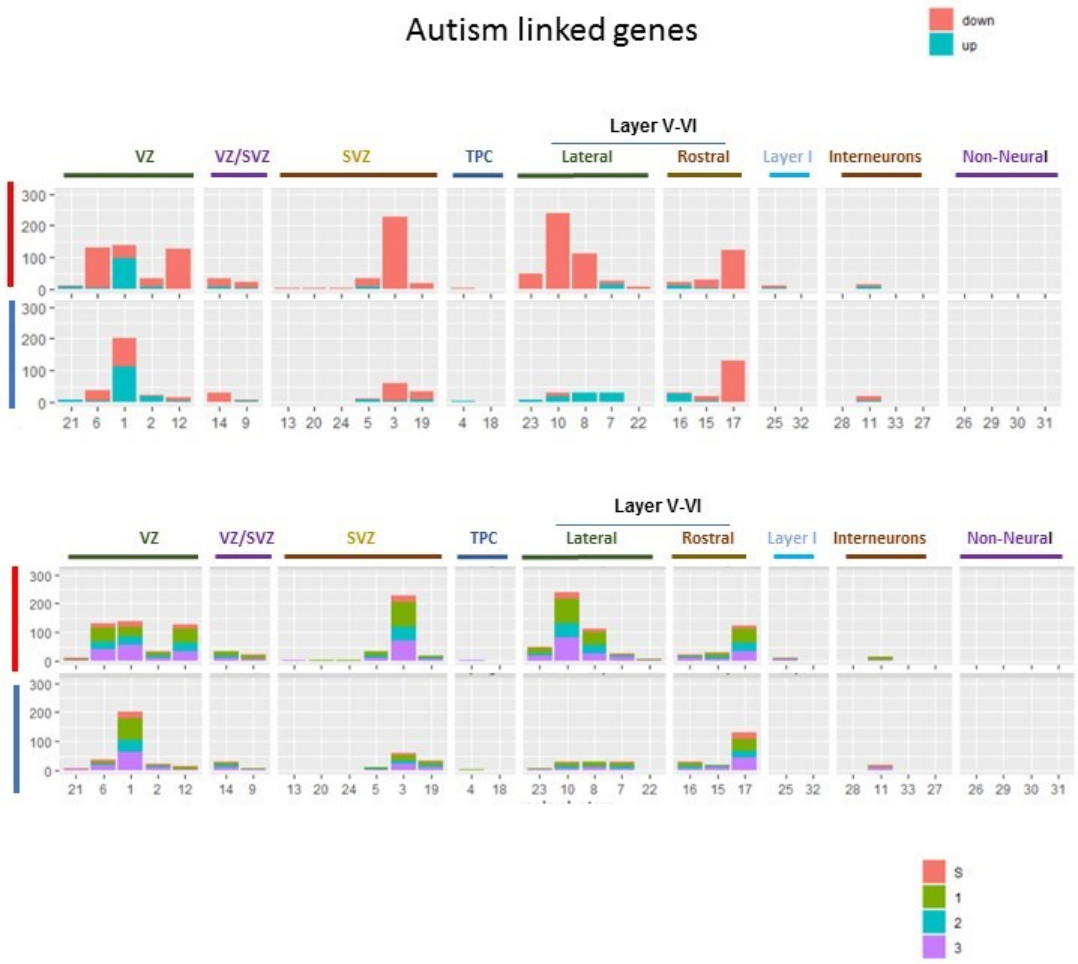

Layer V-VI
VZ
VZ/SVZ
SVZ
TPC
Lateral
Rostral
Layer I
Interneurons
Non-Neural
Xist
Tsix
Ubb
Ubb
21 6 1 2 12 14 9 13 20 24 5 3 19 4 18 23 10 8 7 22 16 15 17 25 32 28 11 33 27 26 29 30 31


Layer V-VI
VZ
VZ/SVZ
SVZ
TPC
Lateral
Rostral
Layer I
Interneurons
Non-Neural
Ctnnb1
Ctnnb1
Eef1a1
Eef1a1
21 6 1 2 12 14 9 13 20 24 5 3 19 4 18 23 10 8 7 22 16 15 17 25 32 28 11 33 27 26 29 30 31



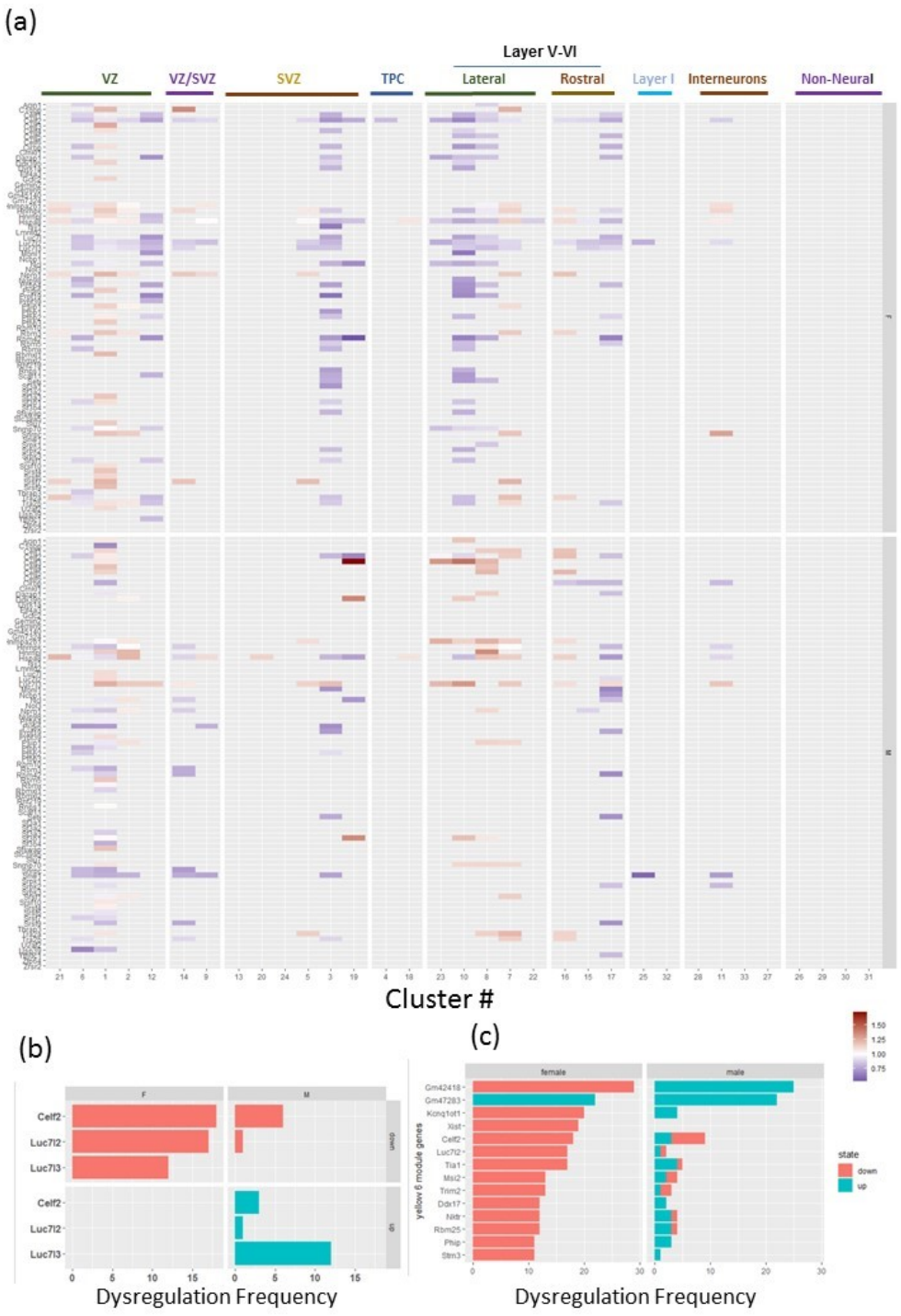
Box plots showing the median and interquartile range of cellular expression level of a group of genes (on y-axis) in the control(orange) and alcohol (green) male (left panel) and female (right panel) clusters, * denotes significant FDR adjusted p-value (<0.05)

Appendix 2 g



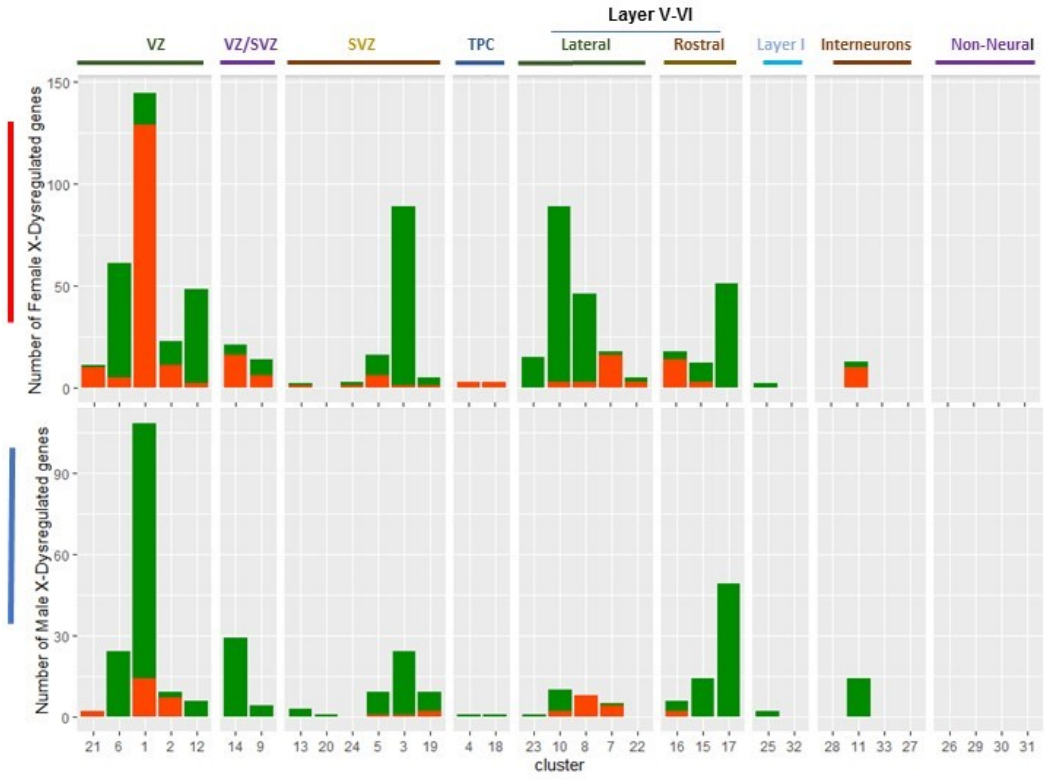
Bar charts presenting the number of up-(green) and down-(orange)regulated autism linked genes in each female (first horizontal panel) and male (second horizontal panel). Breakdown of dysregulated genes by score of strength of evidence is presented in third and fourth horizontal panels

Appendix 2 h



a) Heatmap showing spliceosomal genes (individual vertical lines) regulation by alcohol exposure in female (top panel) and male (bottom panel) clusters. Color legend indicates the fold change of dysregulation.(b) Bar chart showing the number of clusters showing down(top) and up(bottom) regulation of frequently dysregulated (dysregulated in more than 16 clusters) spliceosomal genes (c) Bar chart showing the number of clusters in which each of the cluster 6

yellow modules gene were down-(orange) or up-(green) regulated in females' (left) and males' clusters(right). Shown in figure genes dysregulated in more than 10 clusters.



Bar charts presenting the number of up-(orange) and down-(green)regulated genes located on X-chromosomes in female (first horizontal panel) and male (second horizontal panel) clusters.

Appendix 2 j

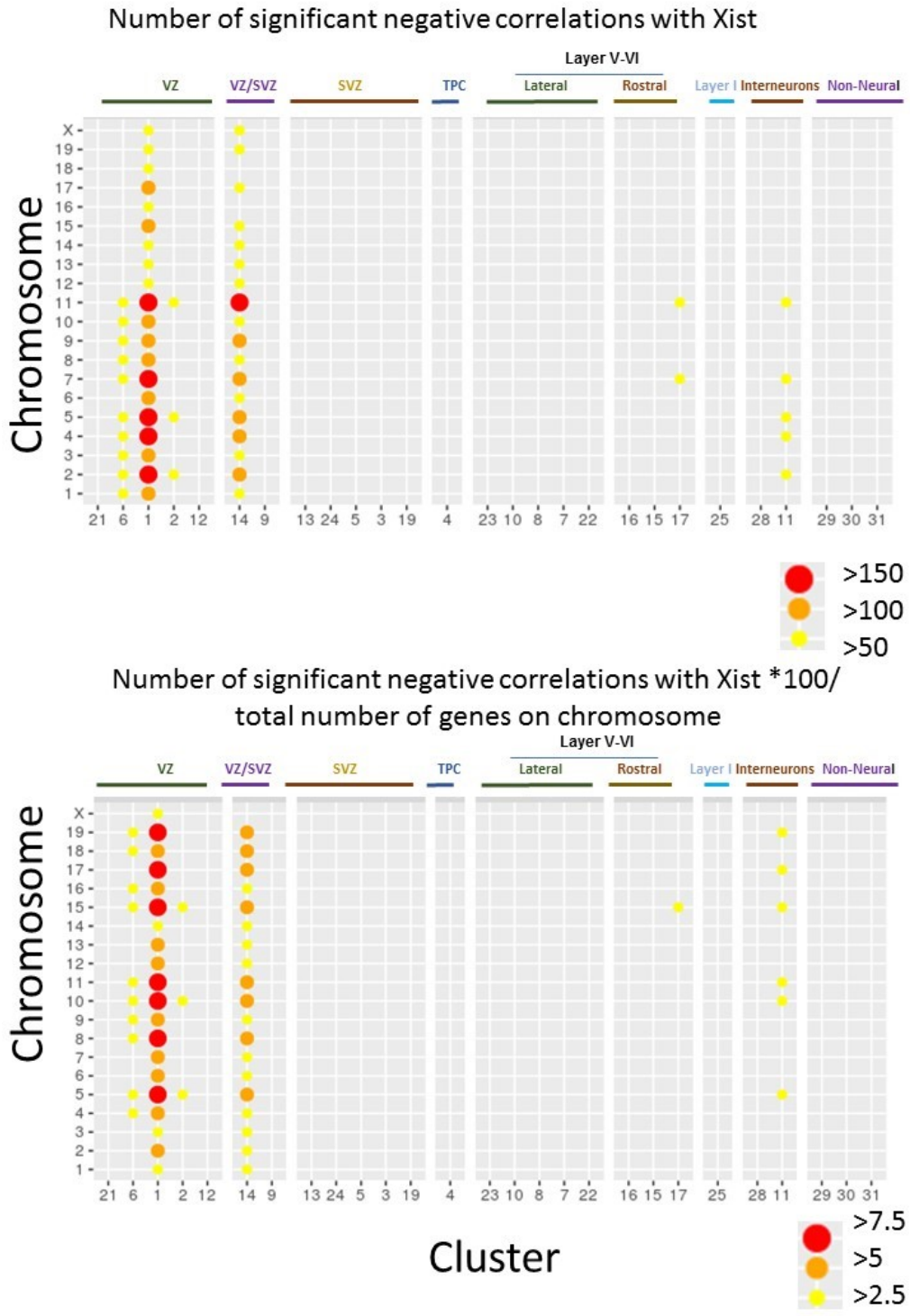

Number of significant negative correlations with Xist
Layer V-VI
VZ
VZ/SVZ
SVZ
TPC
Lateral
Rostral
Layer I
Interneurons
Non-Neural
Chromosome
X
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
21 6 1 2 12
14 9
13 24 5 3 19
4
23 10 8 7 22
16 15 17
25
28 11
29 30 31
>150
>100
>50
Number of significant negative correlations with Xist *100/
total number of genes on chromosome
Cluster
>7.5
>5
>2.5

Figure representing the number (first panel) and normalized to number of each chromosome genes (second panel) significant negative gene expression correlations on each chromosome with Xist in each of the clusters