ASH1L is Necessary for Normal Development of Upper Layer Cortical Neurons

by

Kevin Patrick Toolan

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Genetics and Genomics)
in the University of Michigan
2024

Doctoral Committee:

Professor Sally Camper, Co-Chair
Professor Stephanie Bielas, Co-Chair
Associate Professor Shigeki Iwase
Associate Professor Jacob Kitzman
Research Professor Audrey Seasholtz

Kevin Patrick Toolan

kptoolan@umich.edu

ORCID iD: 0000-0001-7289-569X

Dedication

I dedicate this thesis to Martina. Your unconditional love and support have made this work possible.

I would also like to dedicate this thesis to my parents. You have always supported me, even across the Atlantic Ocean.

Acknowledgements

First and foremost, I want to thank my academic advisors, Sally Camper and Stephanie Bielas. Both of you have provided me with invaluable guidance and advice throughout my PhD career and have always supported me from the very beginning. You have been extremely generous with your time, and never hesitated to provide me with feedback or edits. I have been incredibly lucky to have been co-mentored by such thoughtful and kind scientists.

Members of the Camper lab, both past and present, have been hugely supportive of me. Michelle Brinkmeier has trained and aided me in nearly every experimental technique I have undertaken during my PhD. Words cannot describe how much I appreciate her hard work, attention to detail, and patience. Leonard Cheung has been a fantastic source of knowledge and wisdom and has always provided me with experimental and computational help. Similarly, Amanda Winningham has been incredibly supportive and thoughtful. I have also had guidance and support from Baily Masser, Alex Daly, Maria Ines Perez Millan, Sebastian Vishnapolski, Peter Gergics, Shannon Davis, and Hironori Bando. Lev Prasov, Su Qing Wang, Bill Presley, Gabbi Rozumek, Natalie Michaels and Bridget Blevins have been fantastic lab neighbors and have provided advice and support throughout my time in the Camper lab.

The Bielas lab has been equally supportive of me. Brian McGrath has been so selfless with his time and energy, and I am in awe of his innovation and scientific

knowledge. Amanda Moccia and Liz Werren are fantastic scientists, and both have given me invaluable advice. I will always appreciate their scientific input and feedback. Samantha Regan, Charles Ryan and Emily Peirent have given me incredible guidance and support throughout my many failed experiments and have never hesitated to help me. Similarly, I am eternally grateful for the assistant and feedback Yao-Chang Tsan and Khoa Le have provided me. I have also had guidance and support from Alba Guxholli, Sadie Marlow and Jason Sheingold.

I would also like to thank my thesis committee; Sally Camper, Stephanie Bielas, Audrey Seasholtz, Shigeki Iwase and Jacob Kitzman. You have given me wonderful advice and support throughout my PhD studies. Our meetings have always been a positive experience and have given me confidence as I continue to grow as a scientist. Thank you!

I have been incredibly lucky to be a part of the Department of Human Genetics at the University of Michigan. Tony Antonellis has been an excellent Chair, and prior, a supportive Chair of Education for the department. He has continued to cultivate an environment of collaborative research and mutual respect among staff, students and faculty. The DHG staff have always been supportive and kept the department running smoothly. Special thanks to Molly Martin, Ashley Andrae, Jenny Russell, Dhammika Dewasurendra, Ruth Halsey, Jeff Holden, Tom Sorenson, and Jeff Creech. The DHG faculty and labs have always provided assistance and reagents, for which I am very grateful. A special thank you to Sundeep Kalantry, Clair Harris and Itzaira Mercado-Hernandez for their help regarding Western Blotting experiments, and Sue Hammoud, Mashiat Rabbani and Gabe Manske for a range of reagents and experimental help. I

would also like to thank Ken Kwan for his mentorship and help from the beginning of my graduate studies.

The University of Michigan provides excellent core services, including the Advanced Genomics Core and the Microscopy Core. I would like to thank Eric Rentchler for his time and expertise regarding image analysis. Without Eric's help, I would still be manually counting cells to this day.

I studied my undergraduate degree of Biomedical Sciences at the University of Ulster. I was given the fantastic opportunity to complete an integrated master's degree in conjunction with working in a lab in Indianapolis by Christopher Mitchell, a professor at Ulster. Thank you so much Christopher for orchestrating this connection and allowing me to move to the USA. It truly was a life-changing opportunity. Equally important regarding this opportunity is Tony Firulli, whose lab I had the great fortune of joining, both as a visiting scholar and extended stay as a technician. Thank you, Tony, for your mentorship and inspiration as an incredible scientist. Many thanks to Beth Firulli, Danny Carney, and Rajani George, who all provided me with invaluable support, wisdom, and friendship during my time in the Firulli lab.

Finally, I wish to extend thanks to all my friends and family on both sides of the Atlantic. My parents, Anne and James, and brother, James, have always supported and motivated me in countless ways. Your unending kindness keeps me going. And thank you to my wife, Martina. You are my best friend and greatest cheerleader.

Table of Contents

List of Figures

Abstract

Over 100 genes have been implicated in the genetic etiology of autism spectrum disorder (ASD). Many high confidence ASD genes are involved in chromatin remodeling, histone modifications and DNA methylation. Among these is *ASH1L* (Absent, Small, Homeotic-Like 1), which catalyzes histone H3K36 methylation and has a role in activating transcription by counteracting Polycomb repression. Over 130 autistic individuals have heterozygous loss-of-function *ASH1L* variants, and population studies confirm it is a high-risk autism gene. Several studies report autism-like behaviors in *Ash1l* deficient mice and characterized aspects of the underlying neuropathophysiology. However, the underlying genetic aberrations caused by *Ash1l* loss-of-function are poorly understood in the context of cortical development. We used mice with a cre-inducible deletion of *Ash1l* exon 4, which results in a frame shift and premature stop codon (p.V1693Afs*2). Firstly, we assessed the impact of global *Ash1l* loss-of-function on survival and craniofacial skeletal development. The proportion of homozygous *Ash1l* germline knockout embryos was normal at e18.5, however no live *Ash1l* null pups were present at birth (e18.5: $n = 77$, $P = 0.90$; p0: $Ash1l^{+/+}$ $n = 41$, $P = 0.00095$). $Ash1l^{-/-}$ had shortened nasal bones at e18.5 ($n = 31$, $P = 0.017$) compared to heterozygous and wildtype littermates. To assess *Ash1l* loss-of-function specifically in the cortex, we used a tamoxifen-inducible cre strain to knockout *Ash1l* early in cortical development (*Emx1cre-ER*^{T2}*;* e10.5). We performed immunohistochemistry with

antibodies for SATB2, an upper layer marker, and BCL11B, a layer V marker. We found *Ash1l*^{cKO} had a greater number of SATB2 neurons and they were distributed through the cortical plate (n = 6/genotype, P = 0.0001), whereas BCL11B neurons were similar between genotypes. We performed birthdating of cortical neurons by injected pregnant females with IdU at e14.5 and EdU at e15.5; no differences in neuronal birthdating were identified (n = 4-6/genotype, P = 0.40 for e14.5 and P = 0.057 for e15.5). We utilized single cell RNA sequencing to compare cortical cell populations and identify differentially expressed genes between *Ash1l*^{cKO} and *Ash1l*^{Ctrl} e18.5 embryos. This revealed numerous differences in gene expression that were sufficient to cluster control and mutant upper layer neurons separately. We found that *Ash1l*^{cKO} upper layer neurons had altered pseudotime compared to controls, suggesting underlying differences in cell differentiation trajectory. Together, this analysis reveals that *ASH1L* serves important roles during cortical development, specifically guiding cell fate via gene regulation. Understanding these mechanisms will be fundamental in mitigating *Ash1l*-associated ASD outcomes and aid in treatment development.

Chapter 1: Introduction

Absent, small, or homeotic – like (*ASH1L*), also known as lysine methyltransferase 2H (KMT2H), is a protein-coding gene. The protein product, ASH1L, is a histone methyltransferase (HMTase), which targets residue histone 3 lysine 36 (H3K36) [1]; like other HMTases, it alters chromatin accessibility and regulates gene expression [2]. *ASH1L* is a member of the trithorax group (TrxG) of gene regulators, which counteract polycomb group (PcG) gene repression and regulate expression of many genes. The TrxG-PcG mechanism of gene regulation has been highly conserved throughout metazoan evolution.

ASH1L is a highly conserved gene, and is an ortholog of the *Drosophila* gene *ash1*. ASH1 and ASH1L share greater than 30% protein homology; there is conserved sequence of the Su(var), Enhancer of zeste, Trithorax (SET) domain (66%) and plant homeodomain (PHD) (77%) (**figure 1**) [3]. *Ash1* was initially characterized in 1987 in fly mutagenesis studies; the name *ash1* was derived from the mutant phenotype, in which body structures were absent, small or homeotically transformed [4]. *ASH1L* has come under scrutiny in recent years as a high confidence autism spectrum disorder (ASD) gene. More than one hundred missense, frameshift and nonsense *ASH1L* variants have been found in individuals with ASD, among other neurodevelopmental disorders (NDDs). [5,6]. Several *Ash1l* mouse models have been studied, providing valuable data on how loss of *Ash1l* impacts survival, organ development, gene regulation and

behavior. This chapter aims to summarize previous findings of *ASH1L* function, with specific emphasis on its role in brain development and NDDs, and to highlight outstanding questions in the field.

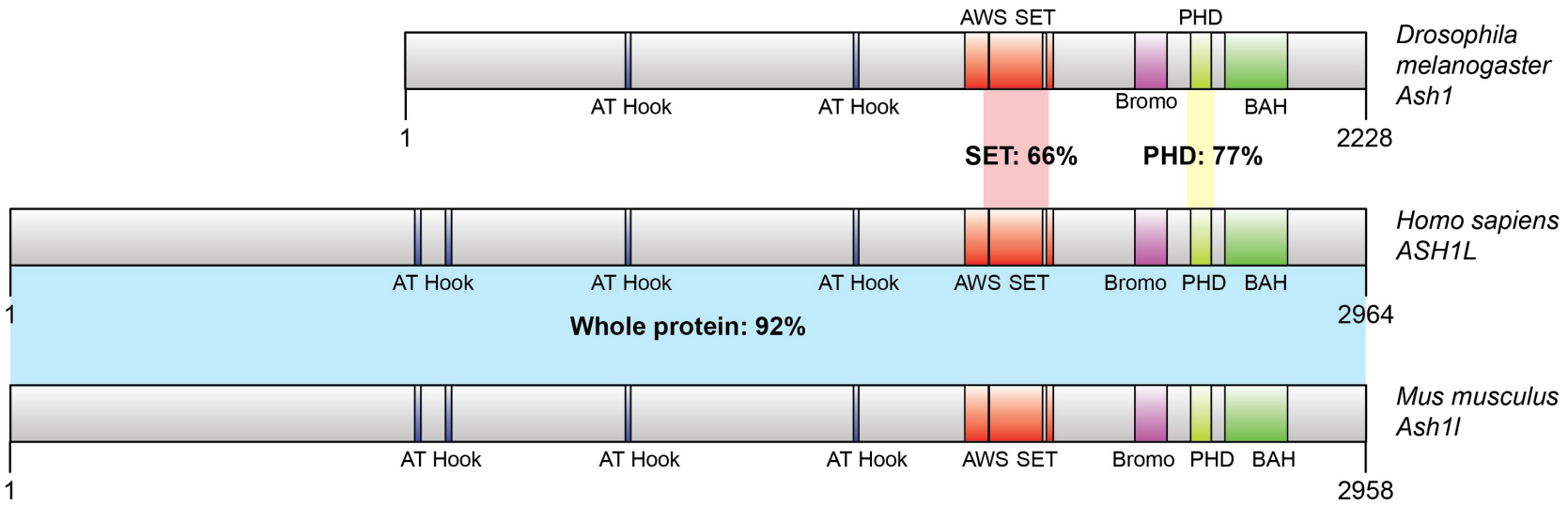


Figure 1: Conservation between Drosophila *Ash1*, and human and mouse ASH1L proteins

Drosophila melanogaster Ash1 protein (top), human ASH1L protein (middle) and mouse ASH1L protein (bottom). There is greater than 30% protein homology between *Ash1* and human ASH1L. The SET domain and PHD (highlighted) have 66% and 77% homology between species, respectively. Human and mouse ASH1L share 92% homology between species.

1.1 Trithorax and Polycomb Gene Regulation

Epigenetic modifications regulate gene expression without altering DNA sequences. They are characterized by being stable and mitotically heritable, allowing for maintenance of active or repressed gene states [7]. Epigenetic mechanisms of gene regulation include DNA methylation, non-coding RNAs and modifying histone tails. To allow for DNA to fit into the nucleus, it must be packaged in a compact manner. The fundamental unit of this packaging is the nucleosome: 147 base pairs (bp) of DNA wrapped around an octamer of core histone proteins, H2A, H2B, H3 and H4 (each with two copies) [8]. These histone proteins have 'tails' – amino acid sequences that extend from the nucleosome core. Histone tails undergo various post-translational modifications, including acetylation and methylation, at specific residues. In turn, these modifications alter DNA accessibility of a nucleosome and how nucleosomes interact with one another [9].

The TrxG and PcG are examples of proteins that modify histone tails. The discovery and study of TrxG and PcG has a relatively long history, and it is intertwined with that of *Hox* genes. *Hox* genes are a conserved family of transcription factors that specify body segmentation. The *Hox* genes, PcG, and TrxG were discovered first in *Drosophila melanogaster* in the 1940s, when homozygous *Drosophila* mutants were found to have extra sex combs. Various other mutant flies arose in subsequent years with altered body structures. In 1978, Lewis described homozygous *polycomb* mutant larvae, with additional abdominal segments, and proposed the idea of a global repressor of bithorax complex genes, a subset of the *Hox* gene family [10,11]. From this sprang the concept of PcG, and soon, TrxG; genes that regulated *Hox* gene expression.

It was soon found that TrxG-PcG genes had orthologs in mammals [12]. *Hox* genes are not the only target for TrxG-PcG gene regulation; over one hundred genes are regulated by the dynamic balance between TrxG-associated activation and PcG-associated repression [7].

Active and repressive states depend on the various histone marks deposited or removed by TrxG and PcG. These proteins can be divided into groups based on biochemical function. Lysine methyltransferases (KMTs; also called histone methyltransferases, HMTs) catalyze the mono-, di- and trimethylation of specific residues along a histone tail. Depending on the residue, this can have an active or repressive effect on gene expression; for example, H3K4me is an active mark, whereas H3K27me3 is a repressive mark. Studies have identified genomic regions of bivalent occupancy – silent domains with both H3K27me3 and H3K4me3. Bivalency reflects a state poised for transcription, which is particularly crucial during neurodevelopment, as these regions may quickly switch to an active state during the course of differentiation [13].

Lysine demethyltransferase (KDMs) remove methyl groups from histone tails. The mammalian gene JARID1d is part of the TrxG, which demethylates H3K4 [14]. Polycomb repressive complex 1 (PRC1) is a member of PcG and is an E3 ubiquitin ligase. Its subunits RING1A/B catalyze monoubiquitination of H2A, mainly lysine 119, resulting in H2AK119ub1 [15]. Ubiquitination is understood to restrict RNA polymerase II (RNAPII) elongation, as well as recruit polycomb repressive complex 2 (PRC2) members. The Polycomb repressive deubiquitinase complex (PR-DUB) removes H2A ubiquitin marks. Interestingly, both ubiquitination and deubiquitination are required for

PcG repression [16]. PR-DUB maintains the level of H2AK119ub1 deposited by PRC1, as excessive ubiquitination counteracts chromatin compaction [17].

Not all TrxG members modify histone tails; ATP-dependent chromatin remodelers can slide or evict nucleosomes, altering gene expression. These complexes can be subdivided into four distinct groups: switch/sucrose non-fermentable (SWI/SNF), imitation switch (ISWI), chromodomain helicase DNA-binding/nucleosome remodeling deacetylase (CHD/NuRD) and inositol auxotroph 80 (INO80/SWR) [13].

Ash1 and *Ash1l* are methylating members of the gene activating TrxG. However, there appears to be a lack of consensus regarding which histone residues they target. It was initially believed that *Drosophila Ash1* was the only methyltransferase capable of methylating H3K4 [18]. However, further studies found other H3K4 methyltransferases in *Drosophila*, such as *dSet1* [19]. Using a histone lysine scanning mutation assay, Tanaka *et al.* screened for targets of histone methyltransferases. Not only did they find that mammalian *ASH1L* specifically methylated H3K36, they repeated their experiment with *Drosophila Ash1* and got the same result [20]. Yuan *et al.* also found that recombinant human *ASH1L* SET domain, purified from *E. coli,* specifically targeted H3K36 [21]. Interestingly, they also demonstrated that H3K36me2/3 inhibited PRC2-mediated spread of H3K27me3, providing the field with a new model of *ASH1L*-associated gene activation.

Several approaches have been taken to assess the effect of ASH1L deficiency on histone modification. Qin *et al.* used short hairpin RNA (shRNA) to knockdown *Ash1l* expression in the prefrontal cortex (PFC) of young mice. They showed significant downregulation of *Ash1l* mRNA and *ASH1L* protein in these mice compared to controls,

which were injected with scrambled shRNA. They observed roughly 50% decrease of H3K4me3 in PFC of *Ash1l* shRNA treated mice compared to controls, whereas no difference in H3K36me2 or H3K36me3 was observed [22]. Miyazaki *et al*. generated an in-frame deletion of *Ash1l*'s SET domain in mouse embryonic stem cells (ESCs), which disrupted the methyltransferase activity of *Ash1l,* but retained its other functional domains. They performed chromatin immunoprecipitation (ChIP) assays and observed significant reduction of both H3K36me2 and H3K36me3, whereas no difference in H3K4me3 levels was observed [23]. In *Ash1l*$^{+/-}$ and control mice, Yan *et al.* investigated *Ash1l*-associated histone methylation across various regions of *EphA7*, which encodes a cell surface protein important for synaptic function and neuronal plasticity. The authors found *EphA7* was downregulated in *Ash1l*$^{+/-}$ compared to control mice via RNA-seq. Further analysis with auditory fear conditioning found that *Ash1l* regulates *EphA7* expression in an activity-dependent manner. To determine if *Ash1l* directly or indirectly regulates *EphA7,* the authors used ChIP to quantify levels of H3K4, H3K36 and H3K27 methylation in *Ash1l*$^{+/-}$ and control mice before and after an auditory-based learning task. Surprisingly, they found that both the H3K4 and K3K36 methylation levels at *EphA7* were unchanged between *Ash1l*$^{+/-}$ and control mice. However, there was an increase of the PcG-mediated silencing H3K27 methylation along the promotor and exons of *EphA7* in *Ash1l*$^{+/-}$ compared to controls. Pulldown with anti-EZH2, the catalytic subunit of PRC2, revealed increased occupancy at the promotor region in heterozygotes. The authors argue that it is likely an indirect mechanism in which *Ash1l* antagonizes PRC2 silencing that allows for *EphA7* expression in an activity dependent manner [24]. Together, these data demonstrate the complexity of deciphering targets of

ASH1L-mediated methylation. These differing results may indicate that *ASH1L* methylation is context dependent; perhaps it targets different residues during different biological processes, such as development and activity-based learning. This theory could be tested by ChIP across various points of development, adulthood, and before and after learning activities. It is also likely that ASH1L function is influenced by interaction with other proteins.

Many mammalian TrxG and PcG proteins function as complexes. PRC1 and PRC2 are composed of core subunits, however there are additional proteins that may be added to alter targets or functionality. PRC1, which catalyzes H2AK119ub1, is composed of RING1A/B and PCGF1–PCGF6. Canonical PRC1 includes CBX proteins, which bind to H3K27me3, whereas non-canonical PRC1 includes KDM2B, a H3K36 demethylase. Interestingly, with the addition of AUTS2, CK2 and CBP/P300, PRC1 functions as a transcriptional activator, rather than silencing gene expression [25]. PRC2 is a methyltransferase which catalyzes H3K27me3, and its core is composed of EZH1/2, SUZ12, EED164 and RBBP4/7. However, its accessory proteins include H3K36me3-binding PCL1–PCL3, and H2Aub-binding JARID2 [26]. The deubiquitinase complex, PR-DUB, is comprised of BAP1, HCFC1, FOXK1/2, and OGT, as well as ASXL1/2/3 proteins [27]. TrxG complexes include COMPASS and COMPASS-like. COMPASS and COMPASS-like have overlapping subunits, including ASH2L, DPY30 and HCF1; however, COMPASS also includes SET1A (which methylates H3K4) and CXXC1 (which binds to CpG). COMPASS-like, on the other hand, includes Mixed Lineage Leukemia (MLL)3, which methylates H3K4 at hormone responsive genes, and UTX (which demethylates H3K27) [28]. ASH1L has been shown to interact with MRG15

(MORF-related gene on chromosome 15) by several independent studies. MRG15 binds to a F-Q-L-P motif in ASH1L, and upon binding, fully activates ASH1L's methyltransferase activity [29–31].

Many mammalian TrxG and PcG proteins are necessary for normal neural proliferation and differentiation. In mouse ESCs, bivalent chromatin is found near transcription start sites of genes essential for neuronal proliferation and differentiation (such as *Sox2* and *Fgf8*, respectively) [32]. PRC1 catalytic subunit *RING1A/B* and PRC2 catalytic subunit *EZH2* have been shown to regulate the fate switch of radial glia from neuron to astrocyte via neural progenitor cells (NPC) knockout experiments [33]. Depending on the combination of accessory proteins added to PRC1, it exhibits distinct functions. For example, *Bmi1* is necessary for self-renewal of neural stem cells by repressing $p16^{Ink4a}$, a cyclin-dependent kinase inhibitor [34]. PRC2 catalytic component *EZH2* is highly expressed in neural stem cells. Conditional knockout of *Ezh2* with *Emx1-cre* results in a shorter neurogenic period in cortical development, resulting in premature fate change to astrocytes [35]. Furthermore, conditional knockout of *Ezh2* with *Wnt1-cre* (a midbrain-specific cre) results in de-repression of *Foxg1* and *Pax6,* which are determinants of forebrain development. This results in a homeotic-like transformation from midbrain to forebrain [36]. TrxG member *Chd7* activates expression of *Sox4* and *Sox11,* which are necessary for neuronal differentiation [37].

The importance of TrxG and PcG proteins in neurodevelopment is underscored by variants found in these genes in individuals with NDDs. Many TrxG and PcG variants result in syndromic disorders, affecting multiple organ systems, including the brain. Weaver syndrome was initially described in two patients [38]; it is a rare disorder

characterized by macrocephaly, intellectual disability, and craniofacial and skeletal anomalies [39]. PRC2 loss of function is believed to be at the heart of Weaver syndrome pathophysiology, with variants reported in subunits *EZH2, SUZ12* and *EED*, and loss of H3K27me3 results in attenuated gene silencing. While there are reports of *SUZ12* and *EED* variants linked to Weaver syndrome [40,41], *EZH2* has come under the most scrutiny. Several heterozygous variants in *EZH2* have been linked to Weaver syndrome [42,43]. Two *EZH2* variants found in individuals with Weaver syndrome (p.V626M and p.A682T) were transfected into mouse chondrocytes lacking endogenous *EZH1/2,* and H3K27me3 activity was measured. Both variants resulted in decreased H3K27me3 compared to wildtype, suggesting that *EZH2* partial loss of function may be implicated in Weaver syndrome [44].

Bainbridge-Ropers syndrome is a rare neurodevelopmental disorder, which has phenotypic overlap with Bohring–Opitz syndrome. Both syndromes are associated with microcephaly, intellectual disability and developmental delay [45,46]. PR-DUB subunits BAP1 and ASXL1/2/3 are implicated in these syndromes, with loss of function variants of these proteins found in patients [47]. *BAP1* is the catalytic component of PR-DUB, and loss of *BAP1* function results in an unregulated increase of H2AK119ub1 at intergenic regions in mouse ESCs. This in turn displaces H3K27me3 from its polycomb target genes and causes its spread through intergenic regions, allowing de-repression of target genes and condensation of off-target chromatin [48].

Rubinstein–Taybi syndrome is another neurodevelopmental disorder, characterized by microcephaly, small corpus callosum, intellectual disability and distinct facial features, with one third of patients also affected by congenital heart defects [49]. A

non-canonical form of PRC1, that actives gene expression with subunits AUTS2, CK2 and CBP/P300, is associated with Rubinstein-Taybi syndrome; patient sequencing has found heterozygous variants in *AUTS2*, *CBP* and *P300* with the syndrome [50]. CBP/P300 are histone acetyltransferases that, as part of this non-canonical PRC1, acetylate heat-shock factor 2 (HSF2), which stabilizes the protein [51]. HSF2 has been shown to be essential in cortical development, in which *Hsf2* null mice have reduced radial glia fibers that in turn disrupt neural migration [52].

Variants in TrxG genes have also been associated with NDDs. *MLL1 - 4* are all associated with distinct disorders. Wiedemann-Steiner syndrome is characterized by developmental delay, intellectual disability, and ASD. Extensive clinical and genetic analysis has found variants in *MLL1* are strongly associated with Wiedemann-Steiner syndrome [53–55]. *MLL2* variants have been found in individuals diagnosed with Kabuki syndrome; features include developmental delay, ASD, as well as extensive craniofacial anomalies [56].

Coffin-Siris syndrome is associated with variants found in a number of TrxG ATP-dependent chromatin remodelers, including *ARID1A/B* and *SMARCB1* [57]. Coffin-Siris syndrome features include microcephaly, growth deficiency and intellectual disability [58]. CHARGE syndrome (Coloboma, Heart defects, choanal Atresia, Retardation of growth, Genitourinary malformation and Ear abnormalities) is linked to heterozygous variants of *CHD7*; individuals with CHARGE may be impacted by a range of clinical phenotypes described in the name [59]. Conditional knockout of *Chd7* with *Atoh1-cre* (cerebellum granule cell specific) found that *Chd7* regulates neuronal differentiation, and loss of function results in cell death [60].

Many *ASH1L* variants have been found in individuals with NDDs. This will be discussed further in the section entitled **Clinical reports and population studies of *ASH1L*.** Overall, these clinical reports stress the importance of PcG and TrxG genes in neurodevelopment, and how variants in these genes can result in severe, multisystem disorders.

ASH1L is a unique trithorax group member as it is the only prominent member identified that methylates H3K36. However, H3K36 methyltransferases exist outside the trithorax classification, and many share the same functional domains as *ASH1L,* mainly the SET and associated domains. H3K36 methylation is an active mark, however H3K36me2 and H3K36me3 mark chromatin differently: H3K36me2 is enriched at the transcriptional start sites of active genes, which gradually decreases and is replaced by H3K36me3 towards the 3' end [61].

In yeast, H3K36 methylation is only carried out by SET2, whereas in mammals there are six H3K36 methyltransferases. These include *SETD2, NSD* (nuclear receptor-binding SET domain) *1, 2* and *3*, and *MES-4*, in addition to *ASH1L* [62]. These proteins likely serve distinct biological functions, as they have differing genomic targets and methylation signatures. In a mouse fibroblast cell line, *Setd2* knockdown was performed using siRNA; global H3K36me3 was almost completely depleted, whereas H3K36me1/2 was largely unaffected [63]. Knockdown of *Nsd1* in ESCs via shRNA resulted in marked decrease in genome-wide H3K36me2, with a consequential increase in PRC2-mediated H3K27me3 [64].

H3K36 methyltransferases have essential roles in neurodevelopment, as several have been implicated in NDDs. Sotos syndrome is caused by variants in *NSD1* [65,66].

The features of individuals with Sotos syndrome are childhood overgrowth, intellectual disability, and ASD. *Nsd1*$^{+/-}$ mice, with a CRISPR-mediated deletion of *Nsd1* exon 3, did not exhibit morphological differences. They did, however, exhibit social behavior defects, in line with the ASD clinical phenotype found in Sotos syndrome individuals [67]. Variants in *NSD2* have been linked to Wolf–Hirschhorn syndrome; features include microcephaly, developmental delay and facial anomalies [68].

Variants in the trimethyltransferase *SETD2* have been found in individuals with NDD [69–72]. Most of these individuals exhibit an overgrowth phenotype, including macrocephaly, and behavioral anomalies including ASD and ID [69]. A clinical review has determined individuals with *SETD2* truncating variants exhibit Luscan-Lumish syndrome, a Sotos syndrome-like disorder which also exhibits obesity. These loss-of-function variants confer hypomethylation, whereas recurrent missense variants (*SETD2* p.Arg1740Trp or p.Arg1740Gln) are found to result in hypermethylation. Interestingly, individuals with these missense variants exhibited developmental delay, feeding difficulties and microcephaly [70,72].

ASH1L is a TrxG regulator that methylates H3K36, in turn counteracting PcG repression. TrxG, PcG and other H3K36 methyltransferases are heavily involved in neurodevelopment, and variants in these genes identified in individuals with an NDD highlight their essential roles. These NDDs have distinctive clinical features, such as macrocephaly or microcephaly, and in some cases, overlapping phenotypes, such as ASD and ID. This further adds to the complexity of deciphering the roles of what these epigenetic regulators perform in neurodevelopment. *ASH1L* variants have been found in

individuals with a range of clinical phenotypes, however there is clear convergence of several disorders, including ASD.

1.2 ASH1L Protein

Human *ASH1L* was first identified by comparing human expressed sequence tags (EST) with *Drosophila ash1* SET domain [3]. It is located on chromosome 1q22 and spans 227.5 kb. *ASH1L* is expressed ubiquitously in the human body, with highest expression in the ovaries, thyroid and brain [73]. The ASH1L protein is 2964 amino acids and contains several functional domains [74]. These include four AT hooks, the Su(var)3-9, Enhancer-of-zeste and Trithorax (SET) domain, Associated With SET (AWS) domain and Post-SET domain. Additional domains include a bromodomain, plant homeodomain (PHD) and bromo-adjacent homology (BAH). A schematic of the *ASH1L* protein is shown in **figure 2**.

AT hooks are a highly conserved domain, comprising of a nine amino acid motif that facilitate DNA binding [75]. AT hooks are found in various chromatin remodelers [76]. While no functional studies have been carried out on *ASH1L, Drosophila Ash1* AT hooks have been assessed for necessity. *Drosophila Ash1* contains three AT hooks, and mutating two or all three AT hooks reduced chromatin binding by almost 70%. In combination of mutating the three AT hooks and mutating the BAH domain in fly, the chromatin binding of *ASH1* was reduced by greater than 90%. Interestingly, this reduction in binding was specific to metaphase, and not in interphase. Furthermore, in a series of rescue experiments with the *ash1* null allele, *ASH1* with mutated BAH partially rescued lethality, and additional mutations in even one AT hook severely impacted

survival [77]. It will be interesting to assess the necessity of AT hooks in *ASH1L* given these findings in *Drosophila Ash1.*

The AWS, SET and post-SET domains are responsible for *ASH1L*'s methyltransferase activity, which has been shown to mono- and dimethylate H3K36 [78]. H3K36 methylation is associated with relaxing chromatin and allowing expression of genes. *ASH1L, NSD1, NSD2, NSD3* and *SETD2* are the only human genes that methylate H3K36, all with AWS, SET and post-SET domains [79]. There have been several studies on the structure of *ASH1L*'s SET and related domains; interestingly, they all indicate mechanisms of self-regulation. The core SET domain is the catalytic component, but also harbors a variable SET-I subdomain within its boundaries. The SET-I and the post-SET domains form loops that surround the SET active site; these undergo conformational changes [80]. MRG15, a chromatin organizer protein, binds to an F-Q-L-P motif in ASH1L [81]. In turn, this interaction increases the disorder of the post-SET domain loop (referred to as the autoinhibitory loop, or AI loop), thus releasing it and activating ASH1L methyltransferase activity [29,31]. Further research showed that this activation is nucleosome dependent, and that MRG15 facilitates ASH1L methyltransferase activity by recruiting it to nucleosomes [30]. The AWS is composed of two α-helices and several loops, which are bound by zinc atoms and coordinated with eight cysteines [82]. X-ray crystallography revealed that the AWS has high flexibility, which allows for more mobility [80]. It is not known whether this also helps regulate *ASH1L* methyltransferase activity.

Miyazaki *et al.* set out to uncover how *Ash1l* establishes gene expression in the context of its methyltransferase activity. First, the authors generated ESCs with a knock-

in allele of *Ash1l* harboring an in-frame deletion in the SET domain, rendering the methyltransferase ability inactive (ΔSET ESCs). Using retinoic acid (RA) to induce *Hox* gene expression, the authors found attenuated activation of *Hoxb4* and *Hoxd4* in ΔSET ESCs. RNA-seq of RA treated ΔSET ESCs and WT cells identified 152 genes that had impaired response to RA in ΔSET ESCs, including *Hox* genes. They also found these genes were enriched for the PRC2-mediated H3K27me3, suggesting that *Ash1l*'s methyltransferase activity counteracts PRC2 at these genes.

Various glutathione S-transferase (GST) fusion *ASH1L* proteins were generated to determine enzymatic activity of the SET domain; these included the entire wildtype (WT) SET domain, SET domain with an amino acid substitution (H2213K) and a recombinant SET domain lacking the N terminal domain of the AWS region. WT SET domain methylated H3 of nucleosomes, whereas the substitution and N terminal SET domain deletion did not. Interestingly, the WT SET domain could methylate H3 of nucleosomes, but not free histone octamers, or H3-H4 tetramers or H2A-H2B dimers bound with DNA, suggesting the full nucleosome unit is needed. Furthermore, Miyazaki *et al.* performed ChIP experiments to elucidate the histone residue targets of *Ash1l* by assessing regions of *Hoxd4* gene before and after addition of RA. The authors found WT, but not ΔSET ESCs, catalyzed di- and trimethylation of H3K36 post RA treatment, but no difference in H3K4 methylation signatures, suggesting *Ash1l* specifically performs H3K36me.

The authors wanted to determine if transcriptional elongation at active genes is necessary to alter histone methylation. To this end, the authors used 5,6-dichloro-1-D-ribofuranosyl-benzimidazole (DRB). DRB inhibits the kinase activity of P-TEFb, which in

turn phosphorylates the Ser2 of RNAPII. This phosphorylation releases RNAPII and transcriptional elongation takes place; when inhibited, RNAPII remains in the poised position. The authors found that addition of DRB to control ESCs still resulted in increased levels of di- and trimethylation at H3K36 compared to ΔSET ESCs, suggesting methylation occurs independently of elongation. Interestingly, H3K36me2 occurs genome-wide, whereas H3K36me3 is gene dependent. The authors sought to determine if *Ash1l*-mediated methylation was excluding PcG components from repressing gene expression, or if the transcriptional elongation itself resulted in this exclusion. The authors conducted ChIP assays of *Hoxd4* chromatin, pulling down MEL18, RNF2 (both components of PRC1) and SUZ12 (component of PRC2). MEL18 and RNF2 were increased in ΔSET ESCs upon RA application, whereas SUZ12 was decreased in both WT and ΔSET ESCs. This data suggests *Ash1l* methyltransferase activity promotes exclusion of PRC1 components *Mel18* and *Rnf2* independently of transcriptional elongation. The fact that the PRC2 component SUZ12 was excluded regardless of H3K36me suggests either that exclusion is *Ash1l*-independent or that the methylation-independent functions of ASH1L are sufficient for exclusion [23].

ASH1L's C-terminal region contains three chromatin-interacting domains: the bromodomain, PHD and BAH domain. Bromodomains form helical structures that recognize acetylated lysine [83]. They are found in several histone methyltransferases, including TrxG members *MLL*, and chromatin remodeling complexes, SWI/SNF [84]. PHDs are a diverse family of histone mark readers. They are zinc binding fingers 50-80 amino acids in length and can recognize a range of marks, including H3K4me2/3, H3K9me3 and H3K14ac [85]. BAH domains are found in chromatin-associated proteins

and have a diverse range of functions [86]. Origin of replication complex component, ORC1, is critical for DNA replication, and contains a BAH domain [87]. ORC1 BAH specifically recognizes H4K20me2, which is enriched at replication origins [88]. Disruption of BAH binding causes reduced occupancy of ORC1 at replication origins and impairs cell-cycle progression. Furthermore, it facilitates ORC2 binding to chromatin as during the shift from M to G1 phase [89]. Interestingly, the BAH domains of the mammalian BAH domain-containing protein 1 (BAHD1) and BAH and Coiled-coil Domain-containing Protein 1 (BAHCC1) recognize and bind to tPRC2-mediated H3K27me3 [90–92]. Mutating the BAHD1 BAH domain increased histone acetylation at polycomb target genes. A CRISPR-mediated mutation of *Bahd1* that specifically disrupts the ability to bind H3K27me3 in mouse resulted in embryonic lethality [91]. Additionally, *Bahd1*$^{+/-}$ male mice have anxiety-like behaviors [93].

ASH1L has an impressive array of functional domains; however, how these domains function in the context of one another is not well understood. It has been shown, in both *Drosophila* and mouse models, that the methyltransferase SET domain is not necessary for survival. Given the binding capabilities of the bromodomain, PHD and BAH domains, *ASH1L* may have important roles in histone marker recognition or facilitate chaperoning other proteins to chromatin. This could be assessed experimentally, such as introducing in-frame loss-of-function mutations of the different domains, or by using biotinylated *ASH1L* and mass spectrometry to determine interacting protein partners.

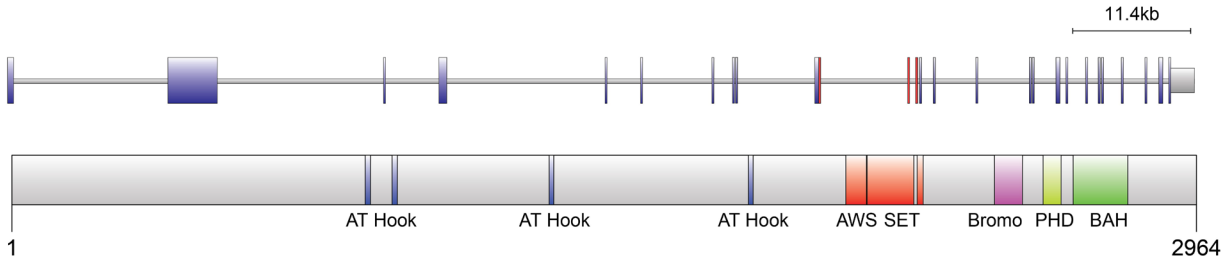


Figure 2: *ASH1L* gene and protein

ASH1L gene (top) and protein structure (bottom). AWS, associated with SET; SET, Su(var)3-9, Enhancer-of-zeste and Trithorax; PHD, plant homology domain; BAH, bromo-adjacent homology.

1.3 Population Studies and Clinical Reports of *ASH1L* in the Context of NDDs

NDD is the medical classification of a range of disorders that impact learning, social and interactive behaviors which manifest during childhood. These include ASD, ID, obsessive compulsive disorder (OCD), attention deficit/hyperactivity disorder (ADHD) and Tourette syndrome (TS). These disorders range from mild to severe, and two or more may be present in a single individual [94]. Substantial evidence indicates that NDDs have an underlying genetic etiology [95,96]. Advances in exome sequencing, both in terms of specificity and affordability, have aided clinical geneticists in identifying underlying patient variants associated with disorders. Furthermore, collecting genetic data from individuals with specific NDDs, such as ASD, and comparing with control groups is proving highly informative. *ASH1L* has come under increased scrutiny as this body of literature grows, as population analyses and clinical reports have highlighted it as an NDD-risk gene.

In one of first population studies to implicate *ASH1L* in NDD risk, Iossfov *et al.* performed exome sequencing of 343 families, with one child diagnosed with ASD, at least one unaffected sibling, and parents. They developed a novel multimodal test to determine *de novo* variants, identifying single nucleotide variants (SNVs) and indels that conferred 'likely gene disruption' (LGD) that perturb protein function. *ASH1L* was one of these 59 LGD candidates [97]. In another study, de Ligt *et al.* performed exome sequencing of 100 individuals with ID (IQ below 50) and unaffected parents. They identified *de novo* candidate genes by excluding variants also found in parents, and confirmed these genes by resequencing 765 individuals with ID. Missense variants in *ASH1L* were found to be associated with ID [98].

Subsequent population studies sought to widen the criteria for assessment by not focusing solely on *de novo* variants. In an ASD study, De Rubeis *et al.* assessed 15,480 DNA samples from 16 sample sets, of which there were 2,270 trios (2 unaffected parents and one affected child). The authors used transmission and *de novo* association (TADA), a statistical model that integrated *de novo,* inherited and case-control variation. Using TADA, they identified 33 autosomal genes, including *ASH1L*, with a false discovery rate (FDR) <0.1. The authors suggest that these genes converge on three biological processes, namely chromatin remodeling, transcription and splicing, and synaptic function [99].

Krumm *et al*. sought to assess the impact of *de novo* and inherited variants. The authors reanalyzed exome sequencing from 8,917 exomes of individuals with ASD and controls, including 2,377 families collected from the Simons Simplex Collection. The analysis included SNV and copy number variants (CNVs) identification; the independent SNV callers in the authors' pipeline allowed them to find *de novo* events in an unbiased manner. Again, *ASH1L* was among the genes that had a LGD variant [100]. Sanders *et al.* conducted further CNV and exome analysis of samples in the Simons Simplex Collection, the Autism Genome Project, and the Autism Sequencing Consortium. They found small *de novo* deletions and exome variants enriched in a common set of ASD risk genes. 65 risk genes had an FDR =<0.1; *ASH1L* had a striking FDR =< 0.01. Interestingly, *ASH1L* was among eight high-risk ASD genes that contributed to chromatin remodeling when gene ontology (GO) analysis was performed, the only significant term [101].

Large studies have been instrumental in gathering genetic data of individuals with an NDD. Wang *et al.* used molecular inversion probes (MIPs) to capture and sequence 213 genes, including *ASH1L*, in an ASD cohort consisting of 1,045 Chinese trios. They assessed genes that had the most recurrent *de novo* variants enriched in affected individuals, and found *ASH1L* ranked 7th in their list [102]. Stessman *et al.* used single-molecule MIPs (smMIPs) of 208 NDD candidate genes to sequence 13,407 probands across 15 international groups. Where available, some unaffected siblings and parents also underwent targeted sequencing. 91 genes were found to be high risk NDD genes, including *ASH1L.* Two *ASH1L* LGD and two *ASH1L* missense variants (all *de novo*) were identified, as well as additional *ASH1L* variants for which inheritance analysis was not available. Follow-up of these individuals confirmed all had ID, with some also exhibiting ASD and/or seizures [103].

Satterstrom *et al.* performed exome sequencing of 35,584 samples, 11,986 of which had ASD, including 6,430 ASD cases, 2,179 unaffected siblings, and both parents. The authors were able to classify protein truncating variants, in addition to missense, as well as determine *de novo* events. TADA analysis was adjusted to include 'probability of loss-of-function intolerance' (pLI) and 'missense badness, PolyPhen-2, constraint' (MPC) scores to assess the burden of variants. Using this sophisticated mathematical model, the authors identified 102 ASD risk genes, all passing FDR =< 0.1, and of these, 26 genes, including *ASH1L,* had a family-wise error rate (FWER) =< 0.05. [104].

Fu *et al.* performed meta-analysis of 63,237 exomes, including 15,036 affected probands and 28,522 parents. The authors found protein truncating and missense

variants enriched in individuals with ASD, with 72 high-risk ASD genes with an FDR =< 0.001. *ASH1L* ranked 15th in their 72 gene list, with protein truncating variants being the most frequent. They found *ASH1L* variants were more likely to be *de novo* than inherited. Interestingly, the authors found *ASH1L* has an association with schizophrenia [105].

The Deciphering Developmental Disorders Study carried out exome sequencing of 4,293 families with a member diagnosed with an NDD. As part of a meta-analysis, they including 3,287 previously published exomes of individuals with an NDD to identify NDD-risk genes [106]. With this data, Faundes *et al*. carried out further investigation of variants residing in lysine methyltransferases and demethyltransferases. The authors identified two individuals with an *ASH1L* protein truncating variant, and five individuals with deletions encompassing the gene. All these individuals presented with variable ID, and in some cases, seizures and developmental delay. The authors argue that *ASH1L* haploinsufficiency contributes to a predominantly NDD phenotype, while physical anomalies may or may not be present [107].

In an investigative study to determine the diagnostic efficacy of using chromosomal microarray analysis (CMA) and WES for ASD, Tammimies *et al*. performed WES on 95 trios and identified a *de novo* loss of function variant of *ASH1L* in an individual with ASD [108]. Aspromonte *et al.* developed a low-cost gene panel for ID and ASD, among other NDDs. They sequenced 150 unrelated individuals from 17 Italian hospitals that presented with one or more NDD. The authors used a multiplex PCR assay for 74 selected genes, including *ASH1L*. The data was then assessed *in*

silico to determine putatively damaging variants. A *de novo* variant in *ASH1L* was identified (p.(Ile1107Val)) in an individual [109].

Epilepsy with myoclonic atonic seizures (MAE) is a rare syndrome associated with severe neurodevelopmental impairment. Interestingly, some individuals with *ASH1L* variants have reported to have seizures [5]. Tang *et al.* performed exome sequencing on 85 individuals with MAE, with 62% patients also exhibiting ID. The authors found pathogenic variants in 12 individuals, including one *ASH1L* variant (p.(Arg1342*)) [110]. While a small sample size, it is an interesting finding, given the role of activity-dependent repression of neurexin 1α, a presynaptic adhesion molecule, and prefrontal cortical knockdown of *Ash1l* in mice resulting in severe seizures [22,111].

TS is a neurodevelopmental disorder, characterized by chronic motor and/or vocal tics. Liu *et al.* performed exome sequencing on 100 trios (proband affected with TS and unaffected parents) to determine *de novo* SNVs and indels that may contribute to the disorder. The authors filtered for rare damaging variants, including variants with an allele frequency of =<0.01 in the 1000 Genomes Project and with a CADD score of greater than 15. Independently, the authors also performed analysis of inherited rare damaging mutations burden at the individual gene level using the rare-variant transmission disequilibrium test (RV-TDT). Two genes, *ASH1L* and *KIF26B,* proved significant in both analyses. The authors identified one *de novo ASH1L* variant and four transmitted *ASH1L* variants in this study. Furthermore, the authors conducted a replication analysis by sequencing exons of both *ASH1L* and *KIF26B* in 524 individuals with TS and 2822 individuals without psychiatric disorders. *ASH1L* had a significantly high variant burden but *KIF26B* did not*;* the authors focused on *ASH1L* for the reminder

of their study. In depth clinical assessment of the five individuals with *ASH1L* variants in the initial trio sequencing was subsequently performed; none of the five had familial history of TS, OCD or ADHD. All individuals had severe symptoms, including facial tics, and three also presented with ADHD and/or OCD. The authors set out to assess the H3K4 methyltransferase activity of two *ASH1L* missense variants (c.6230A>T; (p.Tyr2077Phe) and c.6598A>G; p.(Ser2200Gly)); both variant residues are within or close to *ASH1L*'s SET domain. An ASH1L SET-PHD (amino acids 2069–2636) was cloned from full length human *ASH1L* complementary DNA and the variants were generated by PCR. The constructs were expressed in *E. coli* BL21 (DE3) cells, and the protein products were purified. Using an artificial H3 peptide and methyltransferase assay kit, the authors found a significant decrease in H3K4 methyltransferase activity in both variants compared to wildtype ASH1L SET-PHD, consistent with loss of function [112]. In a different study, Liu *et al.* screened for three *ASH1L* single nucleotide polymorphisms (SNPs) in 271 trios affected with TS and 337 healthy controls in the Chinese Han population. The authors used a TaqMan assay and primers flanking each SNP: rs5005770, rs12734374, and rs35615695. The authors found significant overtransmission of two of the three SNPs in their family-based study, whereas none of the SNPs were significant in the case-control study regarding TS. This analysis suggests that inherited *ASH1L* variants may confer risk for TS [113].

Various other studies have carried out whole genome, exome or targeted sequencing to identify candidate genes in NDD cohorts. They have contributed to clinical databases and add to the statistical significance that *ASH1L* is a high risk NDD gene [114–120].

There are many clinical reports of individuals with an *ASH1L* variant and variable clinical features. In 2017, Okamoto *et al*. reported on a 4-year-old male presenting with growth delay, microcephaly and dysmorphic facial features, including low set ears and hypertelorism. He was nonverbal and exhibited severe ID. Brain magnetic resonance imaging (MRI) found delayed myelination of the cerebral white matter, and a computed tomography (CT) scan found vertebrae fusion of C2 and C3. The authors performed exome sequencing on the proband and parents; they identified a novel, *de novo ASH1L* variant (c.8356G>C; p.(Ala2786Pro)), which lies in the BAH domain [121]. Shen *et al.* described a 9-year-old male with distinctive facial features, including blepharophimosis and small ears. The individual also exhibited mild ID and ADHD. The authors performed trio exome sequencing and found a novel, *de novo ASH1L* variant (c.2422_2423delAAinsT; p.(Lys808TyrfsTer40)). This variant lies in exon 3, putatively resulting in *ASH1L* loss of function [122].

Xi *et al*. reported a 19-year-old pregnant female that presented with severe ID and distinct facial features, including prominent forehead and small ears. The proband's father had low intelligence, and the mother had normal intelligence. The authors performed chromosome microarray analysis (CMA) on the proband, fetus, the proband's parents, paternal grandparents and three paternal uncles. CMA revealed a 1q22 microdeletion, which includes *ASH1L,* that was present in the proband, the fetus and proband's father; other family members did not have the microdeletion. While the microdeletion region encompassed 24 genes, only five (namely *ASH1L, TRIM46*, *CLK2, GON4L* and *ARHGEF2*) loss of function alleles are not tolerated in the general population. The authors postulate that heterozygous loss of *ASH1L* may be driving the

clinical phenotypes, given the prevalence of literature on the gene in the context of NDDs [123].

Liu *et al*. reported a case of monozygotic, female twins that presented with mild intellectual disability and seizures. Exome sequencing was carried out for the twins, parents, and the unaffected brother. The authors found a novel, *de novo* variant in the twins (c.2678dup; p.(Lys894*)), which was confirmed with Sanger sequencing. In addition to the clinical report, the authors interrogated the literature on *ASH1L* patient variants to determine if there was a genotype-phenotype correlation. Of the 14 detailed *ASH1L* variant cases available at the time (including the Liu *et al.* study), four patients had missense variants; three had severe developmental delay, whereas only one patient with a truncating variant had severe developmental delay. The author argue that *ASH1L* missense variants confer greater severity compared to truncating variants [124]. While intriguing, analysis will be needed to be carried out on a greater number of detailed clinic reports of individuals with an *ASH1L* variant to draw conclusions regarding genotype-phenotype correlations.

The Simons Foundation Autism Research Initiative (SFARI) Gene database has recorded two common variants and 120 non-redundant, rare *ASH1L* variants collected from the published literature, all diagnosed with ASD. The two common *ASH1L* variants were found in the TS Chinese Han population study using the SNP-based approach [113]. Both variants reside in introns; however, it is unclear if these variants represent the SNP loci used in the study or are potentially causative variants themselves. The 120 rare variants were identified in the previously cited literature, in addition to several more studies (**figure 3**) [125–129]. Of these rare variants, 66.7% are missense (*n* = 80), 20%

are frameshift (*n* = 24), 10% are stop gained (*n* = 12), 2.5% affect splicing (*n* = 3) and 0.8% is a copy number loss (*n* = 1). 55.8% are of unknown inheritance (*n* = 67), while 27.5% are *de novo* (*n* = 33) and 16.7% are familial (*n* = 20). Of the 24 confirmed inherited variants, 55% are maternal (*n* = 11) and 45% paternal (*n* = 9). The SFARI Gene scoring system, which ranks genes based on published variants, gives *ASH1L* a score of 1, meaning it is classified as a high confidence ASD gene. Additionally, SFARI provides Evaluation of Autism Gene Link Evidence (EAGLE) scoring system. EAGLE assesses the evidence that a gene is specifically relevant to ASD, teasing apart its contribution to other NDDs [130]. *ASH1L* has an EAGLE score of 14.15; this score is considered 'strong', meaning *ASH1L* has been shown repeatedly to have a significant role in ASD.

Population genetics, clinical reports and meta-analyses have provided substantial evidence that variants in *ASH1L* confer risk to NDDs, such as ASD, ID, and TS. The GnomAD Browser has determined *ASH1L*'s pLI score is 1, meaning *ASH1L* truncating variants are not tolerated [131]. Individuals with an *ASH1L* variant are heterozygous, and while two variants have been functionally assessed to determine loss of methyltransferase activity, the vast majority have been predicted to be loss of function [112]. *ASH1L* variants are spread throughout the protein, suggesting there is minimal correlation between domain variants and phenotype. The use of *in vitro* and *in vivo* models has garnered great insight into the roles *ASH1L* plays in neurodevelopment and NDDs.

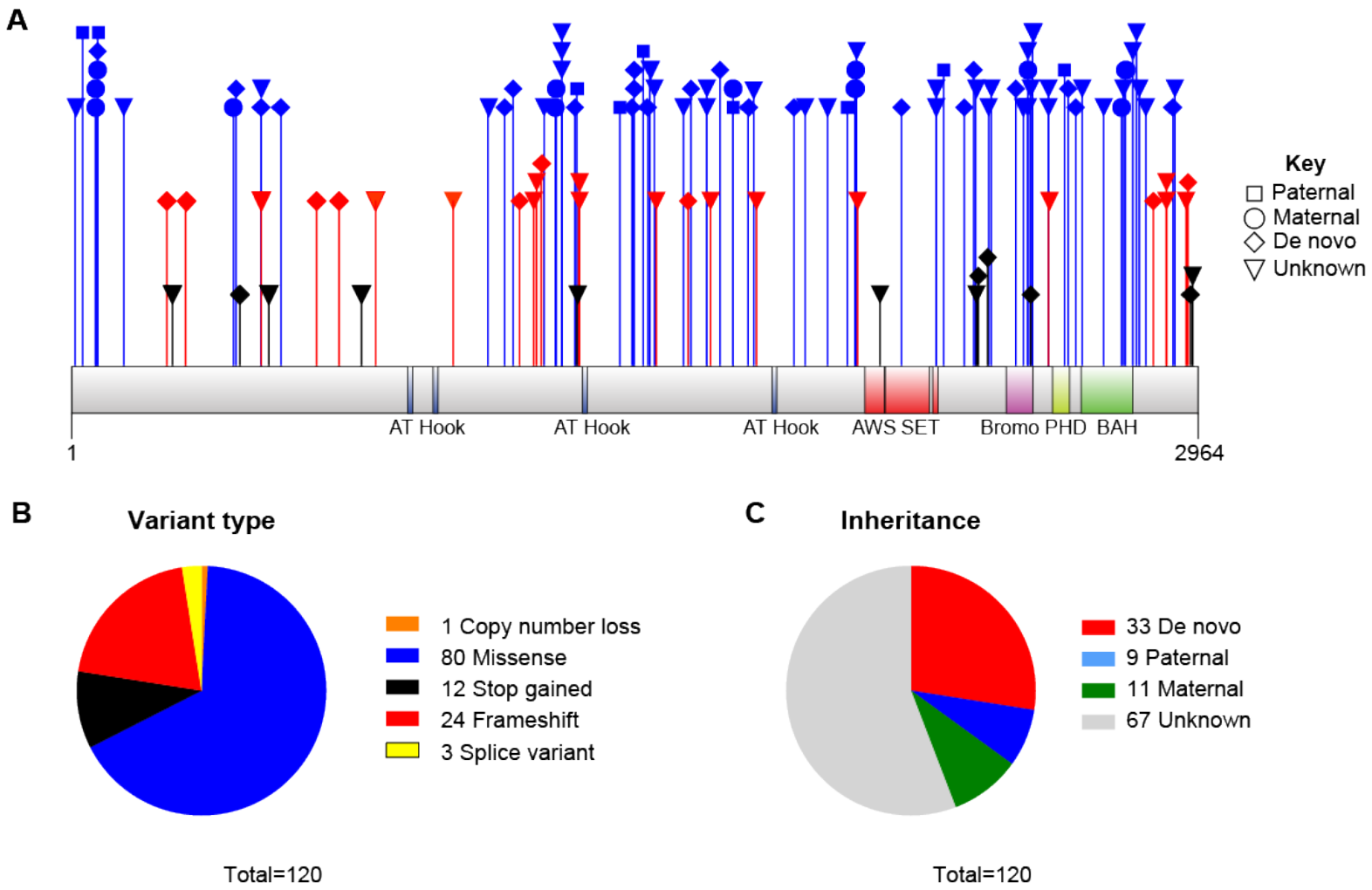


Figure 3: *ASH1L* variants in individuals with ASD

(A) *ASH1L* variants identified in individuals with ASD from the SFARI Gene database [6]. Missense variants are in blue, frameshift in red and premature stop codons in black. Schematic made with Illustrator for Biological Sequences (IBS) [132]. (B) Pie chart of types of variants found. (C) Pie chart of modes of inheritance.

1.4 Mouse and *In Vitro* Models of *ASH1L*

Over the past ten years, several research groups have developed cell lines and mouse models to determine the role of *ASH1L* in different biological systems. The models vary from complete ablation of *ASH1L,* reduction of its expression, or specific mutated domains. According to BLAST, there is 91.8% protein homology of ASH1L in human and mouse; the functional domains of the protein are highly conserved. Mice have similar organ systems to humans, including the brain, which have allowed scientists to make inferences on human biology. Genetic engineering, including genetrap alleles and tissue-specific conditional knockout models, broadens the scope of how these models can be used. Furthermore, gene knockdown with shRNA can be utilized to reduce *Ash1l* expression in a cost-efficient and effective manner. In this section, main conclusions drawn from *ASH1L* models will be discussed. **Table 1** provides a brief overview.

Table 1: Overview of *Ash1l* mouse models

Mouse	Allele description	Phenotypes	References
Ash1l$^{+/-}$ (exon 2)	Heterozygous 11bp deletion in exon 2	Tics, excessive grooming, anxiety, poor discrimination, reduced vocalisation, repetitive and poor social behaviours. Normal brain structure, reduced neuronal pruning, normal neuron density in hippocampus, higher cFos expression.	[24,111,112]
Ash1l$^{-/-}$ (exon 2)	Homozygous 11bp deletion in exon 2	Lethal between e14.5 and p7 from unknown causes.	[111]
Ash1l$^{-/-}$ (exon 4)	Homozygous deletion of exon 4	Lethal at p0, mild skeletal anomalies, Some SATB2 neurons in deeper layer neurons.	[133]
Ash1l Nes-cKO	Homozygous, conditional deletion of exon 4 using neural precursor cell Nestin-cre	Excessive grooming, anxiety, paw clasping, poor discrimination, repetitive and poor social behaviours. Nissl stain of the cortex shows altered barrel organization, reduced myelin basic protein at P21, higher cFos expression.	[133–135]
Ash1l$^{GT/GT}$	Homozygous gene trap cassette in intron 2	Early lethality (~3 weeks), smaller mice, infertile, skeletal defects, lack Meibomian glands and abdominal fat. Altered HOX gene expression in male and female reproductive tract. The long term hematopoietic stem cell compartment is compromised and the mutants progress to bone marrow failure.	[136,137]

Ash1l$^{\Delta SET/\Delta SET}$	Homozygous in-frame deletion in SET domain	Posterior shift of *Hoxb4, Hoxd4* expression; mild skeletal anomalies.	[23]
Ash1l ^{KD}	Short hairpin RNA knockdown in prefrontal cortex	Severe seizures result in death within 10 days. Neuronal hyperactivity.	[22]

1.4.1 ASH1L is Necessary for Survival

Several mouse models have examined the consequence of total ablation or sizable reduction of *Ash1l* expression; they all converge on the same outcome: *Ash1l* is an essential gene that, when reduced or absent, results in lethality. Using CRISPR-Cas9, Zhu *et al.* generated an 11bp deletion in exon 2 of *Ash1l* (c.1002_1013del^{GAATCAGGAAA}). This null allele was made in effort to determine *Ash1l*'s role in the regulation of neurexin-1α (discussed in section **1.4.11**). To generate *Ash1l* null mice (*Ash1l*$^{-/-}$), heterozygous mice (*Ash1l*$^{+/-}$) were crossed; at e14.5, all genotypes were present; however, by p7, no *Ash1l*$^{-/-}$ mice were found. *Ash1l*$^{+/+}$ and *Ash1l*$^{+/-}$ mice were present in expected numbers and *Ash1l*$^{+/-}$ had no discernable phenotypes [111].

Gao *et al.* generated a null allele by crossing *Ash1l* floxed exon 4 (*Ash1l*$^{fl/fl}$) and CMV-cre, a globally expressed *Cre.* Homozygous null mice (*Ash1l*$^{-/-}$) were identified at p0 and had comparable body size and weight to their wildtype littermates. However, the authors reported that the *Ash1l*$^{-/-}$ pups died within 24 hours of birth [133]. Both these models indicate that total loss-of-function of *Ash1l* results in early lethality.

An *Ash1l* gene trap allele is hypomorphic, reducing *Ash1l* expression >80%. The genetrap cassette was inserted into intron 1 of *Ash1l* (*Ash1l*$^{+/GT}$) and these mice were crossed to generate *Ash1l*$^{+/+}$, *Ash1l*$^{+/GT}$ and *Ash1l*$^{GT/GT}$. Over 1000 mice were genotyped to determine viability. E11.5-e18.5 embryos and p1-7 mouse cohorts had genotypes present at normal Mendelian ratios, whereas in the p8-p21 cohort, *Ash1l*$^{GT/GT}$ were significantly underrepresented [136]. This data indicates that reduced *Ash1l* expression in mice confers the risk of lethality after p7.

Together, these studies show that mice with 50% *Ash1l* expression survive into adulthood, whereas if the expression is reduced further, it results in early death. Given the fact that exon 4 floxed mice survive birth, and *Ash1l*$^{GT/GT}$ mice survive to p7 suggests roles of *Ash1l* not limited to *in utero* development. The cause of death in these knockout and genetrap mice has not been determined. An interesting experiment would be to use various tissue-specific *cre* lines to conditionally knockout *Ash1l* in different organ systems.

1.4.2 Loss of ASH1L Results in Reduced Body Weight and Smaller Mice

Ash1l$^{GT/GT}$ and homozygous knockout of *Ash1l's* exon 4 with Nestin-cre (specific to neural progenitor cells; *Ash1l*-Nes-cKO) have been examined regarding their size, weight, and fat composition. At p1-7 *Ash1l*$^{GT/GT}$ hypomorphic mutant mice have same body weight as their *Ash1l*$^{+/GT}$ and *Ash1l*$^{+/+}$ littermates; however, at p14, surviving *Ash1l*$^{GT/GT}$ mice weigh significantly less. At p21, *Ash1l*$^{GT/GT}$ mice are visibly smaller than control littermates. Brinkmeier *et al.* also found a dramatic reduction in abdominal body fat in 2- to 3-month-old *Ash1l*$^{GT/GT}$ mice. *Ash1l*$^{GT/GT}$ mice lacked meibomian glands, which provide lubrication of the eye, and this contributed to chronic blepharitis and ulceration. The female *Ash1l*$^{GT/GT}$ mice had a four-fold reduction in endometrial uterine glands and were infertile. The males were sub-fertile due to defects in development of the epididymis. Overall, *Ash1l*$^{GT/GT}$ exhibited multiple anomalies suggesting *Ash1l* is necessary for growth and development of many organ systems [136].

Gao *et al.* assessed growth and development of *Ash1l*-Nes-cKO. The *Ash1l*-Nes-cKO mice exhibited reduced body weight compared to controls by p6; the mutants were

also visibly smaller. 3-month-old male and female *Ash1l*-Nes-cKO mice weighed significantly less than their control counterparts [133]. Subsequently, Gao *et al.* reported that 8-month-old *Ash1l*-Nes-cKO mice have significantly reduced body weight [135]. They found subcutaneous and visceral white adipose tissue was reduced in *Ash1l*-Nes-cKO, but brown adipose tissue was comparable between genotypes. Interestingly, histological analysis of white adipocytes found cell number was similar between genotypes; however, cell diameter of adipocytes was greatly reduced in *Ash1l*-Nes-cKO. The authors find that *Ash1l*-Nes-cKO mice exhibit greater levels of locomotion and metabolic activity; to this end, they measured oxygen consumption and carbon dioxide production, and confirmed higher respiratory exchange rate in *Ash1l*-Nes-cKO. Furthermore, they found that *Ash1l*-Nes-cKO mice generated more heat, an indicator of heightened metabolism. Thus, *Ash1l* knockout with Nestin-cre results in smaller mice and reduced body weight. However, expression of the Nestin-cre transgene is not restricted to neural progenitor cells and the transgene itself suppresses growth [138]. Gao *et al*. argue that *Ash1l* knockout in the brain results in increased active behavior, and this over-expenditure of energy depletes triglycerides stored in adipocytes. Given the broad expression of ASH1L in many tissues, further experiments will be needed to address the effects of *Ash1l* on metabolism and body weight.

1.4.3 ASH1L Methyltransferase Activity is Not Essential for Survival

Miyazaki *et al.* generated mouse ESCs with an in-frame deletion in the methyltransferase SET domain (ΔSET ESCs). The authors sought to determine if they could recapitulate their findings *in vivo,* so they generated mice homozygous with the

ΔSET domain (*Ash1l*$^{\Delta SET/\Delta SET}$). These mice were viable, fertile, and largely comparable to WT. However, there were some differences, such as a posterior shift in the boundaries of *Hoxb4* and *Hoxd4* mRNA expression in e10.5 *Ash1l*$^{\Delta SET/\Delta SET}$ embryos. Additionally, a mild skeletal phenotype was observed, which is discussed below [23].

This is an interesting finding, given that homozygous knockout of *Ash1l* results in early lethality. Furthermore, assessment of the *Drosophila Ash1* by Steffen *et al.* found that while the AT hooks and BAH domains are necessary for fly survival, the SET domain is not [77]. Together, this data suggests that other domains of ASH1L are playing vital roles that are essential for life; however, at this moment, these functions are unknown. Future experiments, such as generating in-frame mutations that attenuate function of the other domains in mice will be important in assessing necessity.

1.4.4 ASH1L Contributes to Normal Bone Development and Skeleton Formation

Several studies have found *Ash1l* is necessary for bone development and maintenance. Xia *et al.* generated *Ash1l*-silenced mice using the piggyBac (PB) transposon inserted into intron 15 and found these mice were more susceptible to infection and auto-immune diseases, as well as severe bone destruction at knee joints [139].

Yin *et al.* found decreased *Ash1l* expression in osteoporotic mice. C3H10T1/2 (C3) cells can be differentiated to osteocytes, adipocytes or chondrocytes, and found ASH1L expression increased during each of these differentiation protocols. Interestingly, *Ash1l* siRNA knockdown suppressed differentiation to osteocytes and chondrocytes, instead promoting differentiation to adipocytes. Conversely, upregulation

of *Ash1l* in C3 cells promoted differentiation to osteocytes and chondrocytes, suppressing adipogenesis. ChIP assays determined that *Ash1l*-mediated H3K4me3 at *Osx*, *Runx2, Hoxa10*, and *Sox9* promoters. Expression of these genes are necessary to guide fate to osteocytes and chondrocytes. This data suggests that *Ash1l* may be driving differentiation of mesenchymal stem cells by activating select genes, and potentially have a protective role against osteoporosis [140].

In a broader scope, mouse models have uncovered skeletal malformations associated with *Ash1l* loss of function. Gao *et al.,* assessing homozygous knockout of *Ash1l via* CMV-cre driven deletion of exon 4 (*Ash1l*$^{-/-}$), found p0 *Ash1l*$^{-/-}$ pups had an additional rib [133]. The skeletal abnormalities of *Ash1l*$^{GT/GT}$ mice were also mild. An additional rib was frequently attached to the ribcage and the ribs were attached in a staggered manner to the sternum [136].

Mice homozygous for the in-frame SET domain deletion (*Ash1l*$^{\Delta SET/\Delta SET}$) had different but also mild, variable skeletal anomalies. The penetrance of skeletal anomalies was 53% in *Ash1l*$^{+/\Delta SET}$ and 80% in *Ash1l*$^{\Delta SET/\Delta SET}$ mice. The most prevalent anomaly was a broadened neural arch (C2 to C1). Intriguingly, loss of function of group 4 *Hox* genes results in a similar phenotype [141]. Furthermore, homozygous mutations of *Mel18* (*Mel18*$^{-/-}$), a PRC1 component, result in C2-to-C3 vertebrae transformation in mice. Crossing with *Ash1l*$^{+/\Delta SET}$ partially rescued this phenotype, suggesting *Ash1l* methyltransferase activity and *Mel18* play a role in normal skeletal formation [23].

Ash1l loss of function also causes craniofacial anomalies. Using the same *Ash1l* exon 4 deletion allele described above, Gao *et al.* used Nestin-cre to delete *Ash1l* (*Ash1l*-Nes-cKO). Via micro-CT scan and alizarin red/alcian blue bone staining, the

authors found shortened nasal bones in adult *Ash1l*-Nes-cKO mice compared to controls. Together, these data support the role of *Ash1l* in skeletal development. This is perhaps unsurprising, given that TrxG members are implicated in such roles [142].

1.4.5 ASH1L is Required for Development and Function of Reproductive Organs

Development of the reproductive tract in males and females requires region specific expression of *HOX* genes, much like the code that specifies development of the axial skeleton [143]. *Ash1l*$^{+/-}$ and *Ash1l*$^{\Delta SET/\Delta SET}$ mice can reproduce; however, male and female *Ash1l*$^{GT/GT}$ mice were sub-fertile and infertile, respectively. Female *Ash1l*$^{GT/GT}$ mice had defects in the uterine lining; analysis of *Ash1l*$^{GT/GT}$ uteri found they were smaller in size and were twisted compared to WT, suggestive of a homeotic transformation. The stromal layer was thinner, and the paucity of endometrial uterine glands was consistent with failure of implantation and decidualization. *Foxa2* and *Hoxa10* are important for normal uterine development, and expression of both genes was reduced in the *Ash1l*$^{GT/GT}$ uterus.

Male *Ash1l*$^{GT/GT}$ mice were sub-fertile due to structural defects in the epididymis, a part of the male reproductive tract that promotes sperm maturation [144]. In *Ash1l*$^{GT/GT}$ males, the corpus of the epididymis appeared to homeotically transformed to resemble the caput segment, and *Hoxa11* gene expression was reduced. There was evidence of sperm death in the *Ash1l*$^{GT/GT}$ epididymis. Together, this data suggest *Ash1l* plays a vital role in the formation of both male and female reproductive tracts, resulting in impaired reproductive performance. The *Ash1l* genetrap allele is the only mouse model

for which this is reported, given that other homozygous *Ash1l* null alleles result in early death.

1.4.6 ASH1L and Cortical Development

The cortex is a highly organized outer layer of the brain; it is composed of six layers with different neuron subtypes and projections. It is responsible for many higher order functions, including memory and language; anomalies in cortical development are believed to be central to ASD and other NDDs.

Using the Allen Brain Atlas data set, Cheon *et al.* showed that *ASH1L* is expressed in different regions of the cortex during human fetal development. *ASH1L* mRNA and protein are continuously expressed during the differentiation of human ESCs in culture to form deeper layer excitatory neurons. *ASH1L* knockdown with shRNA decreased the length and neuron branching in these differentiating cultures. Furthermore, neuron soma size was enlarged and ectopic phalloidin-expressing structures at neurite growing tips were present in *ASH1L*-shRNA neurons. The authors treated *ASH1L*-shRNA neurons with EI1, an *EZH2* catalytic inhibitor, and found this rescued the shorter length of *ASH1L*-shRNA neurons, but not the dysmorphology phenotype. *ASH1L*-shRNA reduced expression of the neurotrophin receptor *TrkB*, which is essential for neuronal arborization [145]. Thus, *ASH1L* is necessary for normal development of cortical neurons in culture [146].

Gao *et al*. assessed how *Ash1l* ablation impacted cortical development *in vivo*. Nissl stained cortices of newborn germline knockout mice (*Ash1*$^{-/-}$) revealed a loss of cortical organization, specifically the minicolumnar arrangement. There was a greater

distribution of SATB2 neurons in the deeper cortical neurons, suggesting a delay in cortical development. While these data does indicate anomalies in cortical development, the authors reported the *Ash1*$^{-/-}$ mice die within 24 h of birth; this suggests the tissue they analyzed may be undergoing death and confound results [133].

Yan *et al.* used *Emx1-cre* to excise exon 2 of *Ash1l,* resulting in a cortical specific null allele (*Ash1l*$^{Emx1\text{-}cKO}$). NeuN (a pan-neuronal marker) and DAPI were used to assess neuron number or cell density in cortices; mutants were comparable to controls [24].

Collectively, this data suggests *Ash1l*-associated changes in cortical development are likely to be subtle, with minimal differences observable on the macro scale. Use of specific cortical markers on healthy tissue will be a useful analysis to pursue to determine *Ash1l*'s role in cortical development.

1.4.7 Loss of ASH1L Results in Behaviors Associated with ASD, ID

It is no surprise that, given the human data that strongly indicates *ASH1L* is an NDD-risk gene, several *Ash1l* mouse models have been assessed regarding ASD- and ID-like behaviors. In mice, these include alteration in preference to novel objects or newly introduced mice, increased anxiety-like and self-grooming behaviors.

The most widely used behavioral tests include the open field test, marble burying test, three chamber test, elevated maze test, contextual or tone dependent fear discrimination tests, and assessment of paw clasping.

Open field test: the test mouse is placed in lidless box or cage and their range of movement is monitored. Mice with greater anxiety will avoid moving into the center of

the box, instead remaining along the edges or corners, whereas mice with less anxiety are more likely to explore, including entering the open center [147,148].

Marble burying test: this tests repetitive behavior, with the test mouse placed in a cage with twenty or so uncovered marbles or similar objects in a bed of wood shavings. Mice prone to repetition will bury more marbles compared to non-repetitive mice [149].

Three chamber test: this test is used to assess the sociability and the preference of a mouse to engage with a novel object. A large cage is divided into three distinct, accessible chambers, between which the test mouse can move freely. After habituation, a second mouse is added to a cage in one of the flanking chambers. Mice, instinctually, are social animals and will spend more time with this second mouse, rather than the two empty chambers. A mouse with reduced sociability will not spend more time with the second mouse compared to time spent in the two empty chambers. A third mouse, or novel object, is added to the opposite chamber after some time has passed. Mice exhibiting normal behavior will tend to spend more time in that chamber compared to the other two, as they are curious animals [150].

Elevated maze test: another test for anxiety-like behavior. It consists of two perpendicular crossing beams in the shape of a plus sign. The entire platform is elevated, however two opposite beams have walls, while the other beams are open. Mice exhibiting more anxiety will prefer the closed, safe environment of the walled platforms; normal mice do not discriminate, and move within the maze equally [151].

Contextual or tone dependent fear discrimination tests: this test is designed to assess the learning capabilities and memory of mice. Based on contextual fear conditioning, mice are provided a context (either a sound tone or light) and subjected to

a mild electric foot shock in a cage. Continuous subjection of context-shock gradually results in learning to associate the two – mice will learn that context A results in shock, and freeze upon recalling that context. This is the conditioning portion of the experiment, whereas on a later day, the mice will be tested on recall – if they can discriminate between pain associated context A, or a novel, harmless context B [152,153]. An alternative version of this test is the tone-based Go-No-go task. The mouse's head is fixed in front of a modified water lixit, which, at one tone frequency, provides the animal with water upon licking, and at another tone, fires an air puff in their face at licking. Mice with normal learning behavior will learn to avoid licking the valve at the air puff tone [154].

Assessment of paw clasping: this test indicates anomalies in brain-motor function. Normally, mice held by their tail and lowered will extend all four limbs outward. However, disruption in neuronal pathways can result in mice clasping their hind limbs [155].

Yan *et al.* performed a battery of behavioral tests on *Ash1l* mutant mice: germline heterozygous knockout of *Ash1l* with 11bp deletion in exon 2 (c.1002_1013del^{GAATCAGGAAA}) (*Ash1l*$^{+/-}$) and *Emx1-cre* to excise exon 2 of *Ash1l,* resulting in a homozygous, cortical specific null allele (*Ash1l*$^{Emx1\text{-}cKO}$). In the open field test, *Ash1l*$^{+/-}$ spent less time in the center, and in the marble burying test, *Ash1l*$^{+/-}$ buried more marbles compared to controls. P7-9 *Ash1l*$^{+/-}$ pups made less ultrasonic vocalization compared to controls, another behavior indicative of ASD-like behaviors. In the three-chamber test, *Ash1l*$^{+/-}$ and controls showed comparable sociability, whereas when a second strange mouse was added, *Ash1l*$^{+/-}$ did not show preference for it.

Similarly, *Ash1l*$^{Emx1\text{-}cKO}$ spent less time in the center of the open field, as well as less time on the open arms of the elevated maze, compared to controls. Furthermore, *Ash1l*$^{Emx1\text{-}cKO}$ exhibited decreased social preference in the three-chamber test compared to controls. Interestingly, the homozygous *Ash1l*$^{Emx1\text{-}cKO}$ did not exhibit more severe behavioral anomalies compared to the heterozygous *Ash1l*$^{+/-}$. This may be due to only the excitatory neurons being affected in the conditional knockout.

To assess if *Ash1l*$^{+/-}$ mice had anomalies in learning, the authors used discrimination tests. During conditioning acquisition, *Ash1l*$^{+/-}$ mice took slightly longer to learn the context, and they could not discriminate between context A and context B. In contrast, the controls froze only in the shock associated box. Tone-based discrimination tests, including the Go-No-go task, revealed that the *Ash1l*$^{+/-}$ mice had deficits in discrimination [24].

In the three chamber test, *Ash1l*-Nes-cKO mice showed no preference towards another mouse or the other chamber, whereas controls exhibited social behavior and spent more time with the other mouse. Likewise, when another mouse was added to the other chamber, *Ash1l*-Nes-cKO mice did not have a preference compared to controls, which preferred the new mouse.

The authors observed that most *Ash1l*-Nes-cKO mice clasped their paws when lifted by the tail (*n*= 16/20), whereas controls did not exhibit any paw clasping. In a combined experiment, merging the open field test and novel object test, *Ash1l*-Nes-cKO mice did not exhibit preference for the new object, and spent less time in the center. These results suggest conditional knockout of *Ash1l* using Nestin-cre have ASD-like behaviors [133].

Administration of vorinostat, a histone deacetylase inhibitor, to *Ash1l*-Nes-cKO mice did not ameliorate the reduced weight of mutant mice; however, it improved the sociability and cognitive memory when tested with the three-chamber test and novel object test, respectively. Other behaviors, such as anxiety and discrimination tasks, remained unchanged with vorinostat treatment [134].

Qin *et al.* performed knockdown of *Ash1l* with injection of shRNA into the prefrontal cortex (PFC) of 5-week-old mice. These mice had fatal seizures (discussed below); however, treatment with diazepam (DZ) and CaMKII-driven Gi-coupled Designer Receptors Exclusively Activated by Designer Drugs (Gi-DREADD), in turn activated by administration of clozapine-N-oxide (CNO), prevented seizures. Thus, the authors were able to test behaviors of $Ash1l^{KD}$ mice. In the three-chamber test, treated $Ash1l^{KD}$ mice exhibited no preference in the social stimulus (other mouse) over the non-social stimulus (object), whereas control animals (injected with scrambled shRNA and treated with vehicle) spent more time with the other mouse. Furthermore, in the open field test, $Ash1l^{KD}$ mice spent less time in the center compared to controls, indicating anxiety-like behavior [22].

Together, this data strongly suggests *Ash1l* plays a role in modulating behavior. This is demonstrated consistently with three independent models of reduced *Ash1l* expression; all have increased anxiety and decreased social preference, in line with other ASD mouse models [156].

1.4.8 Loss of ASH1L Results in Tic-Like Behavior and Increased Self-Grooming

Germline, heterozygous knockout of *Ash1l* with 11bp deletion in exon 2 (c.1002_1013del^{GAATCAGGAAA}) (*Ash1l*$^{+/-}$) were specifically assessed for tic-like behavior [112]. Mice were videotaped for 10 min, and isolated head jerking or shakes lasting 0.05–0.1 s were characterized as tic-like behavior; tics were quantified as the mean value of total number of tics in the 10 min video. One-month-old *Ash1*$^{+/-}$ mice had an increased number of tics in accordance with these parameters; additionally, they groomed themselves significantly longer than WT littermates. Both tics and grooming behaviors were ameliorated with treatment of haloperidol, a D2 receptor antagonist used to treat human TS patients. One-year-old *Ash1l*$^{+/-}$ mice exhibited skin lesions, likely caused by excessive self-grooming. However, the authors point out that the presence of skin lesions may alter self-grooming behavior in itself [112].

Other *Ash1l* mouse studies found increased self-grooming behavior in mutants, including *Ash1l*-Nes-cKO and *Ash1l*^{KD} with shRNA [22,133,134]. Self-grooming is a common behavior trait of mice displaying ASD-like behaviors, which of course further supports the argument that *ASH1L* is an NDD-risk gene. However, Li *et al.* demonstrated that *Ash1l*-silenced mice (with PB transposon inserted into intron 15) exhibited skin hyperplasia, disrupted keratinocyte differentiation and overproliferation, and impaired wound healing [157]. The resulting inflammation may be further exacerbated by grooming. A counterpoint to this is that the *Ash1l* knockdown via shRNA is localized to the PFC, so it is unlikely that all excessive self-grooming can be attributed to a skin specific phenotype.

1.4.9 Loss of ASH1L Results in Seizures

Qin *et al.* performed knockdown of *Ash1l* in mouse PFCs by stereotaxic injection of shRNA AAV in 5-week-old mice. Controls were injected with scrambled shRNA. The authors confirmed reduction of both *Ash1l* mRNA and protein compared to controls; *ASH1L* protein levels ~75% reduced. Mice infected with *Ash1l* shRNA exhibited severe seizures; hallmarks included head nodding, rearing with forelimb clonus and wild rushing, jumping, and falling over. Of the 20 $Ash1l^{KD}$ mice, all died within 10 days post-infection from apparent seizure related deaths, whereas controls did not exhibit seizures or lethality. Electroencephalography (EEG) recordings found $Ash1l^{KD}$ mice displayed epileptiform activity, a marker of epilepsy, 1-2 days before seizures. Further analysis of brain wave activity showed $Ash1l^{KD}$ mice had increased delta frequency and decreased gamma frequency, which are characteristic of epileptic seizures.

To test treatment options, mice were injected with *Ash1l* shRNA (or scrambled for controls) and Gi-DREADD, a CaMKII-driven neuronal inhibiting system which is activated with CNO. 7-9 days post infection, mice were administered with either DZ or CNO, or both; DZ is a GABAAR positive allosteric modulator, and CNO, a small molecule, activates Gi-DREADD, thereby silencing pyramidal neurons. Interestingly, a combination of both DZ and CNO rescued the mortality phenotype completely, with significant reduction in seizure related EEG readings [22].

The *Ash1l*-Nes-cKO mice have increased seizure susceptibility. The *Ash1l*-Nes-cKO mice and controls were injected with 40mg/kg pentylenetetrazol (PTZ), a GABAA receptor antagonist. This dose is considered sub-convulsive, and it elicited a mild response in wildtype mice, including face twitching. However, the *Ash1l*-Nes-cKO mice

exhibited more severe epileptic behavior, including heavy myoclonic jerks and clonic-tonic spasms. EEG traces before and after PTZ injection revealed that *Ash1l*-Nes-cKO mice had spike-wave electrical discharges with higher amplitude compared to controls [135].

These two studies describing seizure activity in *Ash1l* mutant mice are intriguing. It is curious that knockdown of *Ash1l* in a localized region of the cortex results in fatal seizures, whereas conditional knockout in the whole brain only confers risk to seizure-like behaviors after the administration of PTZ. Further analysis will be needed to confirm seizure susceptibility in *Ash1l* mutant mice, such as administering PTZ to *Ash1l*$^{+/-}$, followed by EEG monitoring. Assessing different conditional knockout lines of *Ash1l,* such as PFC specific (or use of a tamoxifen inducible cre to coincide with PFC development), may provide further data on seizure susceptibility with regards to *Ash1l*.

1.4.10 ASH1L is Necessary for Balance Between Neuronal Excitation and Inhibition

Three *Ash1l* mouse models have been characterized regarding neuronal activity. In normal brains, excitatory and inhibitory (E/I) synaptic transmission is tightly regulated. Disruption in E/I balance has been implicated in a range of NDDs, including ASD and TS [158].

To assess whether their *Ash1l*-Nes-cKO mouse model exhibited differences in neural activity, Gao *et al.* performed IHC with c-Fos, a marker for neuronal activation. The authors sectioned paraffin embedded brains of 1, 4 and 8-week-old *Ash1l*-Nes-cKO mice and controls. Focusing on different brain regions, namely the amygdala,

hypothalamus and motor cortex, c-Fos$^+$ cells were counted. While 1-week-old mice did not have differences in c-Fos$^+$ cell number across the three brain regions between genotypes, 4- and 8-week-old *Ash1l*-Nes-cKO had significantly more c-Fos$^+$ cells in each brain region that was studied. The authors reason that this increase in c-Fos$^+$ cells is the result of hyperactivity in the mutant brains. They relate this finding to the increase in locomotor activity observed in *Ash1l*-Nes-cKO mice [135].

To assess the neural activity of the *Ash1l*$^{+/-}$ that exhibit tic-like behavior and excessive self-grooming, S. Liu *et al.* also performed IHC with c-Fos on vibratome sectioned brain tissue. The authors focused on the dorsal striatum, which is implicated in animal models that exhibit repetitive behaviors and excessive self-grooming [159]. C-Fos$^+$ cells were counted and normalized to total cell number. Both 1-month and 1-year-old *Ash1l*$^{+/-}$ mice exhibited an increased proportion of c-Fos$^+$ cells compared to *Ash1l*$^{+/+}$ controls. Furthermore, *Ash1l*$^{+/-}$ mice with skin lesions had significantly more c-Fos$^+$ cells compared to *Ash1l*$^{+/-}$ mice without skin lesions. The authors also assessed levels of dopamine (DA) releasing events in *Ash1l*$^{+/-}$ mice, as abnormal modulation of DA has been implicated in TS [160]. Using a fluorescent DA sensor and fiberoptic recording, it was found that more DA-releasing events occurred in *Ash1l*$^{+/-}$ mice compared to controls. DA-releasing events were decreased to control levels upon haloperidol administration, a D2 receptor antagonist [112].

Qin *et al.* assessed neural activity in PFC of *Ash1l*^{KD} with shRNA and scrambled shRNA control mice. The authors performed whole-cell patch-clamp recordings of PFC slices, specifically on deep layer glutamatergic pyramidal neurons. These neurons were chosen given obvious anomalies found in children with ASD [161]. *Ash1*^{KD} neurons

exhibited higher frequency of synaptic-driven, spontaneous action potentials compared to controls. *Ash1*^{KD} neurons also fired at hyperpolarized potential (-65mV), which did not occur in any control samples. The authors found the ratio of excitatory postsynaptic currents (EPSCs) to inhibitory postsynaptic currents (IPSCs) was significantly increased in shRNA infected PFCs, indicating hyperactivity in these neurons. Interestingly, spontaneous IPSCs had decreased amplitude and frequency in *Ash1l*^{KD}, and spontaneous EPSCs fired at increased frequency, indicating dysregulation of both excitatory and inhibitory synaptic responses that could contribute to the hyperactivity. Treatment with DZ and CNO activated Gi-DREADD restored this neuronal activity back to normal levels [22].

Together, this data indicates roles of *Ash1l* in maintaining synaptic activity. With c-Fos IHC and whole-cell patch-clamp analysis, these models show hyperactivity results from *Ash1l* loss in the brain. This hyperactivity is likely driving ASD-like, tic-like behaviors and seizures exhibited in the mutants, suggesting that *Ash1l* has roles in post development by modulating genes required for this balance. With E/I imbalance as a potential underlying mechanism for *ASH1L* affected individuals, this informs researchers and clinicians of therapeutic targets.

1.4.11 ASH1L and Gene Expression Regulation

Ash1l is an important gene for viability and for development of skeletal structures and organs including the brain. *Ash1l* has multiple domains, including the H3K36 methyltransferase SET domain, DNA binding AT hooks and histone reading Bromodomain and PHD. The function of these domains converge on gene regulation.

Either through direct or indirect mechanisms, *Ash1l* controls gene expression to co-ordinate proper development and function.

As a TrxG member, its ortholog *Ash1* was shown to regulate *Hox* gene expression. Mammalian *Ash1l* has also been found to modulate *Hox* genes *in vitro* and *in vivo.* With ChIP, Gregory *et al*. found *ASH1L* occupies the entire transcribed region of *Hoxa10* in HEK-293T cells. Using shRNA, the authors reduced *Ash1l* expression and thereby reduced its occupancy at the normally active *Hoxa10* and *Hoxa6* loci. Reduced expression of these genes was observed with RT-PCR [162].

Miyazaki *et al.* demonstrated that in-frame deletion of *Ash1l*'s SET domain modestly reduced *Hoxd4* expression, and in e10.5 *Ash1l*$^{\Delta SET/\Delta SET}$ mice, there was posterior shift in *Hoxb4* and *Hoxd4* expression patterns [23]. In *Ash1l*$^{GT/GT}$ mice, Brinkmeier *et al.* observed reduced expression of *Hoxa10* in uteri and reduced expression of *Hoxa11* and *Hoxd10* in epididymis tissue [136]. Jones *et al.* also assessed the *Ash1l*$^{GT/GT}$ mice*,* specifically with regards to *Ash1l*'s role in hematopoietic stem cells. The authors sorted and purified LSK progenitors (fetal liver and BM lineage–SCA1^{+}c-KIT^{+} cells) and showed that *Hox* genes, *a3-10,* were all reduced in *Ash1l*$^{GT/GT}$. Expression of *Meis1,* a *Hox* related gene, was also reduced.

Wang *et al.* demonstrated *Hoxa10,* in addition to *Col1, Alp* and *Sox9* were all downregulated in siRNA-mediated knockdown of *Ash1l* in C3 cells. Zhang *et al*. microinjected *Ash1l* overexpression (OE) vectors into e16.5 mouse ovaries; this resulted in ~2-fold increase in *ASH1L* protein expression in ovaries compared to controls. Expression of selected genes, including *Dmc1, Rad51, Rec8, Spo11, Msh4* and *Stag3,* was increased in *Ash1l* OE ovaries. These genes are involved in DNA

damage and repair; rather than *Ash1l* OE directly resulting in their increased expression, the authors found *Ash1l* OE resulted in oocyte double strand breaks [163].

Ash1l directly or indirectly regulates genes important for neurodevelopment and learning. Gao *et al.* crossed their *Ash1l* exon 4 floxed mice (*Ash1l*$^{2f/2f}$) with a tamoxifen inducible cre line (*Rosa26-CreERT2*$^{+/+}$). NPCs from *Ash1l*$^{2f/2f}$;*Rosa26-CreERT2*$^{+/+}$ mice were extracted from the subventricular zone (SVZ) and plated with serum-free culture. The addition of tamoxifen resulted in excision of exon 4, resulting in an *Ash1l*-KO NPC line. To assess gene expression differences between *Ash1l*-KO and control cells, the authors differentiated the NPCs and collected at 0, 12 and 24 hours and performed RNA-seq. Sequencing uncovered differentially expressed genes, including downregulation of *Emx2, Foxg1* and *Pcdh10* and confirmed with RT-PCR. Gene ontology (GO) analysis found regulation of cell migration, regulation of developmental growth and brain development enriched in their dysregulated gene sets. [133].

Qin *et al.* performed RNA-seq on PFC *Ash1l*^{KD} and controls. They found DEGs that overlapped with ASD-, ID-, and epilepsy-risk genes, and gene set enrichment analysis (GSEA) found high-risk ASD genes to be a highly affected gene module. The authors also used GO analysis; important neuronal pathways were affected, including neurotransmitter transport and cell adhesion. The authors were able to sort these DEGs in accordance with their involvement in excitatory and inhibitory synapse function. These results were confirmed with RT-PCR. Interestingly, the authors performed a ChIP assay to quantify H3K4me3 at downregulated gene promotors and found a reduction of H3K4me3 at four genes important for E/I balance, including *Grm2* and *Slc6a1*. This data

suggests that, through *Ash1l*-mediated H3K4me3, genes important for synaptic function are regulated. [22].

Yan *et al.* performed RNA-seq on auditory cortices and dorsal striatum of 1-month and 1-year-old *Ash1l*$^{+/-}$ and *Ash1l*$^{+/+}$ mice. The authors compared datasets and found *EphA7,* a gene important for axon guidance and topographical organization, was consistently downregulated in *Ash1l*$^{+/-}$. In the EGR1-EGFP mouse line, active neurons express GFP [164]; using tone-dependent fear conditioning, in which mice learn to anticipate foot shocks upon hearing a tone, the authors used FACS to enrich for activated neurons. They found that *EphA7* was highly expressed in these activated neurons, whereas in *ASH1*$^{+/-}$*;EGR1-EGFP* mice, *EphA7* expression was greatly reduced.

The authors found H3K27me3 was enriched at the *EphA7* gene via ChIP in *Ash1*$^{+/-}$, whereas neither H3K4me3 nor H3K36me2 were different between *Ash1*$^{+/-}$ and controls. The authors also observed *Ash1l*$^{+/-}$ neurons had defects in synaptic pruning, which is essential for learning, and can be rescued by activating *EphA7* via ephrin-A5 application. Together, the authors concluded, the data suggests a specific pathway in which *Ash1l* activates the expression of *EphA7* upon learning, which in turn modulates synaptic pruning [24].

Ash1l has been shown to directly regulate neurexin 1α (*Nrxn1α*) in an activity dependent manner. The *Nrxn* gene family (1, 2 and 3) encodes cell adhesion molecules that are essential for synaptogenesis and pruning. Human variants in these genes have been associated with ASD, as well as epilepsy and schizophrenia. Zhu *et al.* demonstrated that transcription of *Nrxn1α* is activity dependent, by treating embryonic

mouse cortical neuronal cell cultures with potassium or optogenetic stimulation. Stimulation caused a prolonged reduction in *Nrxn1α* transcription. To identify the protein(s) regulating *Nrxn1α* transcription *in vivo*, a synthetic zinc finger protein was used to pull down proteins enriched at the *Nrxn1α* promotor in adult mouse brain, and *ASH1L* was identified by mass spectrometry. This was validated with ChIP-seq of *ASH1L*. Following neuronal activity, *ASH1L* and H3K36me2 were enriched at the promoter region of *Nrxn1α*. *Ash1l* was knocked down using doxycycline-inducible shRNA, and *Nrxn1α* expression was no longer repressed following neuronal stimulation.

An *Ash1l*$^{+/-}$ mouse model was generated; the aforementioned 11bp deletion in exon 2 (c.1002_1013del^{GAATCAGGAAA}). *Ash1l*$^{+/-}$ mice had normal hippocampal neuron density. There was a reduction of H3K36me2 at the *Nrxn1α* promotor in *Ash1l*$^{+/-}$ compared to controls. Other histone modifications, such as H3K4me3 remained unchanged between genotypes. Primary neuronal cell cultures from e14.5 *Ash1l*$^{-/-}$ embryos failed to reduce *Nrxn1α* expression in response to KCl or optogenetic stimulation, confirming *Ash1l* is required for activity-dependent repression of *Nrxn1α* [111].

In summary, analysis of four different mouse models of reduced *Ash1l* expression indicate it is important for regulating neurodevelopmental genes. In some cases, genome wide analysis was carried out, and in other cases candidate genes were analyzed. The leading model that *ASH1L* activates gene activation by counteracting PRC2 repression via H3K36me2 is somewhat challenged by the findings of Zhu *et al*, in which *Ash1l* was found to act as an activity-dependent repression of *Nrxn1α*. Of course, this does not necessarily represent conflicting data. RNA-seq in the studies outlined

above show both up- and downregulated genes, indicating *Ash1l* directly or indirectly regulates gene expression.

This section gives a brief overview of *Ash1l* mouse and *in vitro* models, in which genetic manipulation has altered *Ash1l* expression. From the observed results, roles of *Ash1l* are inferred, particularly in relation to general mouse development, neurodevelopment and neuronal activity. It is clear that, in mice, one functioning allele of *Ash1l* is required for survival, as homozygous germline knockout models exhibit early lethality. Comparing $Ash1l^{+/-}$ and $Ash1l^{GT/GT}$ mice suggests there is a required threshold to generate viable, fertile mice. On the other hand, heterozygous *Ash1l* mice display ASD- and tic-like behaviors, indicating that both *Ash1l* allele are needed for normal neurodevelopment. Neuronal activity studies of several *Ash1l* models suggest hyperactivity is a core phenotype of *Ash1l* loss of function, which will no doubt be further tested as a therapeutic target.

1.5 Conclusions and Future Directions

Individuals with heterozygous, loss of function variants in *ASH1L* present with NDDs, including ASD and ID, and some with variable structural birth defects. *ASH1L* is a member of TrxG, which is part of an ancient mechanism of gene regulation that works in tandem with PcG to specify body segmentation, among other essential processes. The *Drosophila* ortholog, *Ash1,* was identified when mutations at its locus resulted in absent, small or homeotically transformed body parts. Human *ASH1L* was later discovered by comparing EST to *Drosophila Ash1*, and found to harbor the same

functional domains, including the histone methyltransferase SET domain. *ASH1L* has been shown to methylate H3K4 and H3K36; its target residue may differ depending on tissue type or developmental timing. It is obvious from mouse studies that *ASH1L* plays important functions that are required for survival and development. While the focus of this review is on *ASH1L* and neurodevelopment, *ASH1L* has been implicated in various biology processes and diseases, such as cancers [165,166].

ASH1L related genes, such as other TrxG and PcG genes, as well as other H3K36 methyltransferases, have been associated with NDDs. Variants in TrxG gene *MLL1* cause Wiedemann-Steiner syndrome, characterized by developmental delay, intellectual disability and ASD. The catalytic component of PcG PRC2, *EZH2*, is linked to Weaver syndrome, individuals with Weaver syndrome exhibit macrocephaly and intellectual disability. Individuals with Sotos syndrome, characterized with childhood overgrowth, have been found to have variants in the H3K36 methyltransferase *NSD1.* Similarly, variants in *ASH1L* have been found in individuals with NDDs. Genetic analyses, such as exome sequencing, whole genome sequencing and gene-specific targeted sequencing of NDD population and controls have found over 130 pathogenic variants in *ASH1L* which are reported on the ASD-specific SFARI database. Interestingly, clinical reports of individuals with *ASH1L* variants describe a range of phenotypes and severity; it has been suggested that missense variants contribute to a more severe diagnosis for patients. Future investigations may use patient derived ESCs to access methyltransferase activity, interacting partners and protein stability.

Basic science investigations with the aid of several *Ash1l* mouse models have demonstrated the importance of this gene. Homozygous knockout models using

different *Ash1l* alleles have early lethality. Cause of death has yet to be determined; however, tissue specific cre lines, such as neural crest or heart, may help pinpoint the most severely affected organ. *Ash1l*$^{GT/GT}$, with reduced *Ash1l* expression, suffer from undergrowth and infertility, with homeotic transformations occurring in reproductive organs. These mice also exhibit early adult lethality. While *Ash1l* haploinsufficient mice survive and are fertile, they present with a range of NDD behaviors, including heightened anxiety, memory defects and tics. Conditional knockout of *Ash1l* using Nestin-cre and knockdown of *Ash1l* in the PFC present with neural hyperactivity, an underlying marker of several NDDs. Furthermore, PFC *Ash1l* knockdown mice have severe seizures that result in death. With the severe seizure phenotype of *Ash1l* knockdown adult mice, it also indicates *Ash1l* is essential after development. Use of a tamoxifen inducible cre global knockout of *Ash1l* may give greater insight into this.

It may be of clinical importance to compare behaviors and neural activity of heterozygous *Ash1l* mice with homozygous *Ash1l* conditional knockouts and knockdown models. Human patients with a loss of function, heterozygous *ASH1L* variant have variable expressivity, possibly due to modifying genes or other genomic contexts. In mice, *Ash1l* haploinsufficiency presents with increased anxiety and decreased social preference, tic-like behavior and prolonged self-grooming. Homozygous *Ash1l* conditional knockouts and knockdown of *Ash1l,* both resulting in less than 50% *Ash1l* expression in the brain/cortex, also exhibit these behavioral phenotypes. However, these mice also have risk of seizures (fully penetrant in the *Ash1l*^{KD} model), suggesting that below the 50% expression threshold, *Ash1l* loss in mice confers a more severe phenotype. Expression levels of *ASH1L* depending on the

genomic context may drive the variable patient phenotypes and may be taken into consideration regarding genotype-phenotype association.

ASH1L's main function is gene regulation. Several studies have utilized RNA-seq, RT-PCR or ISH to determine dysregulation of genes in the context of *Ash1l.* While genes like the *Hox* family have been shown repeatedly to be downregulated in the absence of *Ash1l,* studies are uncovering other genes required for normal brain development and learning. In combination with ChIP, some studies have found evidence of direct regulation, such as *Nrxn1α* and *EphA7.* Future studies may focus on H3K36me2 differences in mutant mice and controls, as techniques such as CUT&RUN are becoming more widely used. Additionally, spatial transcriptomics will likely be used to tackle complex tissues, like the cortex, in relation to *Ash1l* associated transcriptional regulation. With the goal of improving affected individuals' outcomes in mind, basic science research will help direct treatment options in the future.

Chapter 2: *Ash1l* Loss-of-Function Results in Structural Birth Defects and Altered Cortical Development

2.1 Abstract

The histone methyltransferase *ASH1L* is a critical developmental regulator of gene expression for many organ systems, but its role in brain development has not been deeply investigated. Over 130 autistic individuals have heterozygous loss-of-function *ASH1L* variants, and population studies confirm it is a high-risk autism gene. Several studies report autism-like behaviors in *Ash1l* deficient mice and characterized aspects of the underlying neuropathophysiology.

We used mice with a cre-inducible deletion of *Ash1l* exon 4, which results in a frame shift and premature stop codon (p.V1693Afs*2). We assessed the impact of *Ash1l* loss-of-function on survival and craniofacial skeletal development. We used a tamoxifen-inducible cre strain to knockout *Ash1l* early in cortical development (*Emx1cre-ER*^{T2}*;* e10.5). We used immunohistochemistry to assess cortical lamination, and IdU and EdU incorporation to birthdate cortical neurons. We utilized single cell RNA sequencing to compare cortical cell populations and identify differentially expressed genes.

The proportion of homozygous *Ash1l* germline knockout embryos was normal at e18.5, however no live *Ash1l* null pups were present at birth (e18.5: $n = 77$, $P = 0.90$; p0: *Ash1l*$^{+/+}$ $n = 41$, $P = 0.00095$). *Ash1l*$^{-/-}$ had shortened nasal bones at e18.5 ($n = 31$, $P = 0.017$). In the cortical-specific knockout, SATB2 neurons were increased and

distributed through the cortical plate (n = 6/genotype, P = 0.0001). There were no differences in neuronal birthdating (n = 4-6/genotype, P = 0.40 for e14.5 and P = 0.057 for e15.5). Single cell RNA sequencing revealed numerous differences in gene expression that were sufficient to cluster control and mutant upper layer neurons separately. Pseudotime analysis revealed that the mutant cluster is on an altered cell differentiation trajectory.

2.2 Introduction

During mammalian brain development, the full complement of mature cortical neuron subtypes are born from multipotent neural progenitor cells (NPCs) [167–170]. Transcriptional plasticity is required to generate this cellular diversity. Disruptions in the complex process of gene regulation can impact cellular identity and function [171,172]. This is acutely evident by the large number of genes that encode transcriptional machinery and chromatin components that are the genetic basis of neurodevelopmental disorders (NDDs) [99]. One of these genes, *ASH1L* (Absent, Small, Homeotic-Like), encodes a histone methyltransferase that dimethylates lysine 36 of histone 3 (H3K36me2), a mark associated with gene activation [62]. *ASH1L* is broadly expressed, with the highest levels in the brain [146]. *De novo* variants in *ASH1L* are associated with syndromic intellectual disability (ID) and autism spectrum disorder (ASD), further highlighting an important role in brain development.

Postmortem tissue analysis of individuals with NDDs, particularly ASD, have identified key hallmarks, specifically aberrant mature cortical neuron organization, neuronal morphology, and cortical neuron connections [173]. The cortex is the outermost layer of the cerebral hemispheres, and has been associated with language and memory [174,175]. The generation of the cortex is a complex process that is governed by transcriptional changes; multipotent neural stem cells undergo proliferation and differentiation stages. This gives rise to the vast array of postmitotic neurons arranged in the mammalian-conserved six cortical layers. The first-born cortical neurons form the deepest cortical layer, and as cortical lamination continues, layer V through to II are generated in a sequential manner, each with their own neuronal cell type, axon

projections and gene expression [176]. Transcriptional aberrations that disrupt this process, including cell cycle exit and cellular differentiation, can have seismic effects on corticogenesis, diverging from the normal temporal sequence to an asynchronous phenotype. This deviation can alter cell trajectory and compromise its identity, even resembling a different cell type. Transformation of cellular identities has been characterized in a number of models of various NDD-related genes that govern genetic regulation, suggesting a paradigm of altered cell identity converging on ASD-like cortical phenotypes [177–179]. Transcriptomic profiles are directed by precise chromatin modifications, in turn driving NPCs and immature postmitotic neurons to their correct mature cortical neuron fate; however, how this is coordinated remains an area of open investigation.

ASH1L is a member of the Trithorax group (TrxG) of chromatin regulators. TrxG counteract the transcriptionally repressive Polycomb group (PcG) complexes [180]; this evolutionary conserved relationship is essential for many aspects of metazoan development, including the brain [13]. Their importance in neurodevelopment is underscored by overlapping clinical phenotypes associated with TrxG and PcG variants. Through the use of model systems, convergent mechanisms of differential expression which underlie similar phenotypic outcomes are being identified.

Ash1l mouse models are valuable for revealing the essential functions of this gene in development. Via various loss of function (LOF) alleles, it has been found that global homozygous knockout (KO) of *Ash1l* results in early lethality [111,133]. Furthermore, homozygous *Ash1l* genetrap mice ($Ash1l^{GT/GT}$) exhibit reduced size, poor viability at weaning and infertility [136]. With regards to the brain and behavior,

heterozygous *Ash1l* mice (*Ash1l*$^{+/-}$) present with tic-like behaviors associated with Tourette syndrome [112]. *Ash1l*$^{+/-}$ and conditional knockout using Nestin-cre exhibit autistic-like behaviors, including heightened anxiety and reduced sociability, likely due to neural hyperactivity [105,230]. Adult mice who have been subjected to *Ash1l* knockdown via short hairpin RNA (shRNA) in the prefrontal cortex have fatal seizures [22]. Together, these studies indicate *Ash1l* plays essential roles in the brain, including modulating behavior.

The underlying developmental mechanisms of *Ash1l* action during the course of cortical development are poorly understood. We hypothesize that ASH1L-dependent differential gene expression alters neural development and causes molecular abnormalities in corticogenesis. To better understand the role of *Ash1l* in development, we generated two *Ash1l* mouse lines. The first was an *Ash1l* germline knockout, *Ash1l*$^{-/-}$, which exhibited craniofacial anomalies and perinatal lethality. This recapitulated previously reported features and confirmed phenotypes associated with *Ash1l* LOF [133]. The second *Ash1l* mouse line was a tamoxifen-inducible, conditional knockout model, *Emx1-Cre-ER*^{T2}*;Ash1l*$^{fl/fl}$ (*Ash1l*^{cKO}). Ablation of *Ash1l* in excitatory neurons of the developing cortex at e10.5 resulted in a larger number of SATB2 neurons at e18.5, and broader distribution of them through the cortical plate. Single cell RNA sequencing (scRNA-seq) revealed broad gene expression differences in layer II-VI neurons, and pseudotime analysis affirmed asynchronous differentiation trajectory for cortical neurons.

2.3 Materials and Methods

2.3.1 Mouse Generation

All procedures using mice were approved by the University of Michigan Committee on Use and Care of Animals, and all experiments were conducted in accordance with the principles and procedures outlined in the National Institutes of Health Guidelines of the Care and Use of Experimental Animals.

The *Ash1l*$^{tm2a(EUCOMM)Hmgu}$ allele was generated in C57BL/6N embryonic stem cells as part of the Knockout Mouse Project using the knock-out first targeting vector [181]. *Ash1l*^{tm2a} mice were purchased from EUCOMM (HEPD0792_8_C04, EUCOMM). This *Ash1l* allele has FRT sites flanking a lacZ and *neo* cassette within intron 3 of *Ash1l* and loxP sites flanking exon 4. Excision of exon 4 with *cre* recombinase will result in a frameshift and premature termination, *ASH1L* p.V1693Afs*2, a presumptive null allele. We mated the *Ash1l*$^{tm2a/+}$ mice with *Flp*^{O} transgenic mice purchased from the MMRRC [B6N(B6J)-Tg(CAG-Flpo)1Afst/Mmucd)], stock #036512-UCD to generate *Ash1l*$^{fl/+}$ mice [182]. Mice were crossed to C57BL/6J and genotyped as described to eliminate the *Flp*^{O} transgene and to detect the *Ash1l* floxed allele (detailed below). To generate an *Ash1l* germline knockout line (*Ash1l*$^{-/-}$), *Ash1l*$^{fl/+}$ mice were crossed with (B6.FVB-Tg(EIIa-cre)C5379Lmgd/J strain #003724) [183]. Resulting *Ash1l*$^{fl/-}$*; EIIA-cre* mice were backcrossed with C57BL6/J and genotyped to select for *Ash1l*$^{+/-}$ mating pairs without *EIIA-cre*. To generate breeding pairs for cortical-specific knockout of *Ash1l* (*Ash1l*^{cKO}), we crossed *Emx1-Cre-ER*^{T2} (B6;CBA-Tg(Emx1-cre/ERT2)1Kess/SshiJ, Jackson laboratory strain #027784, and *Ash1l*$^{fl/fl}$ mice to produce *Emx1-Cre-ER*^{T2}*;Ash1l*$^{fl/fl}$ mice [184]. In turn, male *Emx1-Cre-ER*^{T2}*;Ash1l*$^{fl/fl}$ and female *Ash1l*$^{fl/fl}$ were paired for timed

pregnancies. The morning a copulation plug was detected was designated as embryonic day (e) 0.5. To induce *cre* expression, the pregnant females were injected with 100 mg/kg tamoxifen (Sigma Aldrich, Darmstadt, Germany; cat: T5648-1G, source: WXBD8755V) at e10.5 to coincide with neurogenesis of the cortex. Mice were housed in specific pathogen-free conditions in ventilated cages with automatic watering and fed Purina 5020 chow ad libitum.

2.3.2 Genotyping

DNA was extracted from tail tissue as previously described [185]. For *Ash1l* germline knockout genotyping, a multiplex PCR was used to distinguish the presence of exon 4 in one, both or neither *Ash1l* alleles. The F1 forward primer, 5'-GCTTCAAATGGGCAAGCCTC-3', within intron 3, and the R1 reverse primer, 5'-GCTCATCGTGATGTACTATCCGC-3', within exon 4, give a PCR product of 800 bp when exon 4 is present. The additional reverse primer R2, 5'-TAGGCCAGTCAGGCTACGAA-3', lies within intron 4. When exon 4 is absent, the PCR product is 450 bp. The PCR conditions are: 95°C for 5 min, followed by 35 cycles of 95°C for 30 sec, 56°C for 30 sec and 72°C for 30 sec, with a final extension of 72°C for 5 min. PCR products were separated by electrophoresis through a 2% agarose gel. For *Ash1l* cKO genotyping, PCRs were carried out to determine presence of the *Ash1l* 5' loxP site and 3' loxP site, and *Emx1-Cre-ER*^{T2}. 5' loxP forward primer 5'-GTCAGTATAGAAAGGCATGG-3' and reverse primer 5'-ATGGAGCCAGAAGAGGGAGTC-3'. 3' loxP site forward primer 5'-CAGCAAGCATCTTTCCTGGTGACC-3' and reverse primer 5'-ATGGCGAGCTCAGACCATAAC-3'. The PCR conditions for both are: 95°C 5 min,

followed by 35 cycles of 95°C 30 sec, 58°C 30 sec and 72°C 30 sec, with a final extension of 72°C 5 min. The presence of *Emx1-Cre-ER*^{T2} was determined with forward primer 5'-AGAGACTTCAGGGTGCTGGA-3' and reverse primer 5'-GTGCCTGGCTAGAGATCCTG-3'. PCR conditions are: 94°C for 2 min, followed by 30 cycles of 94°C for 30 sec and 54°C for 15 sec, 72°C for 60 sec, followed by a final extension of 68°C for 5 min. PCR products were separated by electrophoresis through a 1% agarose gel.

2.3.3 Tissue Preparation

Pregnant dams were euthanized, and the embryos were removed. Embryo weights were recorded. Whole brains were dissected and weighed, before being fixed in 4% paraformaldehyde overnight at 4°C. Brains were washed in 1xPBS and washed with 10%, 20% and 30% sucrose in 1xPBS solution until tissue sank to bottom of tube. The brains were then washed in a 1:1 30% sucrose, Tissue Plus O.C.T. Compound (Fisher Healthcare, TX, USA; cat: 4585) mix and then transferred to 100% O.C.T. Brains were orientated in O.C.T. before freezing on dry ice and subsequent storage at -80°C. Brains were sectioned on a Leica CM 1950 cryostat at 15µm thickness and mounted onto Super Frost Plus microscope slides (Fisher Scientific, NH, USA; cat: 12-550-15). Slides were stored at -80°C until needed.

2.3.4 Skeletal Preparations

Skeleton preparation was carried out as previously described [186]. In brief, skin, organs and excess tissue were removed from the embryos before submerging in 95% ethanol for three days. Embryos were then added to staining solution (0.3% alcian blue,

acetic acid and 95% ethanol) for three days and dehydrated for a further five days in 95% ethanol. Embryos were then cleared with 1% potassium hydroxide and counterstained (0.1% alizarin red, 1% potassium hydroxide). Skeletons were brought through gradually increasing glycerol concentration solution and finally stored in 100% glycerol.

2.3.5 Immunohistochemistry

Slides with cryosectioned tissue were thawed and washed three times in 1xPBS. Slides were boiled in 10mM critic acid for 10 min. After allowing slides to cool for 1h, they were washed three times in 1xPBS. Sections were blocked with normal donkey serum (Jackson Immunoresearch Laboratories, PA, USA; cat: 017-000-121) or Mouse on Mouse (M.O.M.) diluent (Vector Laboratories, CA, USA; cat: BMK-2202) for 1h at room temperature. Primary antibody diluted in blocking solution was applied to sections for overnight incubation at 4°C. Slides were washed three times in 1xPBS, and then incubated with 1:100 primary host specific biotin for 1h at room temperature. Slides were subsequently washed and incubated with 1:100 fluorescent-conjugated streptavidin (source details below) for 1h at room temperature. Slides were washed with 1xPBS and stained with DAPI (1:200) for 5 min. Slides underwent additional 1xPBS washes and were mounted with Prolong Gold antifade reagent (Invitrogen, OR, USA; cat: P36930) and coverslips applied (Epredia, MI, USA; cat: 152440). The following primary antibodies were used (all from Abcam, MA, USA): mouse anti-SATB2 (1:100; cat: Ab51502, lot: GR70015-13), rat anti-BCL11B (1:500; cat: Ab18465, lot: GR322373-4). The following secondary antibodies were used (all from Jackson Immunoresearch Laboratories, PA, USA): Biotin SP conjugated F(ab')2 Sheep anti-Mouse IgG (cat: 515-

065-072), Biotin SP conjugated F(ab')2 Donkey anti-Rat IgG (cat: 712-066-153) and Biotin-SP (long spacer) AffiniPure F(ab')$_2$ Fragment Donkey Anti-Rabbit IgG (cat: 711-066-152).

2.3.6 Birthdating

Pregnant females were injected with 50 mg/kg IdU (Sigma Aldrich, Darmstadt, Germany; cat: I7125) on e14.5 and 50 mg/kg EdU (Baseclick, Munich, Germany; cat: BCK488-IV-IM-S)) on e15.5. On e18.5, embryonic brains are dissected and undergo tissue processing. EdU was visualized using EdU IV Imaging Kit 488 S (Baseclick, Munich, Germany) in accordance with the manufacturer's protocol. After incubating with the reaction cocktail, slides were washed with 1xPBS and blocked with normal donkey serum (Jackson Immunoresearch Laboratories, PA, USA; cat: 017-000-121). Immunohistochemistry was subsequently carried out with rat anti-BrdU (1:100; AbD Serotec/Bio-Rad, cat: MCA2060, lot: 0109) to detect IdU.

2.3.7 Imaging and Image Analysis

For skull imaging and measurements, heads were removed from skeletons and images of top, bottom and sides were taken on a Leica MZ 10 F microscope with a Leica DFC310 FX camera. The following dimensions of the skull images were measured with ImageJ [187]: total length (tip of the nasal bone to the top of the occipital bone), top width (across the widest part of the parietal bone), bottom width (across the widest part of the basisphenoid bone), height (from the base of the mandible to the junction between the parietal and frontal bones), interorbital distance (width of the

frontal bone at the level of the beginning of the eye socket) and nasal bone length. Both imaging and measurements were conducted without knowledge of genotype.

For birthdating and immunohistochemistry imaging and analysis, sections were imaged on a Nikon X1 Yokogawa spinning disk microscope set-up with Nikon Elements Advanced Research software. Medial and lateral portions of the cortices of both left and right hemispheres were imaged at 400x magnification in a 2x2 multi-tile fashion. 405nm, 475nm and 640nm wavelengths were captured. Z-stack steps were set to ±10μm center of tissue plane. Images were stitched and 100μm x 200μm regions of interest (ROIs) were defined from the apical surface of the cortex. ROIs were cropped and rotated upright. Element's 3D Spot Segmentation was used to mark bright spots greater than 6.3μm in diameter. Thresholds were adjusted based on signal strength and background noise per experiment. For birthdating, brightest spots were chosen, indicating neurons that exited the cell cycle shortly after DNA analogue injection. For distribution analysis, positive 3D spot segmentation points were converted to binary and imported to FIJI [188]. Images were split into ten equal bins and binary spots counted with the 3D Simple Segmentation plug-in [189]. 8 – 12 sections were imaged and counted per embryo.

2.3.8 Cell Collection for scRNA Sequencing

Cells were collected for scRNA sequencing as previously described [190]. In brief, embryonic cortices were dissected in Earle's balanced salt solution (EBSS) and incubated with papain, 5.5 mM L-cysteine-HCL, 1.1 mM EDTA, and 100 mg/ml DNase I (all part of Papain Dissociation System. Worthington Biochemical Corporation, NJ, USA; cat: LK003153) in O_2:CO_2 equilibrated EBSS for 8 min at 37°C. Samples were titrated

and centrifuged. Cells were resuspended with ovomucoid protease inhibitor and 50 mg/ml DNase I in O_2:CO_2 equilibrated EBSS and passed through a 70μm cell strainer (Becton Dickson, NJ, USA; cat: 08-771-2). Cells underwent centrifugation with an ovomucoid protease inhibitor gradient and then resuspended in N2/B27 media (all from Invitrogen, MA, USA. Neurobasal media cat: 21103-049; B27 supplement cat: 17504044; L-glutamine cat: 25030-081). Cells were counted before subsequent scRNA sequencing steps.

2.3.9 Seq-Well scRNA Sequencing, Processing, and Analysis

Single cell RNA sequencing was carried out using the Seq-Well platform, as previously described [191]. Briefly, functionalized Seq-Well arrays were loaded with mRNA Capture Beads (ChemeGenes, Wilmington, MA; cat: MACOSKO-2011-10). Cells were loaded on the arrays and incubated for 15 min. Functionalized membranes (custom ordered from Sterlitech, WA, USA) were sealed onto the arrays and incubated at 37°C for 45 min. Sealed arrays underwent incubations in lysis and hybridization buffers. Beads were removed from the arrays via centrifugation. Beads were incubated with Maxima Reverse Transcriptase (Thermo Fisher Scientific, MA, USA; cat: EPO0753) overnight, then treated with Exonuclease 1 (New England Biolabs, MA, USA; cat: M0293S). Transcriptome amplification was performed using the 2X KAPA Hifi Hotstart Readymix (Roche, Basel, Switzerland; cat: KK-2602). Products were purified with AMPure SPRI beads (Beckman Coulter, CA, USA; cat: A63881). Libraries were prepared using the Nextera XT DNA Library Preparation Kit (Illumina, CA, USA; cat: FC-131-1096) and libraries were sequenced on an Illumina NextSeq 2000 instrument (75 cycles). Digital gene expression matrices were generated from sequencing data

using the Drop-seq Alignment protocol, as previously described [192]. Reads were aligned to mm10 with STAR [193]. Filtering and analysis were carried out using Seurat R package (versions 3.1.2 - 4.3) [194,195]. Cells with transcripts from fewer than 300 genes or more than 2000 genes, and cells with greater than 10% mitochondrial gene transcripts were removed from subsequent clustering and analysis. Data was normalized, scaled, and underwent dimensionality reduction. Clusters representing cell types were generated and annotated using previously described markers of cortical cells [196,197]. Excitatory neurons were isolated and re-clustered using reciprocal PCA. GSEA analysis was carried out using the R package Enrichr [198,199]. Pseudotime analysis was conducted with Monocle3 [200–203]. GO term analysis figures were generated with ShinyGO [204].

2.3.10 Statistics

GraphPad Prism software was used for statistical analysis and making graphs (version 10.0.3). The *P* value for analyzing *Ash1l*$^{-/-}$ lethality was calculated using the chi-square test. *P* values for skull dimension and brain/body weight comparisons were calculated using the analysis of variance (ANOVA) test. Comparing *Ash1l*^{cKO} and *Ash1l*^{Ctrl} data (raw numbers, IHC percentages, scRNA-seq cell type populations) were statistically analyzed using the unpaired t-test. Specific parameters for t-tests are reported in figure legends.

2.4 Results

2.4.1 Ash1l Germline Knockout Mice Exhibit Perinatal Lethality and Variable Craniofacial Anomalies.

We showed that mice homozygous for an *Ash1l* gene trap allele (*Ash1l*$^{GT/GT}$) had residual *Ash1l* expression, reduced body weight, and poor viability at weaning [136]. To assess how total ablation of *Ash1l* expression impacts survival we generated a null allele lacking exon 4, which produces a frame shift and premature termination, *Ash1l* p.V1693Afs*2 **(Fig. 4A, top panel**). We used a multiplex PCR approach for genotyping *Ash1l*$^{+/+}$, *Ash1l*$^{+/-}$ and *Ash1l*$^{-/-}$ mice (**Fig. 4A, bottom panel**). *Ash1l*$^{+/-}$ males and females were crossed for timed pregnancies and embryos were collected at embryonic day (e) 18.5. At e18.5, *Ash1l*$^{+/+}$, *Ash1l*$^{+/-}$ and *Ash1l*$^{-/-}$ embryos were present in a normal Mendelian distribution (1:2:1) (*Ash1l*$^{+/+}$ *n* = 21, *Ash1l*$^{+/-}$ *n* = 37 and *Ash1l*$^{-/-}$ *n* = 19; chi square *P* value (two-tailed) = 0.8955). However, at postnatal day (p) 0, no live *Ash1l*$^{-/-}$ pups were identified, whereas *Ash1l*$^{+/+}$ and *Ash1l*$^{+/-}$ pups were present and alive (*Ash1l*$^{+/+}$ *n* = 15, *Ash1l*$^{+/-}$ *n* = 26 and *Ash1l*$^{-/-}$ *n* = 0; chi-square *P* value (two-tailed) = 0.0009). We found no difference in the expected Mendelian ratio in the number of *Ash1l*$^{+/+}$ and *Ash1l*$^{+/-}$ mice at p21 (1:2) (*Ash1l*$^{+/+}$ *n* = 20 and *Ash1l*$^{+/-}$ *n* = 30; chi-square *P* value (two-tailed) = 0.3679) (**Fig. 4B**). This data shows that homozygous *Ash1l* LOF results in lethality, specifically between e18.5 and birth, demonstrating that *Ash1l* is essential for survival.

Ash1l is a known regulator of *Hox* gene expression [23,136,162]. *Hox* genes are essential for correct body segmentation and development [205]. We used the DECIPHER clinical database to assess the prevalence of skeletal and craniofacial

anomalies in individuals with a heterozygous *ASH1L* variant. Of the 42 patients with clinical information, 50% had documented skeletal anomalies, and 47.6% had craniofacial dysmorphology [5]. Therefore, we hypothesized that *Ash1l*$^{-/-}$ mice may exhibit skeletal and craniofacial defects. Embryos were collected at e18.5 and examined. Some *Ash1l*$^{-/-}$ embryos had visibly shorter snouts (**Fig. 4C**). We quantified differences in skull dimensions following staining with alcian blue and alizarin red (**Fig. 4D**). The nasal bone length was significantly shorter in *Ash1l*$^{-/-}$ embryos compared to *Ash1l*$^{+/+}$ and *Ash1l*$^{+/-}$ embryos (*Ash1l*$^{+/+}$ n = 6, *Ash1l*$^{+/-}$ n = 11 and *Ash1l*$^{-/-}$ n = 14; one-way ANOVA P value = 0.0171). The skull height, width and length were not significantly different (**Supplemental Fig. 1B-I**), suggesting a specific role for *Ash1l* in craniofacial development.

Differences in brain size, particularly macrocephaly, have long had an association with ASD [206,207]. To determine whether *Ash1l* LOF resulted in differences in brain size, we weighed bodies and brains of e18.5 embryos (**Fig. 4E**) and (**Supplemental Fig. 2A, B**). Differences in brain-to-body weight percentage between genotypes at e18.5 (*Ash1l*$^{+/+}$ n = 11, *Ash1l*$^{+/-}$ n = 22 and *Ash1l*$^{-/-}$ n = 9; one-way ANOVA P value = 0.7927) were not statistically significant. We compared gross morphology of *Ash1l*$^{+/+}$, *Ash1l*$^{+/-}$, and *Ash1l*$^{-/-}$ brains. No major anatomical differences were found. The ventricular area of brain sections were measured and normalized to total brain area; no significant difference was observed (**Supplemental Fig. 2C-E**). Together, this data suggests that *Ash1l* deficiency has little or no effect on overall brain growth during prenatal development.

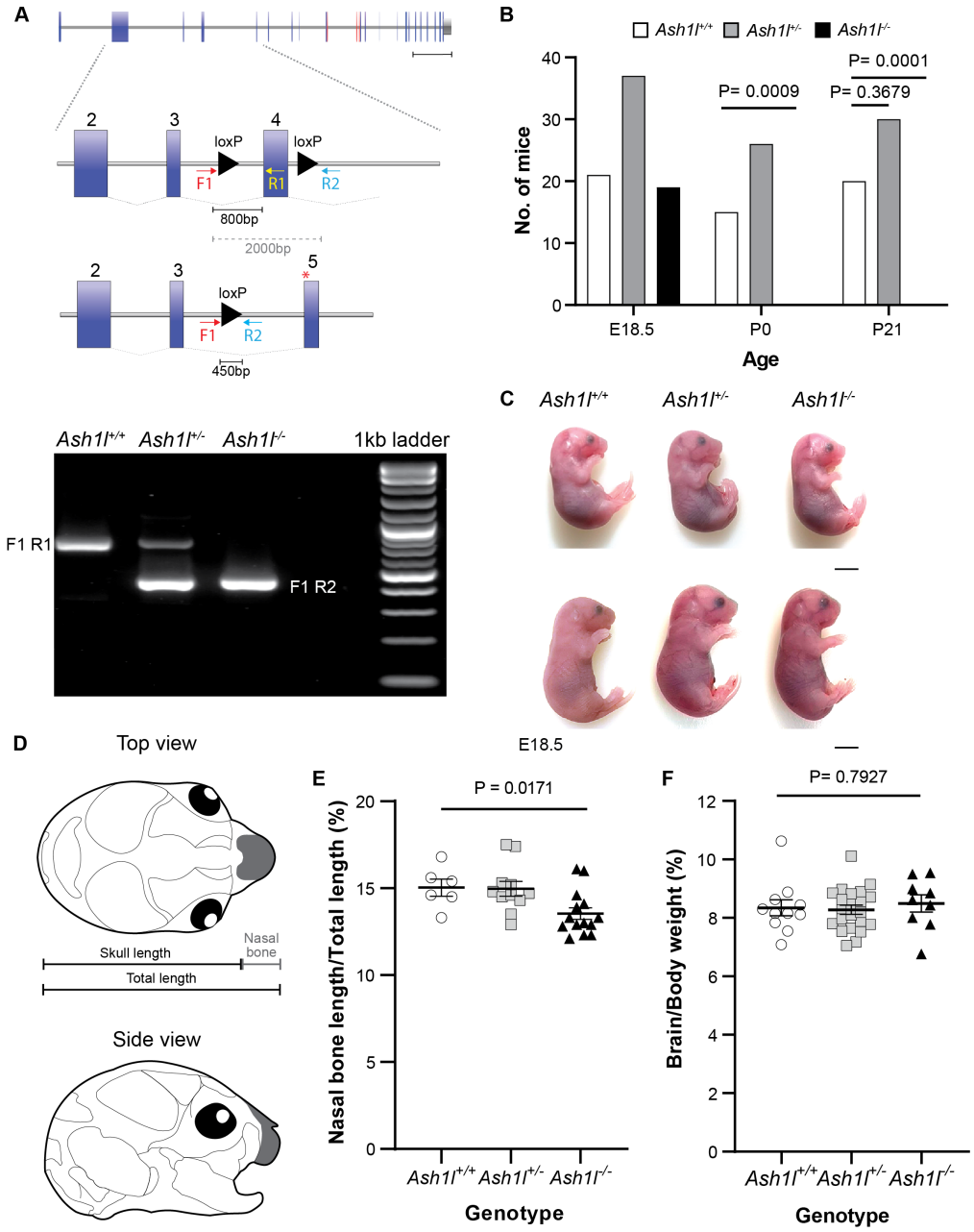


Figure 4: *Ash1l* germline KO mice exhibit perinatal lethality and craniofacial anomalies.

(**A**) Gene structure of *Ash1l*. Scale bar = 11.4 kb. Exons 2-5 are enlarged and loxP sites flanking exon 4 are shown as black arrowheads. SET domain is encoded by exons in red. In the presence of cre, exon 4 is excised and recombination occurs, resulting in a premature stop codon in exon 5 (denoted with red *). Multiplex genotyping is carried out with primers F1, R1 and R2, and results for each genotype are shown in the bottom panel. (**B**) Number of viable mice at e18.5, p0 and p21 of each genotype *Ash1l*$^{+/+}$ (white), *Ash1l*$^{+/-}$ (grey) and *Ash1l*$^{-/-}$ (black) obtained by intercrossing *Ash1l*$^{+/-}$ mice. *P* values were calculated for each timepoint by chi-square. E18.5, *n* = 77; p0, *n* = 41; p21, *n* = 50. (**C**) Images of e18.5 embryos, *Ash1l*$^{+/+}$ (left), *Ash1l*$^{+/-}$ (middle) and *Ash1l*$^{-/-}$ (right). Scale bar = 5mm. (**D**) Schematic of top and side views of e18.5 mouse skull. Nasal bone in dark grey. (**E**) Percentage of nasal bone length that comprises the total skull length (total skull length = skull length + nasal bone length) compared across genotypes at e18.5. *P* value calculated by one-way ANOVA. *Ash1l*$^{+/+}$, *n* = 6; *Ash1l*$^{+/-}$, *n* = 11; *Ash1l*$^{-/-}$, *n* = 14. (**F**) Percentage of brain weight over total body weight compared across genotypes at e18.5. *P* value calculated by one-way ANOVA. *Ash1l*$^{+/+}$, *n* = 11; *Ash1l*$^{+/-}$, *n* = 22; *Ash1l*$^{-/-}$, *n* = 9. Data shown as mean ± SEM.

2.4.2 Ash1l^{cKO} Mice Have Increased SATB2 Cortical Neurons With Altered Layer Distribution.

As *Ash1l*$^{-/-}$ brains were generally comparable to controls, we postulated that *Ash1l* LOF may impact cytoarchitecture. Leading theories on ASD etiology implicate altered development of the cortex, the highly organized outer layer of the brain [208,209]. Additionally, anomalies in cortical organization have been identified in numerous mouse models of ASD [210,211]. For these reasons, we focused on the mouse cortex, and utilized a conditional knockout (cKO) mouse model, given the peri-lethality observed with *Ash1l*$^{-/-}$ mice. We used *Emx1-cre-ER*^{T2}, a tamoxifen inducible *cre* strain that is activated in pallial precursor cells [184]. *Ash1l*$^{fl/fl}$ mice were crossed with *Emx1-creER*^{T2}*,Ash1l*$^{fl/fl}$ to generate *Emx1-creER*^{T2}*,Ash1l*$^{fl/fl}$ (henceforth referred to as *Ash1l*^{cKO}) with *Ash1l*$^{fl/fl}$; *cre*-negative control littermates (*Ash1l*^{Ctrl}). The *Emx1-creER*^{T2} alone does not affect cortical development and has been used as the control line in a number of studies [212–218]. To coincide with the onset of corticogenesis, pregnant females were injected with tamoxifen at e10.5 to induce *cre* expression in *Emx1*-expressing cells. We dissected brains at e18.5, as mouse cortical neurogenesis is largely complete by then (**Fig. 5A**) [219]. To assess whether *Ash1l*^{cKO} embryos exhibited alterations in cortical lamination, we performed immunohistochemistry (IHC) on *Ash1l*^{cKO} and *Ash1l*^{Ctrl} cortices. We performed IHC staining with SATB2, a molecular marker for upper layer neurons II-IV, and BCL11B, a marker for deep layer V neurons [220–224]. Cortical tissue was imaged, regions of interest were selected and uniformly cropped (100μm x 200μm areas, from the basal surface of the cortical plate and extending downwards) for analysis. White boxes in **Fig. 5A** indicate examples of

regions sampled. We performed an assessment of cellular distribution by further dividing these regions into ten equally sized bins (**Fig. 5B**). Overall cell number ($DAPI^+$) was comparable between genotypes (**Supplemental Fig. 3A**). However, there were significantly more SATB2 neurons in *Ash1l*^{cKO} cortices compared to *Ash1l*^{Ctrl} (**Fig. 5C, D**) (*Ash1l*^{Ctrl} n = 6 and *Ash1l*^{cKO} n = 6; unpaired t-test P value = 0.0001). We normalized the number of SATB2 cortical cells to total cells (DAPI) per bin. As expected, *Ash1l*^{Ctrl} SATB2 neurons were more prevalent in the upper bins. Conversely, there was a stark difference in distribution of SATB2 neurons in *Ash1l*^{cKO} tissue. The increase in SATB2 labelled neurons were significantly increased in all ten bins. A significant difference in number of BCL11B neurons was not identified (**Fig. 5F, G**) (*Ash1l*^{Ctrl} n = 6 and *Ash1l*^{cKO} n = 6; unpaired t-test P value = 0.1317). The number of BCL11B neurons normalized to DAPI was comparable between *Ash1l*^{Ctrl} and *Ash1l*^{cKO}, with a slight shift in distribution resulting in a greater number of BCL11B neurons in bin nine in *Ash1l*^{cKO} (**Fig. 5H**). Together, these data show that conditional knockout of *Ash1l* in the developing cortex results in more SATB2 neurons that ectopically reside in deep layers of the cortex, whereas BCL11B neurons remain comparable between *Ash1l*^{Ctrl} and *Ash1l*^{cKO} cortices, at e18.5. *Ash1l* may be necessary to refine the cell fate of upper layer neurons during cortical development.

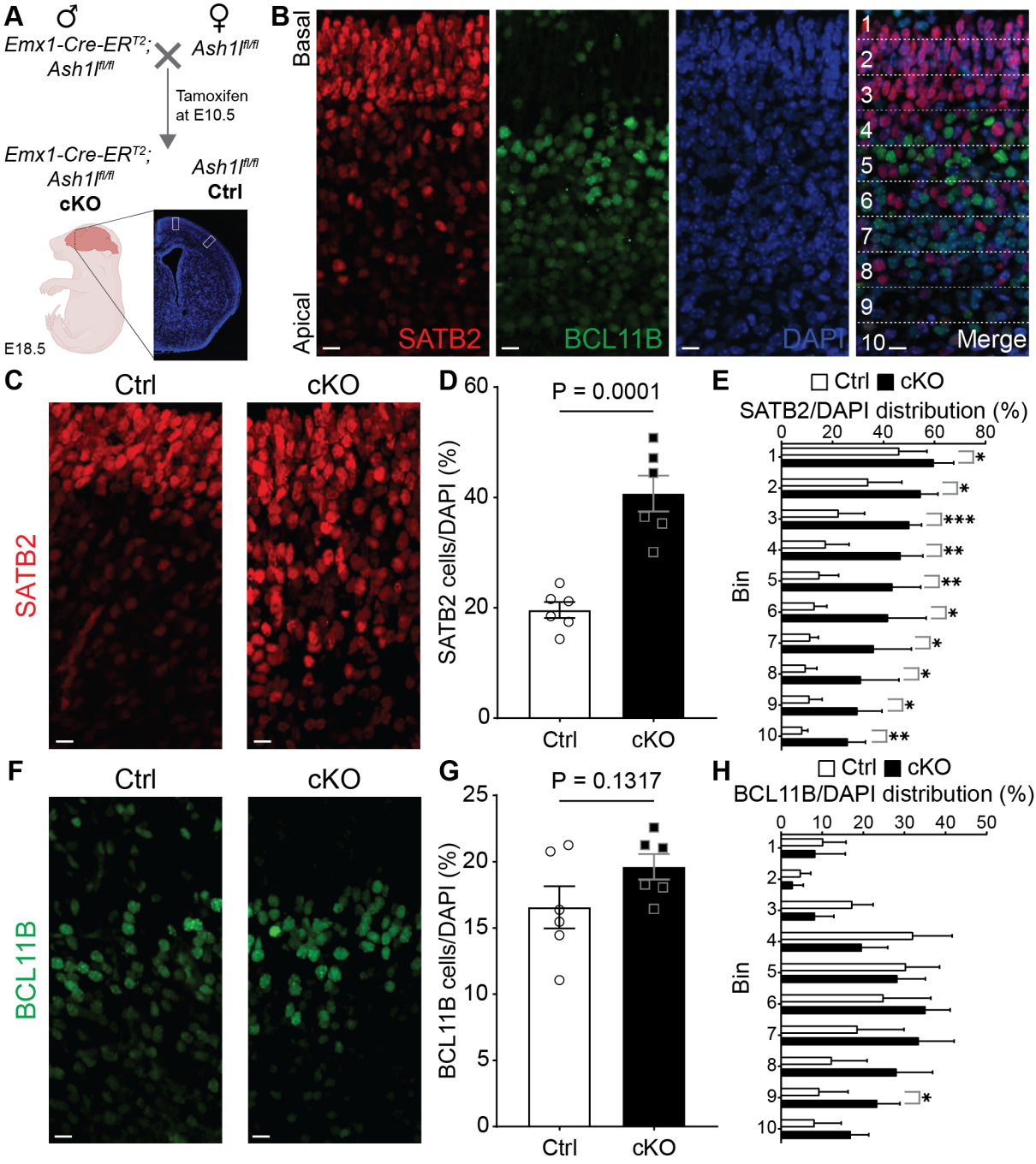


Figure 5: *Ash1l*^{cKO} have increased number of SATB2 cells and altered SATB2 distribution.

(**A**) Representative image of *Ash1l*^{Ctrl} brain section with regions of interest enclosed in white boxes (100μm x 200μm) (**B**) (left to right) SATB2 (layer II-IV marker), BCL11B (layer V marker), DAPI, and merged image. Merged image has ten equal sized bins (1-10) for distribution analysis. (**C**) Representative images of SATB2 IHC *Ash1l*^{Ctrl} and *Ash1l*^{cKO}. (**D**) Percentage of SATB2 cells over total number of cells. (**E**) Distribution of SATB2/DAPI (%) across bins. (**F**) Representative images of BCL11B IHC *Ash1l*^{Ctrl} and *Ash1l*^{cKO}. (**G**) Percentage of BCL11B cells over total number of cells. (**H**) Distribution of BCL11B/DAPI (%) across bins. Statistical analysis of (**D, G**) was determined by two-tailed unpaired t test. Significance shown by *P* values. (**E, H**) statistical analysis was determined by multiple unpaired t tests (Holm-Šídák method). Adjusted *P* values denoted (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$). Data shown as mean ± SEM. White scale bars = 10μm.

2.4.3 Ash1l^{cKO} Cortices Do Not Exhibit Altered Birthdating.

To address the possibility that *Ash1l* is required to orchestrate the timing of cell-lineage differentiation from NPCs, we performed birthdating analysis. To assess a significant difference in SATB2 neurons, we focused on the pivotal timepoints e14.5 (coinciding with the end of deep layer neurogenesis) and e15.5 (coinciding with the start of upper layer neurogenesis). We set up timed pregnancies and injected pregnant females with IdU (5-iodo-2'-deoxyuridine) on e14.5, and EdU (5-ethynyl-2′-deoxyuridine) on e15.5 (**Fig. 6A**). Females were euthanized on e18.5 and embryonic cortices processed. As previously found with the lamination analysis, cell numbers were comparable between genotypes (**Supplemental Fig. 3B**). We observed no differences in the proportion of IdU neurons or EdU neurons when comparing genotypes (**Fig. 6B, C**) (IdU: *Ash1l*^{Ctrl} n = 4 and *Ash1l*^{cKO} n = 6; unpaired t-test P value = 0.3993. EdU: *Ash1l*^{Ctrl} n = 4 and *Ash1l*^{cKO} n = 6; unpaired t-test P value = 0.0565). We performed co-labelling IHC for SATB2 and EdU to determine if the additional SATB2 neurons in the lower bins (8-10) of the *Ash1l*^{cKO} cortices were born at e15.5 (**Fig. 6D**). The majority of SATB2/EdU double-labelled neurons were found in the upper bins of *Ash1l*^{Ctrl} and *Ash1l*^{cKO}. We observed significantly more SATB2/EdU neurons in bin one in *Ash1l*^{cKO}, while the lower nine bins were comparable between genotypes. Together, this data suggests that the majority of the *Ash1l*^{cKO} SATB2 neurons in the lower bins were born prior to e15.5.

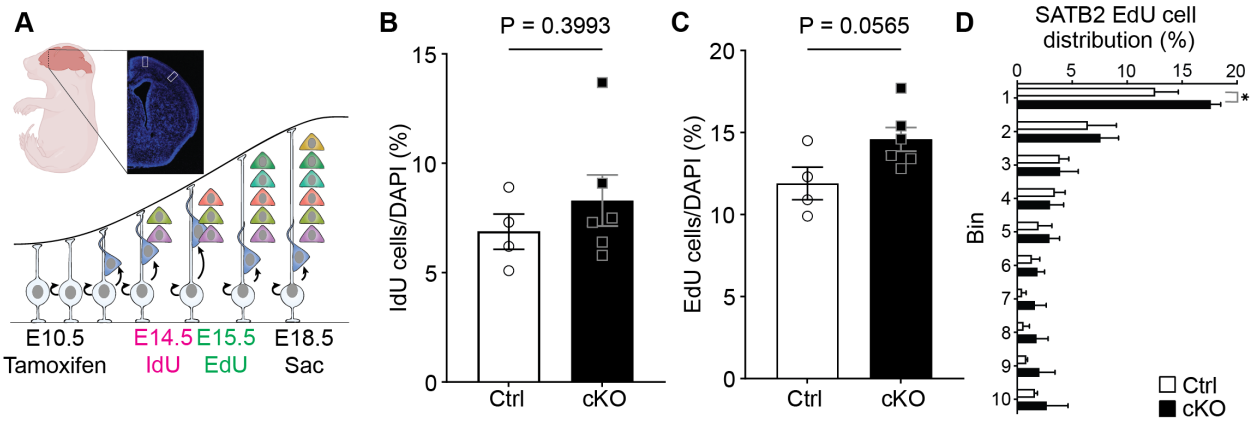


Figure 6: *Ash1l*^{cKO} cortices do not exhibit altered birthdating.

(**A**) Developmental timeline of corticogenesis e8.5 to e18.5. Experimental times shown on bottom axis. (**B**) Percentage of IdU cells over total number of cells. (**C**) Percentage of EdU cells over total number of cells. (**D**) Distribution as percentage of SATB2, EdU cells across ten bins (Percent of SATB2, EdU cells within a bin over total number of cells in 100μm x 200μm image). Statistical analysis of (**B, C**) was determined by two-tailed unpaired t test. (**D**) Statistical analysis was determined by multiple unpaired t tests (Holm-Šídák method). Adjusted *P* values denoted (* $P < 0.05$). Data shown as mean ± SEM.

2.4.4 Ash1l^{cKO} Cortices Have an Altered Population of Layer II-IV Neurons.

As a chromatin-based transcriptional regulator, *Ash1l* may be altering the expression of genes required for correct lineage fate specification in the developing cortex. The cortex is highly heterogeneous, so we used scRNA-seq to perform a broad and unbiased analysis of cell types. We collected cortices from three *Ash1l^{Ctrl}* and four *Ash1l^{cKO}* e18.5 embryos, and carried out scRNA-seq on the SeqWell platform for all seven samples (**Fig. 7A**) [191]. Before clustering, we assessed sample quality by comparing percentage of mitochondrial gene mapped, number of genes and number of transcripts detected (**Supplemental Fig. 4A**). There was remarkable consistency in sample quality. *Ash1l* transcripts were examined using the Integrative Genomics Viewer (IGV), and as expected, very few reads mapped to exon 4 (floxed in our allele) in *Ash1l^{cKO}* compared to *Ash1l^{Ctrl},* indicating consistent, efficient, tamoxifen-induced gene ablation (**Supplemental Fig. 4B**).

Excitatory neurons were identified based on gene expression, bioinformatically isolated and re-clustered (**Fig. 7B**). Cell clustering unveiled nine specific cell types across both genotypes. Select marker genes are shown in **Fig. 7C**. Eight out of the nine identified cell types have been well characterized and expected in e18.5 wild type mouse cortex: four layer II-IV neuron clusters, layer V subcerebral neurons, layer V-VI hippocampus callosal projection neurons (CPN), layer VI corticothalamic neurons and neural progenitor-like cells (**Fig. 7D**). Interestingly, we found an additional layer II-IV neuron cluster (named layer II-IV^{cKO}). This cell type aligned with the other layer II-IV cluster expression profiles (**Fig. 7C**). In *Ash1l^{cKO}* cortices, there was a significant decrease in layer II-VI neuron clusters (1) and (2), and a significant increase in layer II-

IV^{cKO} (**Fig. 7E, F**). This finding was consistent across individual samples (**Supplemental Fig. 4C**). These data suggest marked changes in gene expression in the cortex of *Ash1l*^{cKO} embryos, specifically in upper layer neurons II-IV.

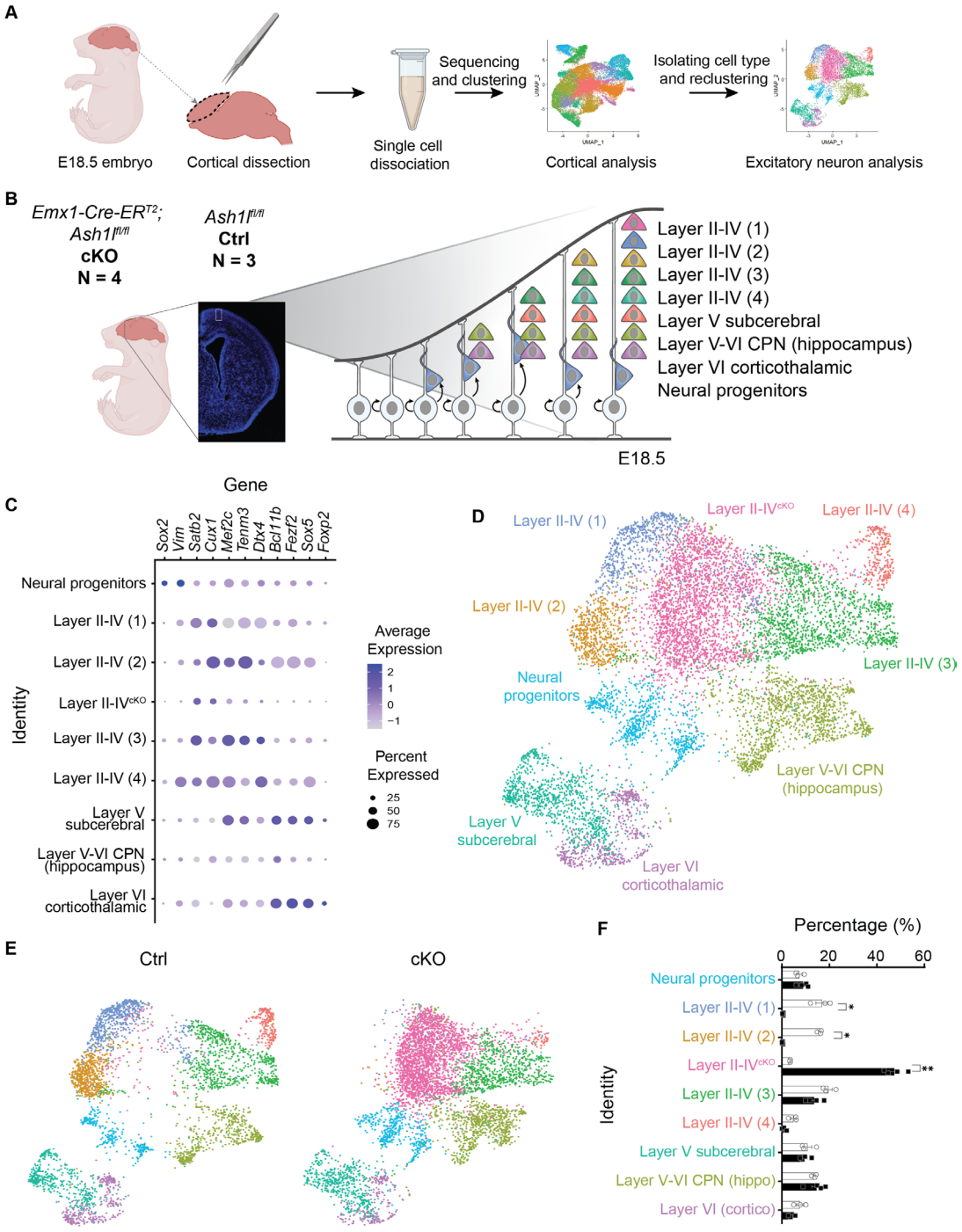


Figure 7: Single cell RNA sequencing unveils molecular differences in *Ash1l*^{cKO}.

(**A**) Experimental overview of scRNA-seq experiment following tamoxifen injection at e10.5. (**B**) Breakdown of e18.5 embryo genotypes used in scRNA-seq experiment. Developmental timeline of corticogenesis with annotated cell types at e18.5. Experimental times shown on bottom axis. (**C**) Dot plot of marker genes (x axis) by cell type (y axis). Dot size represents percentage of cells within a given cluster that express each marker, and color denotes expression level. (**D**) UMAP of *Ash1l*^{Ctrl} and *Ash1l*^{cKO} cell types identified. (**E**) UMAPs split by genotype; *Ash1l*^{Ctrl} and *Ash1l*^{cKO}. (**F**) Bar chart showing percentage of *Ash1l*^{Ctrl} and *Ash1l*^{cKO} cells corresponding to each cell type. Statistical analysis was determined by multiple unpaired t tests (Holm-Šídák method). Adjusted *P* values denoted (* $P < 0.05$, ** $P < 0.01$). Data shown as mean ± SEM.

2.4.5 ScRNA-seq Reveals Differential Expression of Altered Cell-Lineage Differentiation.

The size of the genotype-specific layer II-IV^{cKO} cluster correlates with the reciprocal reduction of clusters layer II-IV (1) and (2) in *Ash1l*^{cKO}. Thus, we first sought to identify the differential expression underlying these changes (**Fig. 8A**). We isolated these three clusters and performed reclustering to assess differentially expressed genes that may be driving the layer II-IV^{cKO} identity. Surprisingly, we found only small, but significant, changes in gene expression (**Fig. 8B)**. Upregulated genes in *Ash1l*^{cKO} included *Ndn* and *P-Rex1,* both of which have been associated with abnormal neuron development and activity. *Ndn* overexpression has been found to induce hyperexcitability in cortical neurons [225]. *P-Rex1* overexpression alters neuronal spine density and results in anxiety like behaviors in mice [226]. Interestingly, the most upregulated gene was *Malat1,* a lncRNA that binds to polycomb repressive complex 2 (PRC2) and recruits it to target loci [227]. We found increased expression of *Satb2,* which correlates with an increase in SATB2 neurons in *Ash1l*^{cKO} cortices by immunohistochemistry. Downregulated genes included synaptic associated genes *Cadm2* and *Homer2. Egr1,* another downregulated gene*,* has been extensively studied for its role in neural activity and memory in knockout mouse models [228,229]. Overall, the majority of the dysregulated genes found in *Ash1l*^{cKO} impact neuron development or activity.

We sought to determine what biological processes might be affected by dysregulated gene expression. Gene ontology (GO) term analysis showed that neurogenesis, neuron differentiation, and neuron development were impacted across

up- and downregulated genes (**Supplemental Fig. 5A, D, G**). To home in on more specific mechanisms, we performed gene set enrichment analysis (GSEA) on the differentially expressed genes (**Fig. 8C**). The most significant term for downregulated genes was chromatin remodeling (GO: 0006338), which was unsurprising given that *Ash1l* is a histone modifier. Interestingly, positive regulation of both transcription by RNA polymerase II (GO: 0045944) and DNA-templated transcription (GO: 0045893) were enriched in the *Ash1l*^{cKO} downregulated gene set, suggesting that loss of *Ash1l* results in broad, genome-wide transcriptional changes.

We postulated that layer II-IV clusters (1) and (2) had undergone a different cell fate in *Ash1l*^{cKO} due to altered gene expression, resulting in a unique transcriptomic identity found in the scRNA-seq data. We aimed to assess the cell trajectory of the layer II-IV^{cKO} neurons in the context of the other upper layer neurons layer II-IV clusters (1, 2, 3, 4). To this end, we bioinformatically isolated these clusters and included the neural progenitor cluster to aid in establishing a pseudotime root. These groups underwent reclustering (**Fig. 8D**). We found significant differences in cell numbers in most clusters between genotypes (**Fig. 8E**). We assessed *Sox2* and *Vim* expression to identify the progenitor-like cluster, which was then used as the root node for further pseudotime analysis (**Fig. 8F**). We utilized Monocle 3, a popular R package for cell trajectory analysis [200–203]. We found contiguous pseudotime paths with distinct branch points across the clusters (**Fig. 8G, left**). There was a clear, genotypic difference in pseudotime values when the UMAP was split by genotype (**Fig. 8G, right**). We next plotted pseudotime by density (scaled cell numbers) (**Fig. 8H**). *Ash1l*^{Ctrl} formed three distinct peaks: one at 0 (the remaining neural progenitor-like cells), and two at 12 and

14. These later peaks were driven by transcriptomics that represent mature layer II-IV neurons. *Ash1l*^{cKO} formed a less defined curve; like *Ash1l*^{Ctrl}, there was a peak at 0, however the large peak at 6 suggests a large cell population that have not reached their final fate. The smaller peak at 10 may suggest that *Ash1l*^{cKO} cells can proceed further along this trajectory. Together, this data indicate that *Ash1l* plays a role in establishing correct cell fate in the developing cortex by coordinating gene expression.

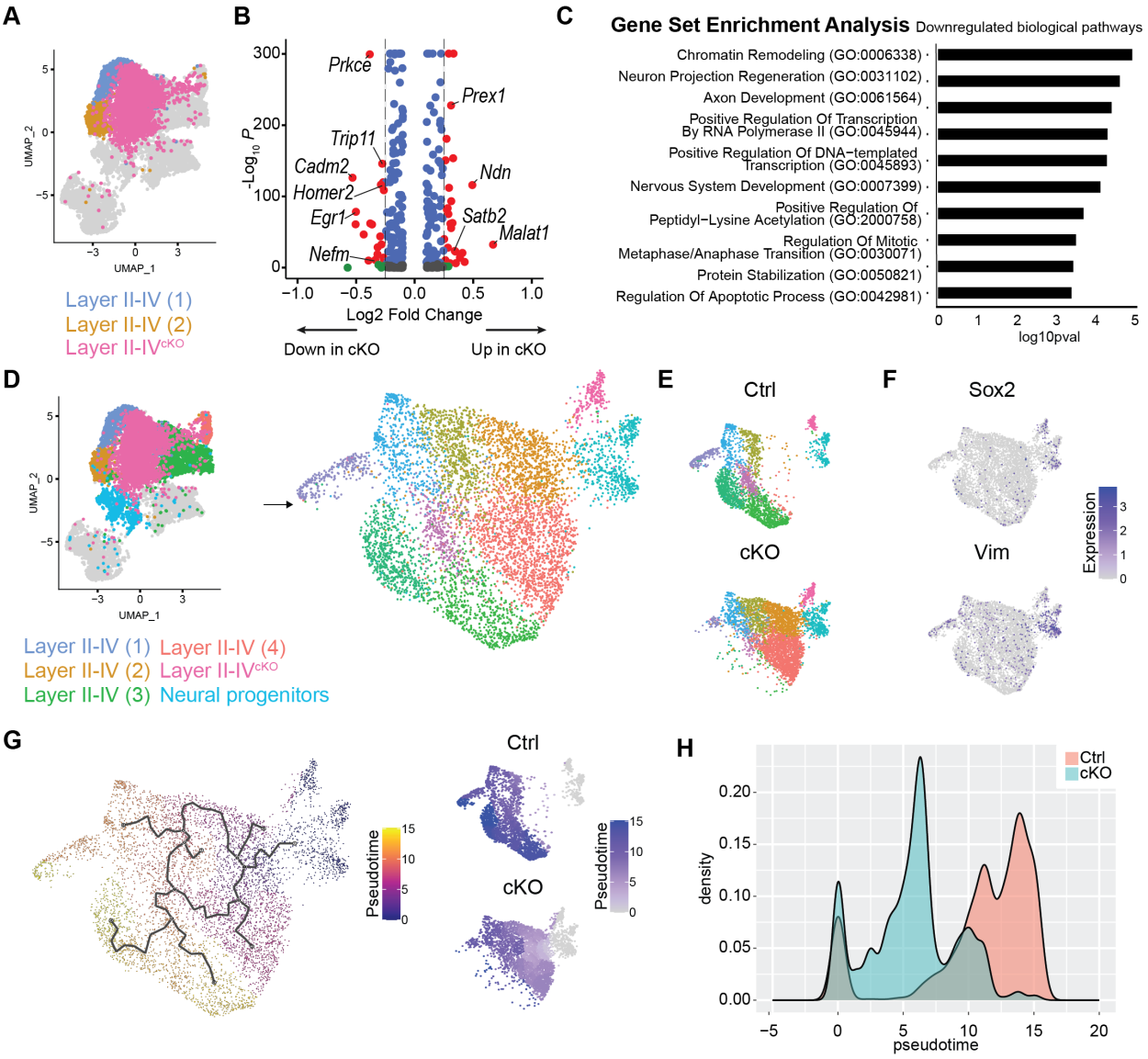


Figure 8: Analysis of scRNA-seq suggests altered cell trajectory due to differentially expressed genes.

(**A**) Upper layer neurons (I, II and cKO) were selected. (**B**) Volcano plot showing differentially expressed genes found between *Ash1l*^{Ctrl} and *Ash1l*^{cKO} neurons in these clusters. Some differentially expressed genes are labeled; upregulated genes in *Ash1l*^{cKO} on right and downregulated genes on left. X axis $\log_2$ fold change cutoff at 0.25. (**C**) Gene set enrichment analysis (GSEA) terms of biological processes that were negatively impacted in *Ash1l*^{cKO}. (**D**) Upper layer neurons II-IV (1, 2, 3, 4 and cKO) and neural progenitors were bioinformatically isolated and reclustered. UMAP shown on right. (**E**) UMAPs split by genotype. (**F**) *Sox2* (top) and *Vim* (bottom) mRNA expression; both are markers for neural progenitor-like cells. Darker blue represents higher expression. (**G**) Pseudotime analysis conducted with Monocle3. Gradient from navy blue to yellow represents increasing pseudotime. (Right) UMAPs split by genotype; darker blue represents higher pseudotime. (**H**) Density plot comparing distribution of cells along pseudotime axis. *Ash1l*^{Ctrl} in red and *Ash1l*^{cKO} in green.

2.5 Discussion

ASH1L has come under increased scrutiny in recent years, given the growing body of evidence that it is a high confidence ASD gene [105,230]. A number of *Ash1l* studies have investigated its role in the brain and behavior [22,24,111,112,133,135,230]. These studies have provided valuable insight into *Ash1l*-dependent neurophysiology and ASD-like behaviors. In this study, we set out to investigate the molecular consequences of *Ash1l* LOF in the developing mouse cortex. We confirmed that ablation of *Ash1l* in the germline causes perinatal lethality, and we documented specific craniofacial anomalies in e18.5 mutant embryos. There were no significant differences in brain weight relative to body weight. In the $Ash1^{cKO}$ cortices, there was a significant increase and ectopic placement of SATB2 neurons in the cortical plate. Birthdating analysis revealed that SATB2 neurons in deep layers of $Ash1l^{cKO}$ cortex were born prior to e15.5, like in the $Ash1l^{Ctrl}$. We compared cell populations in $Ash1l^{cKO}$ and $Ash1l^{Ctrl}$ via scRNA seq and found a layer II-IV cluster unique to mutants, in lieu of two distinct layer II-IV clusters found in controls. Further analysis indicated that cells in the layer II-IV^{cKO} cluster have gene dysregulation that drives its unique identity, and these cells have an altered cell-lineage fate trajectory.

Prior to birth, $Ash1l^{-/-}$ embryos were present at Mendelian ratios; however, no live $Ash1l^{-/-}$ pups persist at p0 in our germline *Ash1l* null line, due to *E2a-cre* mediated recombination of exon 4. Perinatal lethality is a conserved phenotype observed for a series of germline *Ash1l* LOF alleles, yet the cause of death is unknown. Highly penetrant neonatal lethality is observed in the CRISPR-mediated deletion, *Ash1l* c.1002_1013del^{GAATCAGGAAA}[111,133]. In the genetrap (GT) model, $Ash1l^{GT/GT}$ mice

showed reduced growth and postnatal lethality, with the surviving animals lacking Meibomian glands and exhibiting vertebral, reproductive organ, and fertility defects [136]. Lethality cannot be attributed to defects in brain development, as the *Emx1-cre* and *Nestin-cre* mediated deletion of *Ash1l* exon 4 produces viable mice. This highlights a systemic role for *Ash1l* in development that has implications for multiple organ involvement.

Human *ASH1L* heterozygous LOF variants are associated with a variety of structural birth defects, including some skeletal (50%) and craniofacial abnormalities (47.6%) and either macrocephaly (7.1%) or microcephaly (4.8%) [5]. There are also cases with heart, genital, and/or feeding abnormalities consistent with a syndromic developmental disorder. We did not detect any change in brain size or body weight of wild type, heterozygote or *Ash1l*$^{-/-}$ embryos at e18.5 However, e18.5 *Ash1l*$^{-/-}$ mice did reveal tissues vulnerable to *Ash1l* LOF during development. *Ash1l*$^{-/-}$ mice exhibited craniofacial anomalies, specifically a shorter nasal bone. The absence of craniofacial defects in the *Emx1-cre* (not expressed in neural crest) and presence in the *Nestin-cre* (expressed in neural crest) mice indicate that the *Ash1l* contribution of craniofacial abnormalities is through its impact on neural crest development [133]. Mice with craniofacial anomalies may have difficulty feeding after birth and die due to starvation [231,232]. This may explain perinatal lethality exhibited by *Ash1l*$^{-/-}$ animals; however, further investigation will needed to determine cause of death.

As a genetic etiology of ASD, efforts to reveal the role of *ASH1L* in brain development have intensified. Conditional *Ash1l* knockout during brain development reveals mammalian ASH1L neuropathology. We observed cortical lamination defects

that were similar to those observed in *Ash1l* mouse models [133]. We found a significant increase in the number of SATB2 neurons in *Ash1l*^{cKO} cortices at e18.5. We determined that this genotypic difference is due to a greater number of SATB2 neurons in deeper cortical layers of *Ash1l*^{cKO} brain. Immature postmitotic neurons destined to become upper cortical layer SATB2 neurons migrate from the ventricular zone past deep layers to contribute to upper cortical layers. Thus, ectopic deep layer SATB2 neurons can be attributed to altered timing of differentiation or migration of neurons. To tease apart these developmental mechanisms, IdU and EdU birthdating was performed in *Ash1l*^{cKO} and *Ash1l*^{Ctrl} at e14.5 and e15.5, the latter of which representing the onset of upper layer neuron development. No genotypic difference in the number or distribution of IdU or EdU neurons were observed. E15.5 EdU labelled neurons resided in upper cortical layers at e18.5, with a comparable number of SATB2 double positive cells observed between genotypes. This suggests that neurons generated at this developmental timepoint properly migrate to upper cortical layers. Importantly, SATB2 neurons that ectopically reside in deep layers were not EdU labelled, indicating that their cortical position was not due to migration defects, and is instead the result of premature differentiation. This finding is supported by the scRNA-seq transcriptomic data that revealed a unique cKO cluster in the upper cortical neuron UMAP space and asynchronous differentiation by pseudotime analysis. Cortical heterotopias and altered lamination are cytoarchitectural defects associated with ASD neuropathology, highlighting the relevance of these findings [208,209].

To explore asynchronous fate changes in *Ash1l*^{cKO} cortical development, we reclustered all layer II-IV clusters and neural progenitors and performed pseudotime

analysis. Interestingly, we did not see a simple shift in the pseudotime peaks depicted by *Ash1l*^{Ctrl}; we saw a build-up of *Ash1l*^{cKO} cells partially through the pseudotime axis. This data further supported our argument that upper layer cortical neurons in *Ash1l*^{cKO} cortices exhibit a distinct differentiation trajectory from that of control upper layer cortical neurons. This may represent convergent neuropathology between ASD risk genes, as asynchronous cell differentiation has been noted in a range of studies involving neural differentiation [233–237]. Paulsen *et al.* generated organoids harboring heterozygous LOF variants in one of three ASD risk genes, namely *SUV420H1*, *ARID1B* and *CHD8*. Using scRNA-seq at different time points, the authors found all three genes shared molecular phenotypes corresponding to asynchronous development. *ARID1B*$^{+/-}$ genotype resulted in delayed generation of deep layer projection neurons, whereas both *SUV420H1*$^{+/-}$ and *ARID1B*$^{+/-}$ genotypes had premature development of GABAergic neurons [237].

Through its chromatin associated activities, ASH1L is situated to alter transcriptional plasticity and cell fate. We directly compared the transcriptomic profiles of layer II-IV (1) and (2) with layer II-IV^{cKO}. Surprisingly, we found small, but significant, changes in gene expression that implicate pathogenic mechanisms of altered cortical neuron fate with implications for ASD biology. *Ndn, P-Rex1, Satb2,* and *Malat1* expression were upregulated in *Ash1l*^{cKO} neurons. Overexpression of *Ndn* and *P-Rex1* in mice causes neuronal hyperactivity and behavioral anomalies [225,226]. The lncRNA *Malat1* binds PRC2 and recruits it to target loci [227]. This is interesting given that ASH1L is thought to dismiss PRC2 in order to activate target gene expression. Various downregulated genes include *Cadm2* and *Homer,* which are involved in synapse

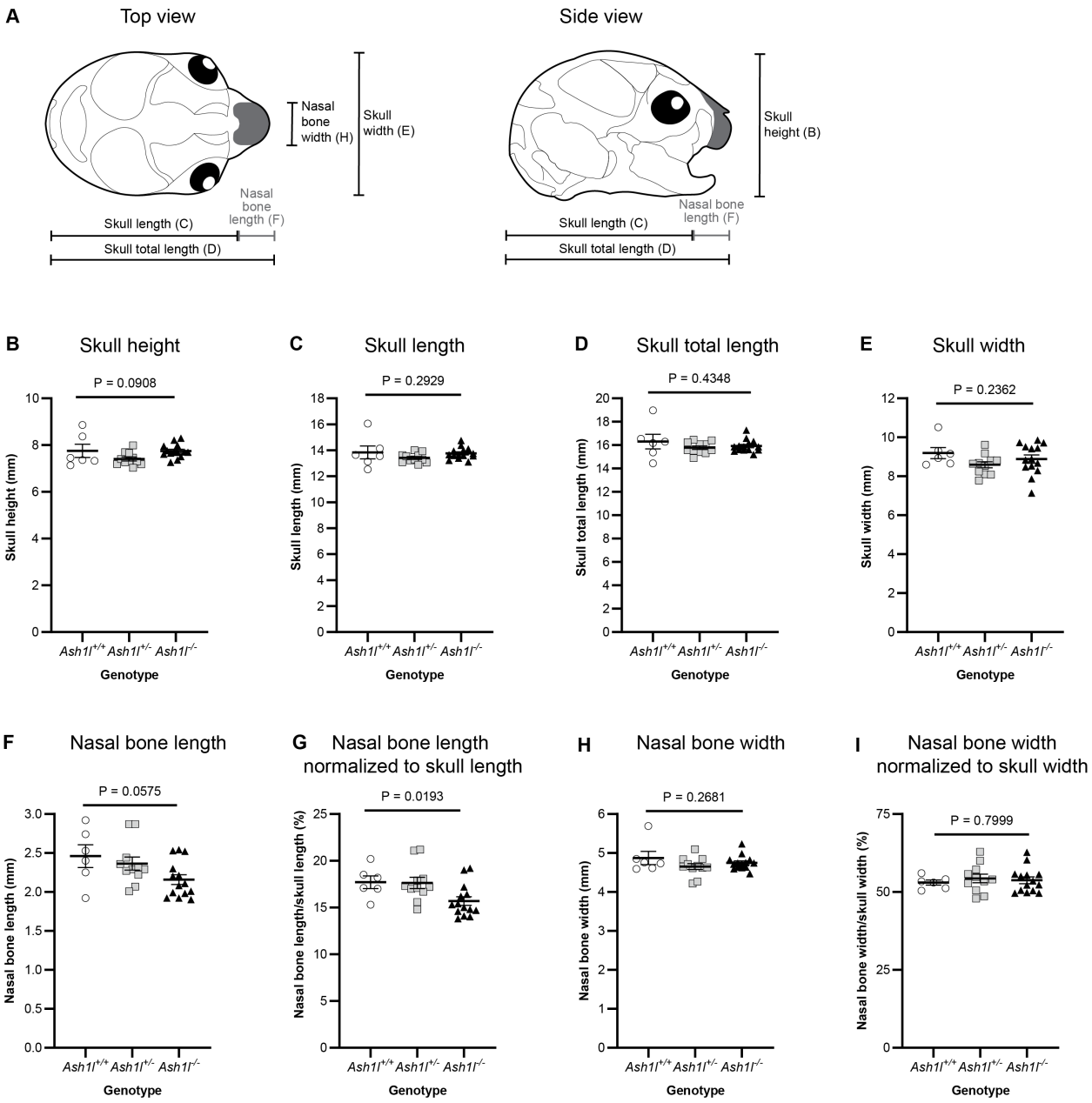
formation, and *Egr1,* prominently associated with neural activity and memory [228,229,238,239]. GSEA output suggested that downregulated genes found in the *Ash1l* layer II-IV^{cKO} played important roles in overall transcription; terms included positive regulation of DNA-templated transcription (GO:0045893), positive regulation of transcription by RNA polymerase II (GO:0045944), and chromatin remodeling (GO:0006338). Bulk RNA seq conducted by other groups studying *Ash1l* loss was assessed using GSEA or GO term analysis. We found overlap with a range of GO terms, including synapse organization, cell adhesion molecule binding and small GTPase binding [24].

Previous studies of *Ash1l* in the brain have employed bulk RNA-seq to determine differentially expressed genes. These models include *Ash1l-KO* primary neurons, knockdown of *Ash1l* via short hairpin RNA (shRNA) in the prefrontal cortex, and heterozygous *Ash1l* mice; subsequent GO term analysis implicate *Ash1l* serving broad neurogenesis roles [22,24,133]. ScRNA-seq analysis provides greater resolution to the cellular basis of these changes. We found most of the cell clusters overlapped in both genotypes, including those representing deeper layer neurons. However, clear genotypic differences were observed with the upper layer neuron clusters. The *Ash1l*^{cKO} had very few cells associated with the layer II-IV (1) and (2) clusters, whereas they contained the unique layer II-IV^{cKO} cluster. While these clusters are closely related based on their close association on the UMAP plot (**Fig. 7D**), their transcriptional profiles defined by upper cortical neuron transcripts were sufficiently different to generate distinct clusters. It is not evident how these transcriptional changes correlate to previously described functional changes. This finding also raises the likelihood that

the ectopic, deep layer SATB2 labelled neurons may not represent functionally normal upper layer neurons. In combination, these findings support a model where ASH1L is required for both the timing of differentiation and mature fate of cortical neurons that profoundly impacts upper layer cortical neuron heterogeneity. Interestingly, knockdown of *Ash1l* with shRNA was shown to alter the differentiation trajectory *in vitro* of mesenchymal stem cells, promoting adipogenic differentiation while suppressing osteogenic and chondrogenic differentiation, suggesting this may represent a broader developmental mechanism of ASH1L [140].

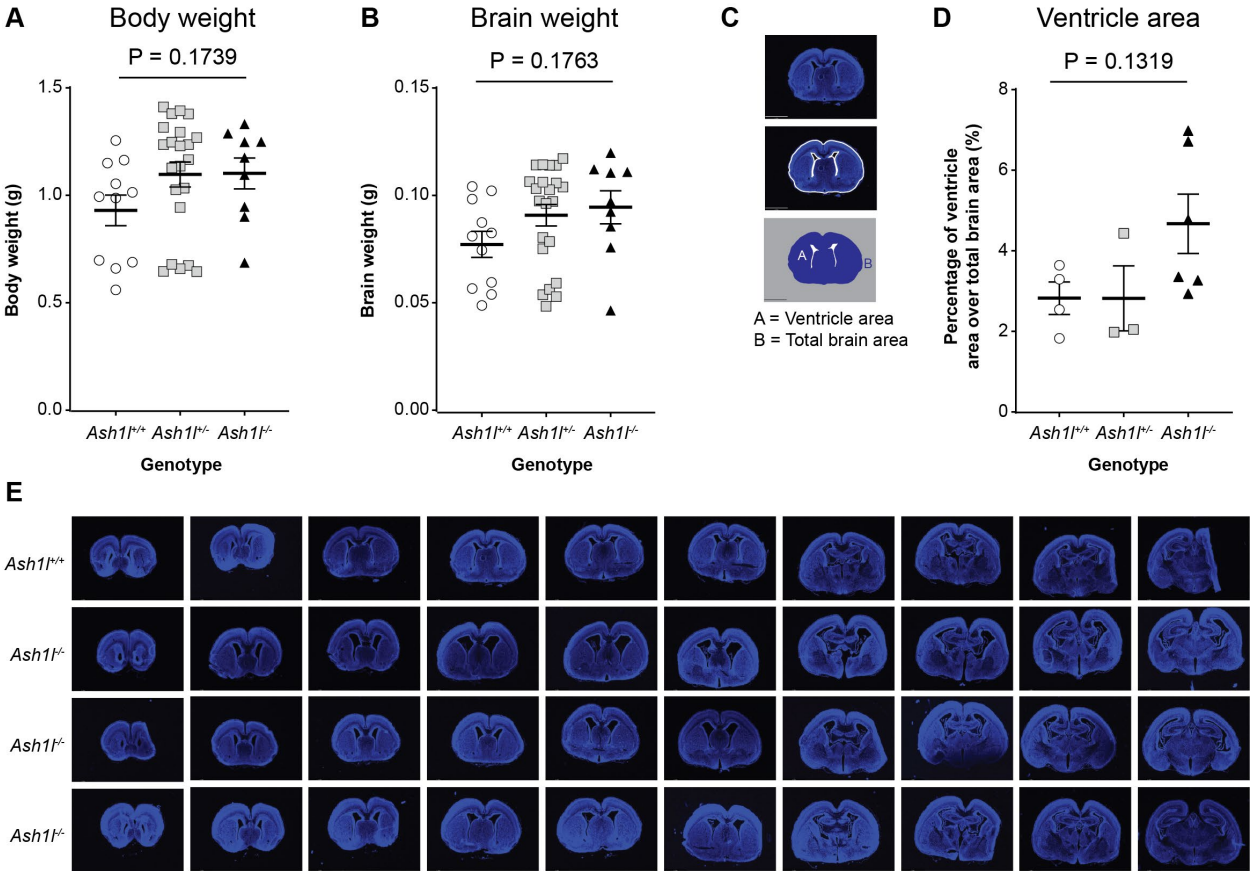
The underlying molecular etiologies of ASD remain elusive, however studying high risk genes, like *ASH1L*, on a gene-by-gene basis helps uncover overlapping mechanisms. One potential mechanism is that altered trajectories of cell fate may contribute to asynchronous differentiation. As a histone modifier, *Ash1l* regulates a broad range of genes that play roles in cell fate. This study provides evidence that *Ash1l* plays a critical role in mouse cortical development. This role may underlie some of the ASD features in individuals heterozygous for *ASH1L* LOF.

2.6 Supplemental Figures



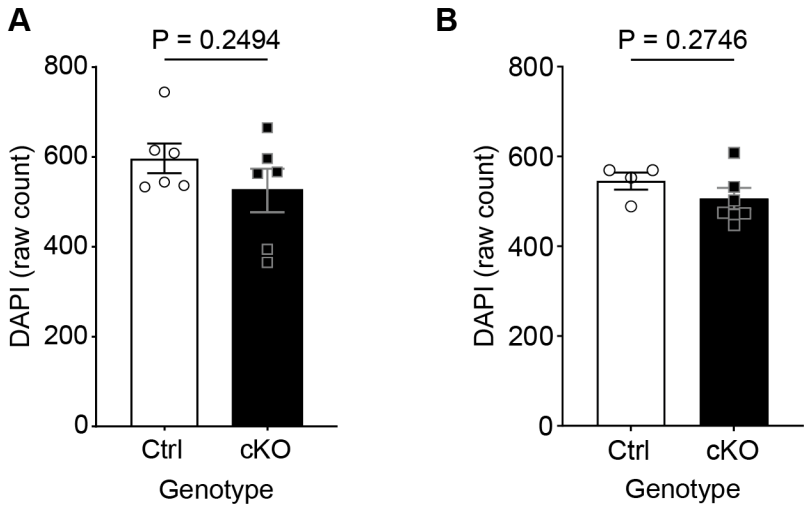
Supplemental figure 1: *Ash1l*$^{-/-}$ have a reduction in nasal bone length, but most other skull dimensions are normal.

(**A**) Schematic of top and side views of e18.5 mouse skull. Nasal bone in dark grey. Measurements denoted with black and grey lines. Letter indicates corresponding figure part. (**B**) Skull height, (**C**) skull length, (**D**) skull total length, (sum of skull length and nasal bone length) and (**E**) skull width. (**F**) Nasal bone length and (**G**) percentage of nasal bone length normalized to skull length. (**H**) Nasal bone width and (**I**) percentage of nasal bone length normalized to skull width. *P* values calculated by one-way ANOVA. *Ash1l*$^{+/+}$, *n* = 6; *Ash1l+/-*, *n* = 11; *Ash1l*$^{-/-}$, *n* = 14. Data shown as mean ± SEM.



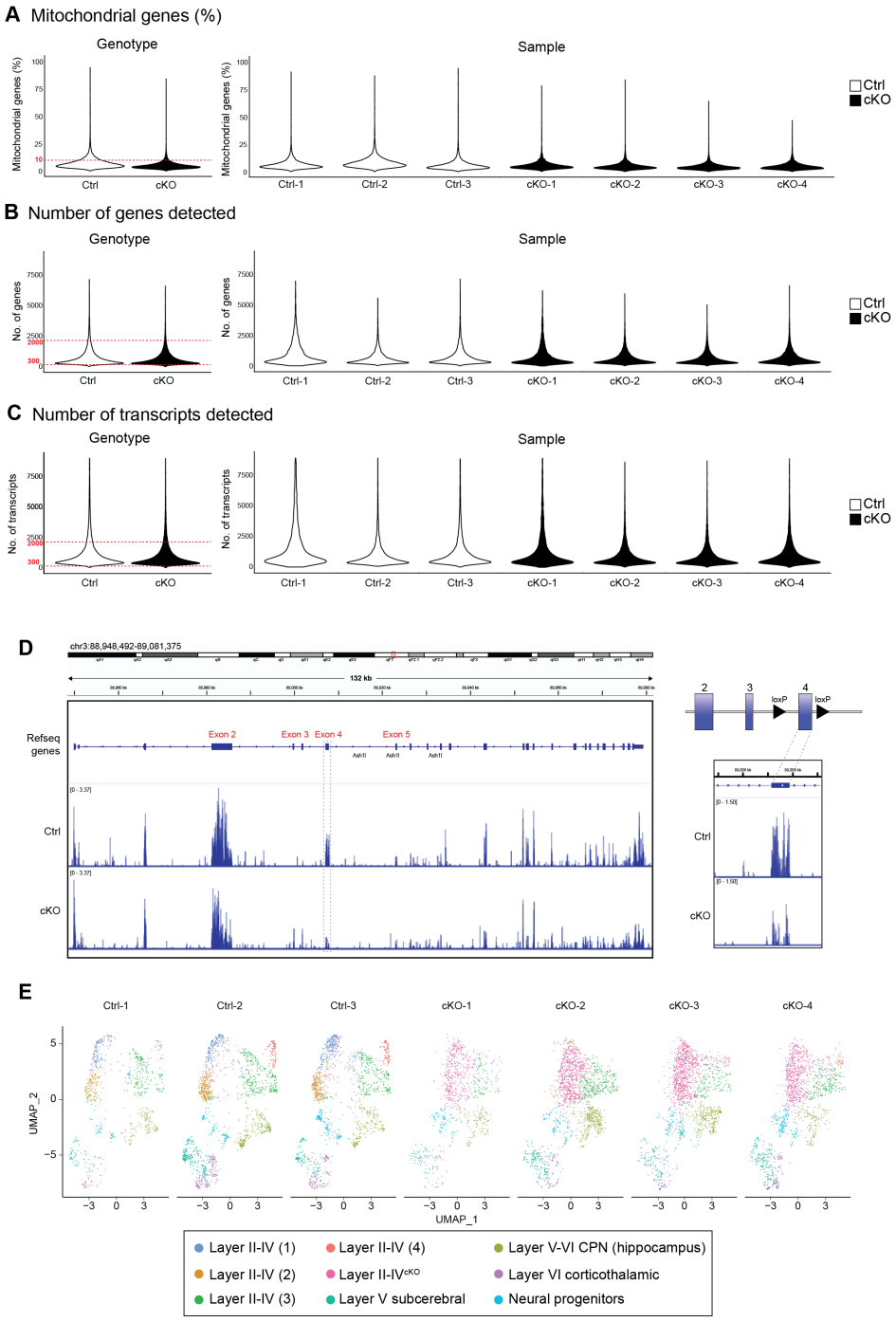
Supplemental figure 2: *Ash1l* deficiency does not affect fetal growth.

Body (**A**) and brain (**B**) weights of *Ash1l*$^{+/+}$, *Ash1l*$^{+/-}$ and *Ash1l*$^{-/-}$ mice collected at e18.5. *Ash1l*$^{+/+}$, n = 11; *Ash1l*$^{+/-}$, n = 22; *Ash1l*$^{-/-}$, n = 9. (**C**) Example of ventricle and whole brain section area analysis. (**D**) Percentage of ventricle area over total brain area at e18.5. *Ash1l*$^{+/+}$, n = 4; *Ash1l*$^{+/-}$, n = 3; *Ash1l*$^{-/-}$, n = 6. P values calculated by one-way ANOVA. Data shown as mean ± SEM. (**E**) Coronal sections of *Ash1l*$^{+/+}$ and *Ash1l*$^{+/-}$ brains.



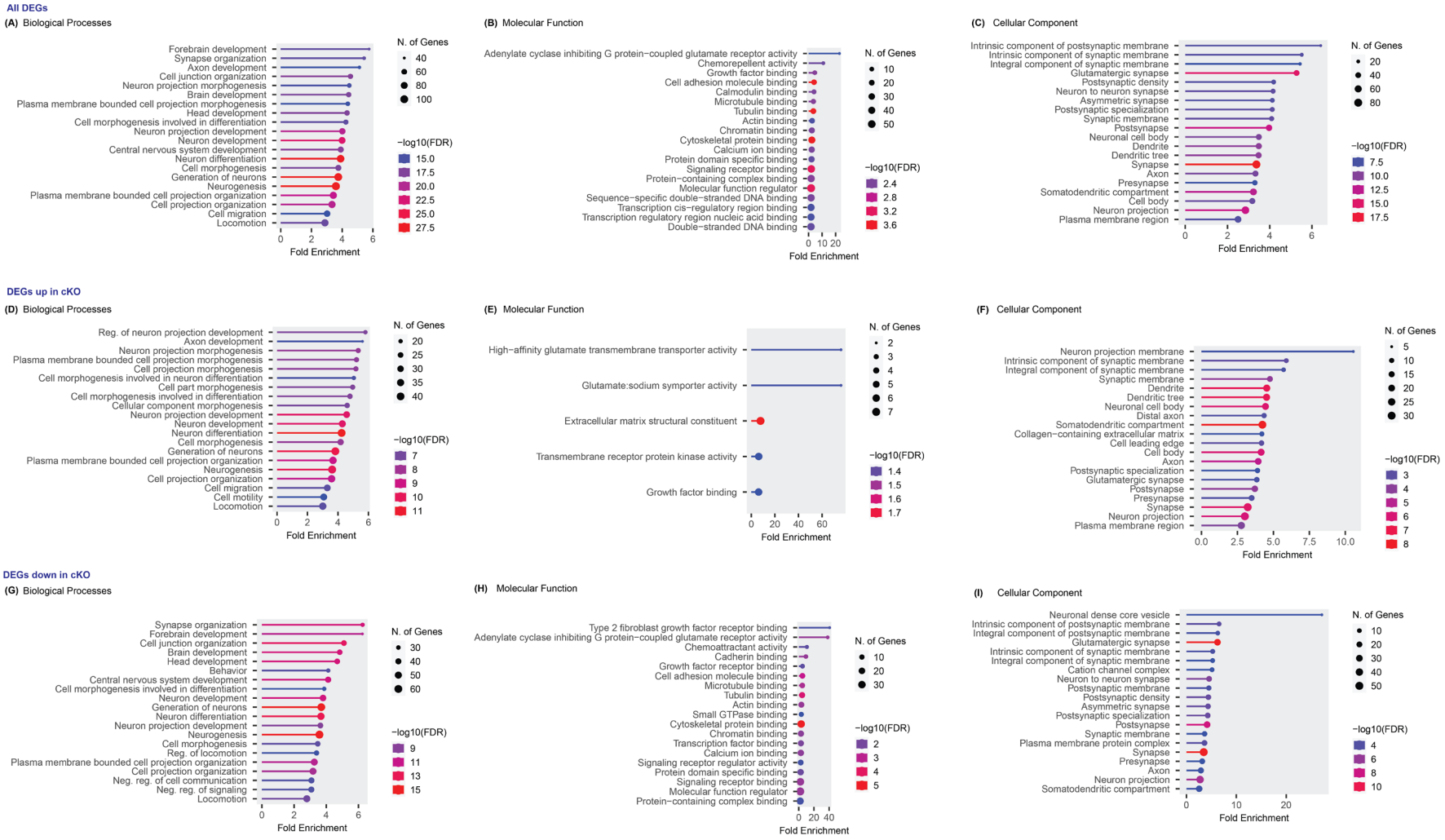
Supplemental figure 3: Conditional deletion of *Ash1l* does not affect cortical cell number.

DAPI (total cell number) raw counts of *Ash1l*^{Ctrl} and *Ash1l*^{cKO}. (**A**) DAPI count average for each sample used in lamination analysis figure 2. (**B**) DAPI count average for each sample used in birthdating analysis figure 3. Statistical analysis for both was determined by two-tailed unpaired t test. Significance shown by *P* values. Data shown as mean ± SEM.



Supplemental figure 4: Single cell RNA sequencing quality control reveals consistency across samples of the same genotype.

(**A**) Percentage of mitochondrial genes mapped between genotypes (left) and individual samples (right) before filtering. Red dotted line on genotype violin plot denotes threshold cut-off. (**B**) Number of genes detected and (**C**) number of molecules detected between genotypes (left) and individual samples (right) before filtering. Red dotted lines on genotype violin plots denote threshold range. (**D**) Integrative Genomics Viewer (IGV) screenshot of *Ash1l* gene and *Ash1l*^{Ctrl} and *Ash1l*^{cKO} reads. Barcodes of cells postfilter were used as a reference for the original BAM files that were generated during the Drop-seq Alignment protocol. These BAM reads were loaded into IGV and depletion of *Ash1l* exon 4 was seen in *Ash1l*^{cKO} compared to controls. Zoomed in reads of exon 4 shown on the right. (**E**) UMAPs of individual samples. Cell type key shown below.



Supplemental figure 5: Conditional deletion of *Ash1l* causes altered cortical gene expression.

ShinyGO lollipop charts showing affected biological processes, molecular functions, and cellular components. Differentially expressed genes found in reclustered UMAP shown in figure 5A. (**A-C**) biological processes, molecular functions, and cellular components of all differentially expressed genes, respectively. (**D-F**) biological processes, molecular functions, and cellular components of upregulated genes, respectively. (**G-I**) biological processes, molecular functions, and cellular components of downregulated genes, respectively.

Chapter 3: Conclusions and Future Directions

Haploinsufficiency of the histone methyltransferase *ASH1L* is associated with NDD. *ASH1L* regulates many genes necessary for organogenesis, including brain development. Various *Ash1l* mouse models have uncovered several central nervous system (CNS) phenotypes, including hyperactivity and ASD-like behaviors. However, *Ash1l*'s role in the developing cortex has not been characterized; this represents a sizable gap in knowledge, given the link between NDDs and the cortex.

This body of work aimed to determine the role of *Ash1l* in the developing mouse cortex. Ablation of *Ash1l* in the germline resulted in perinatal lethality, and craniofacial anomalies were observed, specifically a shorter nasal bone, in e18.5 *Ash1l*$^{-/-}$ embryos. There were no differences in brain weight relative to body weight, or differences in general brain structures between genotypes. E18.5 *Ash1l*^{cKO} cortices had twice as many SATB2 neurons compared to controls. The distribution of SATB2 was also altered; there were 18-29% more SATB2 neurons located in deeper cortical bins in *Ash1l*^{cKO} compared to *Ash1l*^{Ctrl}. EdU was injected at e15.5 and co-stained with SATB2 at e18.5; The analysis showed there were similar proportions of SATB2 neurons born at e15.5 between genotypes, suggesting that the surplus SATB2 neurons in the mutants had different birth timing. Cell populations in *Ash1l*^{Ctrl} and *Ash1l*^{cKO} were compared using scRNA-seq, and *Ash1l*^{cKO} had a unique layer II-IV cluster in lieu of two distinct layer II-IV clusters found in controls. Further analysis identified gene dysregulation in the layer II-

IV^{cKO} that is likely driving the unique identity. Layer II-IV neurons and neural progenitors were reclustered from both genotypes and pseudotime analysis was performed. Mutant upper layer neurons exhibited a stalled differentiation trajectory compared to controls, in line with delayed differentiation. These data point to an asynchronous cell fate acquisition phenotype, encompassing both premature and delayed differentiation. This suggests *ASH1L* is required for guiding cell fate during cortical development, and in its absence results in altered neuronal identity, similar to that of homeotic transformation caused by *ASH1L* deficiency in the reproductive tract and axial skeleton [136]. Homeotic-like transformations of neuronal cell identities have been documented previously, including *GATA2* conditional knockout using *En1*-cre transforming GABAergic neurons into glutamatergic neurons [178,179]. While this data furthers the field's understanding of how *ASH1L* loss of function impacts cortical development, it raises additional questions. In this section I will outline these questions and suggest various experiments that may be used to answer them.

3.1 Where are the Unique Layer II-IV^{cKO} Cells in the Cortex?

The scRNA-seq experiment of *Ash1l*^{cKO} and *Ash1l*^{Ctrl} cortices generated a large dataset that led to the discovery of the unique layer II-IV^{cKO} neurons in the mutants. While these cells had a unique transcriptomic identity, the gene markers driving this identity were not highly expressed. Conversely, markers expressed in control layer II-IV clusters were not hugely downregulated in layer II-IV^{cKO} neurons. This made it challenging to identify II-IV^{cKO} neurons *in vivo*; without a highly expressed marker, IHC or ISH could not be performed for these mutant specific neurons. In lieu of targetable

genes, spatial transcriptomics may be employed to determine where these unique neurons are in the cortex.

Spatial transcriptomics merges scRNA-seq technology with conventional histology techniques. It allows for sequencing of cells in a tissue section while preserving the histological data, resembling a map of the tissue. This can be carried out via two methods, namely imaging-based and sequencing-based spatial transcriptomics technologies [240]. For the purposes of locating layer II-IV^{cKO} neurons in cortical tissue, the sequencing-based approach is the more appropriate. Slide-seq is a commercially available technology that is well suited for this analysis [241,242]. The microarray slide or 'puck' has a grid of micro-wells, 10μm in diameter; as with scRNA-seq, these wells contain barcoded beads that have unique molecular identifiers (UMIs) for transcript identification, positional sequences that are mapped back to the position on the array and poly-dT tails that capture mRNA. Frozen tissue is cryosectioned and mounted onto the puck. The wells capture a cell and sequencing is carried out. One limitation of this technology is the resolution, as wells may capture more than one cell; however, the technology is constantly improving to allow for true single cell resolution.

Spatial transcriptomics is becoming a widely used tool, particularly in surveying highly heterogenous tissue like the cortex. Bhattacherjee *et al*. performed spatial transcriptomic analysis on adult mouse PFCs to better understand mechanisms underlying chronic pain. The authors used multiplexed error-robust fluorescence *in situ* hybridization (MERFISH) and identified a vast population of neuronal subtypes across the cortex. Test mice underwent spared nerve injury (SNI) to induce chronic pain and the transcriptomes were compared to control mice. The authors discovered that layer 5

extra-telencephalic neurons, specifically projecting into the periaqueductal gray area, were involved in a chronic pain circuit by comparing gene expression genes between test and control mice [243]. In another study, Bayraktar *et al.* performed spatial transcriptomics with single-molecule, fluorescent *in situ* hybridization (smFISH) on mouse cortices to determine if glial cells have distinctive lamination. Interestingly, astrocytes had broad expression gradients in the cortical plate, indicative of distinctive layers. Furthermore, the authors analyzed *Emx1-cre; Satb2*^{cKO} mice and found this lamination is disrupted, suggesting neurons help guide this astrocyte organization [244].

In the case of locating *Ash1l*^{cKO} specific layer II-IV^{cKO} neurons, using spatial transcriptomics would provide important data regarding the role of *Ash1l* in the developing cortex. For example, I hypothesize that the SATB2 neurons deposited in the deeper layers are members of the layer II-IV^{cKO} neuron population, which could be tested with the spatial analysis. If it was observed that the layer II-IV^{cKO} neurons were located close to one another, rather than spread out, it could suggest that these cells arise from a specific common progenitor. Furthermore, it could help establish a model for how *Ash1l* loss of function impacts neuronal circuitry.

3.2 What are the Molecular Mechanisms Underlying Gene Expression Changes in *Ash1l*^{cKO}?

ASH1L regulates gene expression; however, both H3K4 and H3K36 have been linked to its methyltransferase activity. I found gene dysregulation was driving the identity of *Ash1l*^{cKO} specific layer II-IV^{cKO} neurons, so understanding the underlying mechanism would provide greater insight into its role in cortical development. I attempted to answer this question with CUT&RUN (Cleavage Under Targets and

Release Using Nuclease) with qPCR [245]. Pseudo-bulk RNA-seq from the scRNA-seq experiment provided a list of DEGs, both up- and downregulated, in cortical tissue of *Ash1l*^{cKO} compared to *Ash1l*^{Ctrl}. I selected several downregulated genes, including *Basp1, Calm1* and *Jun,* and reviewed H3K36me2 peaks at these loci using the ENCODE database of p0 wildtype mouse cortices. I hypothesized that *Ash1l*-mediated H3K36me2 methylation activates these genes in normal tissue. I designed primers at H3K36me2 peaks, both at promotor and intragenic regions. Using a commercially available CUT&RUN kit (EpiCypher), I pulled down bound DNA with a H3K36me2 antibody from cortical lysates of e18.5 *Ash1l*^{cKO} and *Ash1l*^{Ctrl} embryos. Using the primers, I performed qPCR to determine quantity of H3K36me2 bound DNA between mutants and controls. The data was variable within genotype groups such that the results were inconclusive.

CUT&RUN with sequencing would be a useful next step in determining *Ash1l* targets in *Ash1l*^{cKO} and *Ash1l*^{Ctrl} cortices. The main advantage is the unbiased, broad approach offered by the sequencing. In addition to H3K36me2, I would pull down with H3K4me3, H3K36me3, PRC2-mediated H3K27me3 and *ASH1L* antibodies in mutant and control cortices. Determining where *ASH1L* is bound to would provide data on direct targets. Comparing dysregulated genes from the scRNA-seq dataset and levels of H3K4, K36 and K27 methylation would give a better understanding to the underlying molecular mechanics of *Ash1l*-associated gene regulation.

To complement this analysis, I would perform single cell assay for transposase-accessible chromatin with sequencing (scATAC-seq) [246,247]. This method would define open chromatin that is accessible for transcriptional components in individual

cells. ScATAC-seq can be integrated with the scRNA-seq data to find correlation between open/closed chromatin with up/downregulated genes, respectively. Comparing this data derived from *Ash1l*^{cKO} and *Ash1l*^{Ctrl} would help establish how *Ash1l* regulates gene expression, and further our understanding of how *Ash1l* deficiency impacts cortical development.

It would be of clinical relevance to assess how *ASH1L* haploinsufficiency impacts human neurons. Using induced pluripotent stem cells (iPSCs) from patients with various *ASH1L* variants and matched controls, I would differentiate them into neurons and perform RNA-seq, CUT&RUN and ATAC-seq. This data would provide information on how different variants can impact gene expression and chromatin accessibility.

3.3 What Cortical Cell Lineages Does *Ash1l* Loss-of-Function Impact?

The cortex is derived from a pool of common neural progenitor cells (NPCs), which, over the course of development, differentiate into distinct lineages. Intermediate progenitor cells (IPCs) undergo symmetric and asymmetric cell division, the latter generating a more specialized cell type at the loss of multipotency. A number of factors mediate this highly complex process, including cell cycle length and extracellular signaling [248]. The scRNA-seq data of *Ash1l*^{Ctrl} and *Ash1l*^{cKO} indicates that *Ash1l* has a role in guiding cell fate of a subset of upper layer neurons. However, specifically what lineages *Ash1l* loss-of-function impacts is not clear.

Post-mitotic neurons have a defined lineage that encompasses cells of varying relatedness; more closely related neurons are likely to be more similar in terms of gene expression and function. Cell lineage tracing is a commonly used technique to

investigate the developmental history and fate changes of a cell over time. Fluorescent proteins offer a way of labelling cells through a lineage. For example, a specific NPC may have a green fluorescent protein (GFP) expressing construct; after multiple rounds of division, all progeny will also express GFP. However, given the complexity, density and sheer number of neurons in the cortex, it is difficult to distinguish specific lineages using a single fluorescent protein per cell, even if there are multiple options (e.g. green and red fluorescent proteins; GFP and RFP respectively).

Stochastic multicolor labelling is a powerful technique that uses multiple fluorescent proteins in various combinations within a given cell, generating a spectrum of colors. Brainbow is one such approach to multicolor labelling [249]. Briefly, three or more genes encoding fluorescent proteins in a transgene are separated by loxP sites. In the presence of cre, each copy of the transgene will undergo independent recombination, either excising or inverting a random combination of fluorescent genes. The first gene along this sequence is the expressed color. With multiple copies of the transgene in a given cell, the combined fluorescing proteins emit a distinct color; furthermore, the progeny of that cell will also be labelled with that color [250]. The R26-Confetti mouse line utilizes the Brainbow2.1 sequence (*loxP*-STOP-*loxP*-GFP-PFY-*Pxol*-*loxP*-RFP-PFC-*Pxol*), with a CAG promotor and floxed-STOP cassette upstream [251]. Variations of this line include tamoxifen-inducible cre. Snippert *et al.* first used this mouse line to investigate stem cell proliferation in the intestinal crypt. The authors used a tamoxifen inducible cre specific to intestinal stem cells (Lgr5-CreER) with the R26-Confetti mouse and found the stem cell population is regulated via neutral competition [251]. R26-Confetti mice have been used to trace progenitor populations in the

developing mouse brain. Pilz *et al.* crossed R26-Confetti mice with GLAST^{CreERT2}, a radial glia specific cre. Using low doses of tamoxifen, the authors were able to sparsely label radial glia and found differing proliferative behavior among them [252].

To determine what cortical lineages are guided in part by *Ash1l,* I would cross an *Ash1l* specific, inducible cre line (*Ash1l*^{CreERT2}) with R26-Confetti mice. At e15.5, coinciding with upper layer neuron development, I would inject low doses of tamoxifen to induce sparse labelling. At e18.5, I would collect brains for analysis. All fluorescent neurons represent *Ash1l* expression at e15.5; neurons of the same color represent progeny of one cell. Mapping and quantifying these neurons would provide an overview of the impact of *Ash1l* driven fate of upper layer neurons. To further investigate this role, I would compare *Ash1l*^{cKO} and *Ash1l*^{Ctrl} cortices with the R26-Confetti construct. I would likely need to adjust the crossing parameters and use a non-inducible *Ash1l* conditional knockout (*Emx1-cre),* as the initial tamoxifen injection will also activate the fluorescence. Again, I would inject at e15.5. Comparing the size and number of same-colored neurons would provide spatial information of how *Ash1l* impacts neuronal lineages. Injection of tamoxifen at different time points will also indicate when *Ash1l* is important for fate guidance.

3.4 How Does *Ash1l* Loss of Function Impact Other Organ Systems?

The literature on *Ash1l* mouse models indicates it is an important gene, as germline knockout or genetrap alleles result in early lethality and severe phenotypes. Brinkmeier *et al.* demonstrated that multiple organ systems were affected in *Ash1l*$^{GT/GT}$ mice, which is perhaps unsurprising given the broad expression of *Ash1l* across tissues

[136]. Clinical reports of individuals with an *ASH1L* variant have various diagnoses, including craniofacial and skeletal anomalies. Deciphering how *ASH1L* loss of function impacts organ systems other than the brain may provide therapeutic targets that can be addressed, improving overall quality of life of the patient.

Conditional knockout mice allow for ablation of gene expression in a tissue specific manner. I would cross the *Ash1l* floxed allele, used in this study, with various tissue specific cre lines. *Drosophila ash1* has been shown to be required for normal heart development and function [253]. αMHC-MerCreMer mice have a cardiac myocyte specific, tamoxifen-inducible cre which can be used to investigate the role of *Ash1l* in heart development [254]. Importantly, later induction of the cre, such as the adult stage, would provide information on maintenance roles. *Ash1l*$^{GT/GT}$ females exhibited smaller and twisted uteri, which could not support implantation of fertilized eggs because they lacked the endometrial glands of the uterus [136]; using a uterus specific cre, such as homozygous *Pgr*^{iresCre}, would tease apart origins of this phenotype [255]. By employing CUT&RUN and pulling down with *ASH1L*, I may assess where *ASH1L* is binding to in the genome and what genes are being regulated that result in the phenotypes.

Several *Ash1l* mouse models exhibit various skeletal anomalies [23,133,136]. Several skeletal conditional knockout mouse lines, including *Prx1-cre* and *Osx1-cre,* have been used to establish gene function in skeleton development [256,257]. Interestingly, the *Prx1-cre* is active in the developing head and limb mesenchyme, which may be used to investigate *Ash1l*-associated craniofacial anomalies. A further avenue of investigation would be the role of *Ash1l* in bone marrow. *Ash1l* has been shown to regulate hematopoietic stem cell (HSC) self-renewal and quiescence, and

implicated in various blood cancers [137,258,259] . Conditional knockout of *Ash1l* in the bone marrow, using lines such as *Lyz-cre*, would be a powerful model to assess for cancer.

This thesis work set out to determine the role of *ASH1L* during cortical development in the context of ASD and other NDDs. In summary, *Ash1l*^{cKO} cortices had more SATB2 expressing neurons that were present in upper and deeper cortical layers. Birthdating experiments determined there was no difference in proliferation at e15.5, and co-staining with SATB2 antibody suggested these excess SATB2 expressing neurons were born prior to e15.5. ScRNA-seq found a unique layer II-IV cluster in lieu of two layer II-IV neuron clusters found in controls. After reclustering all layer II-IV clusters, pseudotime analysis revealed altered differentiation trajectory in *Ash1l*^{cKO}. These new data suggest that *Ash1l* is required for timing of differentiation and mature fate of upper layer cortical neurons. To better understand the impact of *Ash1l* ablation in the cortex, several future directions have been suggested. This includes the use of spatial transcriptomics, that would determine where these unique layer II-IV^{cKO} neurons reside. This would provide valuable data on this cell type origin and timing of differentiation. Furthermore, using CUT&RUN to uncover *ASH1L* mediated histone methylation will be essential in deciphering direct and indirect gene expression changes. Stochastic multicolor labelling would a useful future direction to better understand affected cortical lineages by *Ash1l* loss of function. More broadly, determining phenotypes found with various *Ash1l* conditional deletion in different organ systems would provide insight into the range of clinical reports attributed to *ASH1L* loss of function. By better

understanding the underlying mechanisms and resulting phenotypes using mice and other models, this will inform researchers on potential treatments for improving quality of life for those affected by *Ash1l* variants.

Bibliography

[1] Cabianca DS, Casa V, Bodega B, Xynos A, Ginelli E, Tanaka Y, et al. A Long ncRNA Links Copy Number Variation to a Polycomb/Trithorax Epigenetic Switch in FSHD Muscular Dystrophy. Cell 2012;149:819–31. https://doi.org/10.1016/j.cell.2012.03.035.

[2] Rea S, Eisenhaber F, Ponting CP, Allis CD, Jenuwein T. Regulation of chromatin structure by site-specific histone H3 methyltransferases 2000;406.

[3] Nakamura T, Blechman J, Tada S, Rozovskaia T, Itoyama T, Bullrich F, et al. huASH1 protein, a putative transcription factor encoded by a human homologue of the *Drosophila ash1* gene, localizes to both nuclei and cell–cell tight junctions. Proc Natl Acad Sci USA 2000;97:7284–9. https://doi.org/10.1073/pnas.97.13.7284.

[4] Shearn A, Hersperger E, Hersperger G. Genetic studies of mutations at two loci of Drosophila melanogaster which cause a wide variety of homeotic transformations. Roux's Arch Dev Biol 1987;196:231–42. https://doi.org/10.1007/BF00376347.

[5] ASH1L - DECIPHER v11.22 n.d. https://www.deciphergenomics.org/gene/ASH1L/patient-overlap/cnvs (accessed October 2, 2023).

[6] ASH1L - SFARI n.d. https://gene.sfari.org/database/human-gene/ASH1L (accessed October 30, 2023).

[7] Steffen PA, Ringrose L. What are memories made of? How Polycomb and Trithorax proteins mediate epigenetic memory. Nat Rev Mol Cell Biol 2014;15:340–56. https://doi.org/10.1038/nrm3789.

[8] Zhou K, Gaullier G, Luger K. Nucleosome structure and dynamics are coming of age. Nat Struct Mol Biol 2019;26:3–13. https://doi.org/10.1038/s41594-018-0166-x.

[9] Iwasaki W, Miya Y, Horikoshi N, Osakabe A, Taguchi H, Tachiwana H, et al. Contribution of histone N-terminal tails to the structure and stability of nucleosomes. FEBS Open Bio 2013;3:363–9. https://doi.org/10.1016/j.fob.2013.08.007.

[10] Kassis JA, Kennison JA, Tamkun JW. Polycomb and Trithorax Group Genes in *Drosophila*. Genetics 2017;206:1699–725. https://doi.org/10.1534/genetics.115.185116.

[11] Lewis EB. A gene complex controlling segmentation in Drosophila. Nature 1978;276:565–70. https://doi.org/10.1038/276565a0.

[12] Gould A. Functions of mammalian Polycomb group and trithorax group related genes. Current Opinion in Genetics & Development 1997;7:488–94. https://doi.org/10.1016/S0959-437X(97)80075-5.

[13] Kuehner, Yao. The Dynamic Partnership of Polycomb and Trithorax in Brain Development and Diseases. Epigenomes 2019;3:17. https://doi.org/10.3390/epigenomes3030017.

[14] Lee MG, Norman J, Shilatifard A, Shiekhattar R. Physical and Functional Association of a Trimethyl H3K4 Demethylase and Ring6a/MBLR, a Polycomb-like Protein. Cell 2007;128:877–87. https://doi.org/10.1016/j.cell.2007.02.004.

[15] Wang H, Wang L, Erdjument-Bromage H, Vidal M, Tempst P, Jones RS, et al. Role of histone H2A ubiquitination in Polycomb silencing. Nature 2004;431:873–8. https://doi.org/10.1038/nature02985.

[16] Scheuermann JC, De Ayala Alonso AG, Oktaba K, Ly-Hartig N, McGinty RK, Fraterman S, et al. Histone H2A deubiquitinase activity of the Polycomb repressive complex PR-DUB. Nature 2010;465:243–7. https://doi.org/10.1038/nature08966.

[17] Bonnet J, Boichenko I, Kalb R, Le Jeune M, Maltseva S, Pieropan M, et al. PR-DUB preserves Polycomb repression by preventing excessive accumulation of H2Aub1, an antagonist of chromatin compaction. Genes Dev 2022;36:1046–61. https://doi.org/10.1101/gad.350014.122.

[18] Byrd KN, Shearn A. ASH1, a *Drosophila* trithorax group protein, is required for methylation of lysine 4 residues on histone H3. Proc Natl Acad Sci USA 2003;100:11535–40. https://doi.org/10.1073/pnas.1933593100.

[19] Hallson G, Hollebakken RE, Li T, Syrzycka M, Kim I, Cotsworth S, et al. dSet1 Is the Main H3K4 Di- and Tri-Methyltransferase Throughout *Drosophila* Development. Genetics 2012;190:91–100. https://doi.org/10.1534/genetics.111.135863.

[20] Tanaka Y, Katagiri Z, Kawahashi K, Kioussis D, Kitajima S. Trithorax-group protein ASH1 methylates histone H3 lysine 36. Gene 2007;397:161–8. https://doi.org/10.1016/j.gene.2007.04.027.

[21] Yuan W, Xu M, Huang C, Liu N, Chen S, Zhu B. H3K36 Methylation Antagonizes PRC2-mediated H3K27 Methylation. Journal of Biological Chemistry 2011;286:7983–9. https://doi.org/10.1074/jbc.M110.194027.

[22] Qin L, Williams JB, Tan T, Liu T, Cao Q, Ma K, et al. Deficiency of autism risk factor ASH1L in prefrontal cortex induces epigenetic aberrations and seizures. Nat Commun 2021;12:6589. https://doi.org/10.1038/s41467-021-26972-8.

[23] Miyazaki H, Higashimoto K, Yada Y, Endo TA, Sharif J, Komori T, et al. Ash1l Methylates Lys36 of Histone H3 Independently of Transcriptional Elongation to Counteract Polycomb Silencing. PLoS Genet 2013;9:e1003897. https://doi.org/10.1371/journal.pgen.1003897.

[24] Yan Y, Tian M, Li M, Zhou G, Chen Q, Xu M, et al. ASH1L haploinsufficiency results in autistic-like phenotypes in mice and links Eph receptor gene to autism spectrum disorder. Neuron 2022;110:1156-1172.e9. https://doi.org/10.1016/j.neuron.2021.12.035.

[25] Gao Z, Zhang J, Bonasio R, Strino F, Sawai A, Parisi F, et al. PCGF Homologs, CBX Proteins, and RYBP Define Functionally Distinct PRC1 Family Complexes. Molecular Cell 2012;45:344–56. https://doi.org/10.1016/j.molcel.2012.01.002.

[26] Schuettengruber B, Bourbon H-M, Di Croce L, Cavalli G. Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. Cell 2017;171:34–57. https://doi.org/10.1016/j.cell.2017.08.002.

[27] Kolovos P, Nishimura K, Sankar A, Sidoli S, Cloos PA, Helin K, et al. PR-DUB maintains the expression of critical genes through FOXK1/2- and ASXL1/2/3-dependent recruitment to chromatin and H2AK119ub1 deubiquitination. Genome Res 2020;30:1119–30. https://doi.org/10.1101/gr.261016.120.

[28] Geisler SJ, Paro R. Trithorax and Polycomb group-dependent regulation: a tale of opposing activities. Development 2015;142:2876–87. https://doi.org/10.1242/dev.120030.

[29] Hou P, Huang C, Liu C-P, Yang N, Yu T, Yin Y, et al. Structural Insights into Stimulation of Ash1L's H3K36 Methyltransferase Activity through Mrg15 Binding. Structure 2019;27:837-845.e3. https://doi.org/10.1016/j.str.2019.01.015.

[30] Al-Harthi S, Li H, Winkler A, Szczepski K, Deng J, Grembecka J, et al. MRG15 activates histone methyltransferase activity of ASH1L by recruiting it to the nucleosomes. Structure 2023:S0969212623002447. https://doi.org/10.1016/j.str.2023.07.001.

[31] Lee Y, Yoon E, Cho S, Schmähling S, Müller J, Song J-J. Structural Basis of MRG15-Mediated Activation of the ASH1L Histone Methyltransferase by Releasing an Autoinhibitory Loop. Structure 2019;27:846-852.e3. https://doi.org/10.1016/j.str.2019.01.016.

[32] Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, et al. A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. Cell 2006;125:315–26. https://doi.org/10.1016/j.cell.2006.02.041.

[33] Hirabayashi Y, Suzki N, Tsuboi M, Endo TA, Toyoda T, Shinga J, et al. Polycomb Limits the Neurogenic Competence of Neural Precursor Cells to Promote Astrogenic Fate Transition. Neuron 2009;63:600–13. https://doi.org/10.1016/j.neuron.2009.08.021.

[34] Molofsky AV, Pardal R, Iwashita T, Park I-K, Clarke MF, Morrison SJ. Bmi-1 dependence distinguishes neural stem cell self-renewal from progenitor proliferation. Nature 2003;425:962–7. https://doi.org/10.1038/nature02060.

[35] Pereira JD, Sansom SN, Smith J, Dobenecker M-W, Tarakhovsky A, Livesey FJ. Ezh2, the histone methyltransferase of PRC2, regulates the balance between self-renewal and differentiation in the cerebral cortex. Proc Natl Acad Sci USA 2010;107:15957–62. https://doi.org/10.1073/pnas.1002530107.

[36] Zemke M, Draganova K, Klug A, Schöler A, Zurkirchen L, Gay MH-P, et al. Loss of Ezh2 promotes a midbrain-to-forebrain identity switch by direct gene derepression and Wnt-dependent regulation. BMC Biol 2015;13:103. https://doi.org/10.1186/s12915-015-0210-9.

[37] Feng W, Khan MA, Bellvis P, Zhu Z, Bernhardt O, Herold-Mende C, et al. The Chromatin Remodeler CHD7 Regulates Adult Neurogenesis via Activation of SoxC Transcription Factors. Cell Stem Cell 2013;13:62–72. https://doi.org/10.1016/j.stem.2013.05.002.

[38] Weaver DD, Graham CB, Thomas IT, Smith DW. A new overgrowth syndrome with accelerated skeletal maturation, unusual facies, and camptodactyly. The Journal of Pediatrics 1974;84:547–52. https://doi.org/10.1016/S0022-3476(74)80675-X.

[39] Majewski F, Ranke M, Kemperdick H, Schmidt E. The weaver syndrome: A rare type of primordial overgrowth. Eur J Pediatr 1981;137:277–82. https://doi.org/10.1007/BF00443257.

[40] Cohen ASA, Tuysuz B, Shen Y, Bhalla SK, Jones SJM, Gibson WT. A novel mutation in EED associated with overgrowth. J Hum Genet 2015;60:339–42. https://doi.org/10.1038/jhg.2015.26.

[41] Imagawa E, Higashimoto K, Sakai Y, Numakura C, Okamoto N, Matsunaga S, et al. Mutations in genes encoding polycomb repressive complex 2 subunits cause Weaver syndrome. Hum Mutat 2017;38:637–48. https://doi.org/10.1002/humu.23200.

[42] Gibson WT, Hood RL, Zhan SH, Bulman DE, Fejes AP, Moore R, et al. Mutations in EZH2 Cause Weaver Syndrome. The American Journal of Human Genetics 2012;90:110–8. https://doi.org/10.1016/j.ajhg.2011.11.018.

[43] The Childhood Overgrowth Collaboration, Tatton-Brown K, Hanks S, Ruark E, Zachariou A, Duarte SDV, et al. Germline mutations in the oncogene EZH2 cause Weaver syndrome and increased human height. Oncotarget 2011;2:1127–33. https://doi.org/10.18632/oncotarget.385.

[44] Lui JC, Barnes KM, Dong L, Yue S, Graber E, Rapaport R, et al. Ezh2 Mutations Found in the Weaver Overgrowth Syndrome Cause a Partial Loss of H3K27 Histone Methyltransferase Activity. The Journal of Clinical Endocrinology & Metabolism 2018;103:1470–8. https://doi.org/10.1210/jc.2017-01948.

[45] Bainbridge MN, Hu H, Muzny DM, Musante L, Lupski JR, Graham BH, et al. De novo truncating mutations in ASXL3 are associated with a novel clinical phenotype with similarities to Bohring-Opitz syndrome. Genome Med 2013;5:11. https://doi.org/10.1186/gm415.

[46] Bölicke N, Albert M. Polycomb-mediated gene regulation in human brain development and neurodevelopmental disorders. Developmental Neurobiology 2022;82:345–63. https://doi.org/10.1002/dneu.22876.

[47] Tamburri S, Conway E, Pasini D. Polycomb-dependent histone H2A ubiquitination links developmental disorders with cancer. Trends in Genetics 2022;38:333–52. https://doi.org/10.1016/j.tig.2021.07.011.

[48] Conway E, Rossi F, Fernandez-Perez D, Ponzo E, Ferrari KJ, Zanotti M, et al. BAP1 enhances Polycomb repression by counteracting widespread H2AK119ub1 deposition and chromatin condensation. Molecular Cell 2021;81:3526-3541.e8. https://doi.org/10.1016/j.molcel.2021.06.020.

[49] Stevens CA. Rubinstein-Taybi Syndrome n.d.

[50] Ajmone PF, Avignone S, Gervasini C, Giacobbe A, Monti F, Costantino A, et al. Rubinstein–Taybi syndrome: New neuroradiological and neuropsychiatric insights from a multidisciplinary approach. American J of Med Genetics Pt B 2018;177:406–15. https://doi.org/10.1002/ajmg.b.32628.

[51] De Thonel A, Ahlskog JK, Daupin K, Dubreuil V, Berthelet J, Chaput C, et al. CBP-HSF2 structural and functional interplay in Rubinstein-Taybi neurodevelopmental disorder. Nat Commun 2022;13:7002. https://doi.org/10.1038/s41467-022-34476-2.

[52] Chang Y, Östling P, Åkerfelt M, Trouillet D, Rallu M, Gitton Y, et al. Role of heat-shock factor 2 in cerebral cortex formation and as a regulatorof p35 expression. Genes Dev 2006;20:836–47. https://doi.org/10.1101/gad.366906.

[53] Aggarwal A, Rodriguez-Buritica DF, Northrup H. Wiedemann-Steiner syndrome: Novel pathogenic variant and review of literature. European Journal of Medical Genetics 2017;60:285–8. https://doi.org/10.1016/j.ejmg.2017.03.006.

[54] Baer S, Afenjar A, Smol T, Piton A, Gérard B, Alembik Y, et al. Wiedemann-Steiner syndrome as a major cause of syndromic intellectual disability: A study of 33 French cases. Clinical Genetics 2018;94:141–52. https://doi.org/10.1111/cge.13254.

[55] Jones WD, Dafou D, McEntagart M, Woollard WJ, Elmslie FV, Holder-Espinasse M, et al. De Novo Mutations in MLL Cause Wiedemann-Steiner Syndrome. The American Journal of Human Genetics 2012;91:358–64. https://doi.org/10.1016/j.ajhg.2012.06.008.

[56] Boniel S, Szymańska K, Śmigiel R, Szczałuba K. Kabuki Syndrome—Clinical Review with Molecular Aspects. Genes 2021;12:468. https://doi.org/10.3390/genes12040468.

[57] Moccia A, Martin DM. Nervous system development and disease: A focus on trithorax related proteins and chromatin remodelers. Molecular and Cellular Neuroscience 2018;87:46–54. https://doi.org/10.1016/j.mcn.2017.11.016.

[58] Tsurusaki Y, Okamoto N, Ohashi H, Kosho T, Imai Y, Hibi-Ko Y, et al. Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. Nat Genet 2012;44:376–8. https://doi.org/10.1038/ng.2219.

[59] Hsu P, Ma A, Wilson M, Williams G, Curotta J, Munns CF, et al. CHARGE syndrome: A review. J Paediatrics Child Health 2014;50:504–11. https://doi.org/10.1111/jpc.12497.

[60] Feng W, Kawauchi D, Körkel-Qu H, Deng H, Serger E, Sieber L, et al. Chd7 is indispensable for mammalian brain development through activation of a neuronal differentiation programme. Nat Commun 2017;8:14758. https://doi.org/10.1038/ncomms14758.

[61] Kuo AJ, Cheung P, Chen K, Zee BM, Kioi M, Lauring J, et al. NSD2 Links Dimethylation of Histone H3 at Lysine 36 to Oncogenic Programming. Molecular Cell 2011;44:609–20. https://doi.org/10.1016/j.molcel.2011.08.042.

[62] Huang C, Zhu B. Roles of H3K36-specific histone methyltransferases in transcription: antagonizing silencing and safeguarding transcription fidelity. Biophys Rep 2018;4:170–7. https://doi.org/10.1007/s41048-018-0063-1.

[63] Edmunds JW, Mahadevan LC, Clayton AL. Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation. EMBO J 2008;27:406–20. https://doi.org/10.1038/sj.emboj.7601967.

[64] Streubel G, Watson A, Jammula SG, Scelfo A, Fitzpatrick DJ, Oliviero G, et al. The H3K36me2 Methyltransferase Nsd1 Demarcates PRC2-Mediated H3K27me2 and H3K27me3 Domains in Embryonic Stem Cells. Molecular Cell 2018;70:371-379.e5. https://doi.org/10.1016/j.molcel.2018.02.027.

[65] Kurotaki N, Imaizumi K, Harada N, Masuno M, Kondoh T, Nagai T, et al. Haploinsufficiency of NSD1 causes Sotos syndrome. Nat Genet 2002;30:365–6. https://doi.org/10.1038/ng863.

[66] Tatton-Brown K, Douglas J, Coleman K, Baujat G, Cole TRP, Das S, et al. Genotype-Phenotype Associations in Sotos Syndrome: An Analysis of 266 Individuals with NSD1 Aberrations. The American Journal of Human Genetics 2005;77:193–204. https://doi.org/10.1086/432082.

[67] Oishi S, Zalucki O, Vega MS, Harkins D, Harvey TJ, Kasherman M, et al. Investigating cortical features of Sotos syndrome using mice heterozygous for *Nsd1*. Genes Brain and Behavior 2020;19:e12637. https://doi.org/10.1111/gbb.12637.

[68] Barrie ES, Alfaro MP, Pfau RB, Goff MJ, McBride KL, Manickam K, et al. De novo loss-of-function variants in *NSD2* (*WHSC1*) associate with a subset of Wolf–Hirschhorn syndrome. Cold Spring Harb Mol Case Stud 2019;5:a004044. https://doi.org/10.1101/mcs.a004044.

[69] Chen M, Quan Y, Duan G, Wu H, Bai T, Wang Y, et al. Mutation pattern and genotype-phenotype correlations of SETD2 in neurodevelopmental disorders. European Journal of Medical Genetics 2021;64:104200. https://doi.org/10.1016/j.ejmg.2021.104200.

[70] Lee S, Menzies L, Hay E, Ochoa E, Docquier F, Rodger F, et al. Epigenotype–genotype–phenotype correlations in *SETD1A* and *SETD2* chromatin disorders. Human Molecular Genetics 2023;32:3123–34. https://doi.org/10.1093/hmg/ddad079.

[71] Luscan A, Laurendeau I, Malan V, Francannet C, Odent S, Giuliano F, et al. Mutations in *SETD2* cause a novel overgrowth condition. J Med Genet 2014;51:512–7. https://doi.org/10.1136/jmedgenet-2014-102402.

[72] Pappas J, Rabin R. SETD2 Neurodevelopmental Disorders. Neurodevelopmental Disorders n.d.

[73] Fagerberg L, Hallström BM, Oksvold P, Kampf C, Djureinovic D, Odeberg J, et al. Analysis of the Human Tissue-specific Expression by Genome-wide Integration of Transcriptomics and Antibody-based Proteomics. Molecular & Cellular Proteomics 2014;13:397–406. https://doi.org/10.1074/mcp.M113.035600.

[74] ASH1L - Histone-lysine N-methyltransferase ASH1L - Homo sapiens (Human) | UniProtKB | UniProt n.d. https://www.uniprot.org/uniprotkb/Q9NR48/entry#Q9NR48-1 (accessed October 31, 2023).

[75] Reeves R, Nissen MS. The A.T-DNA-binding domain of mammalian high mobility group I chromosomal proteins. A novel peptide motif for recognizing DNA

structure. Journal of Biological Chemistry 1990;265:8573–82. https://doi.org/10.1016/S0021-9258(19)38926-4.

[76] Aravind L. AT-hook motifs identified in a wide variety of DNA-binding proteins. Nucleic Acids Research 1998;26:4413–21. https://doi.org/10.1093/nar/26.19.4413.

[77] Steffen PA, Altmutter C, Dworschak E, Junttila S, Gyenesei A, Zhu X, et al. The Trithorax group protein ASH1 requires a combination of BAH domain and AT hooks, but not the SET domain, for mitotic chromatin binding and survival. Chromosoma 2021;130:215–34. https://doi.org/10.1007/s00412-021-00762-z.

[78] Wagner EJ, Carpenter PB. Understanding the language of Lys36 methylation at histone H3. Nat Rev Mol Cell Biol 2012;13:115–26. https://doi.org/10.1038/nrm3274.

[79] Lam UTF, Tan BKY, Poh JJX, Chen ES. Structural and functional specificity of H3K36 methylation. Epigenetics & Chromatin 2022;15:17. https://doi.org/10.1186/s13072-022-00446-7.

[80] Rogawski DS, Ndoj J, Cho HJ, Maillard I, Grembecka J, Cierpicki T. Two Loops Undergoing Concerted Dynamics Regulate the Activity of the ASH1L Histone Methyltransferase. Biochemistry 2015;54:5401–13. https://doi.org/10.1021/acs.biochem.5b00697.

[81] Schmähling S, Meiler A, Lee Y, Mohammed A, Finkl K, Tauscher K, et al. Regulation and function of H3K36 di-methylation by the trithorax-group protein complex AMC. Development 2018:dev.163808. https://doi.org/10.1242/dev.163808.

[82] An S, Yeo KJ, Jeon YH, Song J-J. Crystal Structure of the Human Histone Methyltransferase ASH1L Catalytic Domain and Its Implications for the Regulatory Mechanism. Journal of Biological Chemistry 2011;286:8369–74. https://doi.org/10.1074/jbc.M110.203380.

[83] Filippakopoulos P, Knapp S. The bromodomain interaction module. FEBS Letters 2012;586:2692–704. https://doi.org/10.1016/j.febslet.2012.04.045.

[84] Zaware N, Zhou M-M. Bromodomain biology and drug discovery. Nat Struct Mol Biol 2019;26:870–9. https://doi.org/10.1038/s41594-019-0309-8.

[85] Li Y, Li H. Many keys to push: diversifying the “readership” of plant homeodomain fingers. ABBS 2012;44:28–39. https://doi.org/10.1093/abbs/gmr117.

[86] Chambers AL, Pearl LH, Oliver AW, Downs JA. The BAH domain of Rsc2 is a histone H3 binding domain. Nucleic Acids Research 2013;41:9168–82. https://doi.org/10.1093/nar/gkt662.

[87] Duncker BP, Chesnokov IN, McConkey BJ. The origin recognition complex protein family. Genome Biol 2009;10:214. https://doi.org/10.1186/gb-2009-10-3-214.

[88] Kuo AJ, Song J, Cheung P, Ishibe-Murakami S, Yamazoe S, Chen JK, et al. The BAH domain of ORC1 links H4K20me2 to DNA replication licensing and Meier–Gorlin syndrome. Nature 2012;484:115–9. https://doi.org/10.1038/nature10956.

[89] Noguchi K, Vassilev A, Ghosh S, Yates JL, DePamphilis ML. The BAH domain facilitates the ability of human Orc1 protein to activate replication origins in vivo. EMBO J 2006;25:5372–82. https://doi.org/10.1038/sj.emboj.7601396.

[90] Fan H, Lu J, Guo Y, Li D, Zhang Z-M, Tsai Y-H, et al. BAHCC1 binds H3K27me3 via a conserved BAH module to mediate gene silencing and oncogenesis. Nat Genet 2020;52:1384–96. https://doi.org/10.1038/s41588-020-00729-3.

[91] Fan H, Guo Y, Tsai Y-H, Storey AJ, Kim A, Gong W, et al. A conserved BAH module within mammalian BAHD1 connects H3K27me3 to Polycomb gene silencing. Nucleic Acids Research 2021;49:4441–55. https://doi.org/10.1093/nar/gkab210.

[92] Zhao D, Zhang X, Guan H, Xiong X, Shi X, Deng H, et al. The BAH domain of BAHD1 is a histone H3K27me3 reader. Protein Cell 2016;7:222–6. https://doi.org/10.1007/s13238-016-0243-z.

[93] Pourpre R, Naudon L, Meziane H, Lakisic G, Jouneau L, Varet H, et al. BAHD1 haploinsufficiency results in anxiety-like phenotypes in male mice. PLoS ONE 2020;15:e0232789. https://doi.org/10.1371/journal.pone.0232789.

[94] Morris-Rosendahl DJ, Crocq M-A. Neurodevelopmental disorders—the history and future of a diagnostic concept. Dialogues in Clinical Neuroscience 2020;22:65–72. https://doi.org/10.31887/DCNS.2020.22.1/macrocq.

[95] Hu WF, Chahrour MH, Walsh CA. The Diverse Genetic Landscape of Neurodevelopmental Disorders. Annu Rev Genom Hum Genet 2014;15:195–213. https://doi.org/10.1146/annurev-genom-090413-025600.

[96] Parenti I, Rabaneda LG, Schoen H, Novarino G. Neurodevelopmental Disorders: From Genetics to Functional Pathways. Trends in Neurosciences 2020;43:608–21. https://doi.org/10.1016/j.tins.2020.05.004.

[97] Iossifov I, Ronemus M, Levy D, Wang Z, Hakker I, Rosenbaum J, et al. De Novo Gene Disruptions in Children on the Autistic Spectrum. Neuron 2012;74:285–99. https://doi.org/10.1016/j.neuron.2012.04.009.

[98] De Ligt J, Willemsen MH, Van Bon BWM, Kleefstra T, Yntema HG, Kroes T, et al. Diagnostic Exome Sequencing in Persons with Severe Intellectual Disability. N Engl J Med 2012;367:1921–9. https://doi.org/10.1056/NEJMoa1206524.

[99] De Rubeis S, The DDD Study, Homozygosity Mapping Collaborative for Autism, UK10K Consortium, The Autism Sequencing Consortium, He X, et al. Synaptic, transcriptional and chromatin genes disrupted in autism. Nature 2014;515:209–15. https://doi.org/10.1038/nature13772.

[100] Krumm N, Turner TN, Baker C, Vives L, Mohajeri K, Witherspoon K, et al. Excess of rare, inherited truncating mutations in autism. Nat Genet 2015;47:582–8. https://doi.org/10.1038/ng.3303.

[101] Sanders SJ, He X, Willsey AJ, Ercan-Sencicek AG, Samocha KE, Cicek AE, et al. Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci. Neuron 2015;87:1215–33. https://doi.org/10.1016/j.neuron.2015.09.016.

[102] Wang T, Guo H, Xiong B, Stessman HAF, Wu H, Coe BP, et al. De novo genic mutations among a Chinese autism spectrum disorder cohort. Nat Commun 2016;7:13316. https://doi.org/10.1038/ncomms13316.

[103] Stessman HAF, Xiong B, Coe BP, Wang T, Hoekzema K, Fenckova M, et al. Targeted sequencing identifies 91 neurodevelopmental-disorder risk genes with autism and developmental-disability biases. Nat Genet 2017;49:515–26. https://doi.org/10.1038/ng.3792.

[104] Satterstrom FK, Kosmicki JA, Wang J, Breen MS, De Rubeis S, An J-Y, et al. Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. Cell 2020;180:568-584.e23. https://doi.org/10.1016/j.cell.2019.12.036.

[105] Fu JM, Satterstrom FK, Peng M, Brand H, Collins RL, Dong S, et al. Rare coding variation provides insight into the genetic architecture and phenotypic context of autism. Nat Genet 2022;54:1320–31. https://doi.org/10.1038/s41588-022-01104-0.

[106] Deciphering Developmental Disorders Study. Prevalence and architecture of de novo mutations in developmental disorders. Nature 2017;542:433–8. https://doi.org/10.1038/nature21062.

[107] Faundes V, Newman WG, Bernardini L, Canham N, Clayton-Smith J, Dallapiccola B, et al. Histone Lysine Methylases and Demethylases in the Landscape of Human Developmental Disorders. The American Journal of Human Genetics 2018;102:175–87. https://doi.org/10.1016/j.ajhg.2017.11.013.

[108] Tammimies K, Marshall CR, Walker S, Kaur G, Thiruvahindrapuram B, Lionel AC, et al. Molecular Diagnostic Yield of Chromosomal Microarray Analysis and Whole-Exome Sequencing in Children With Autism Spectrum Disorder. JAMA 2015;314:895. https://doi.org/10.1001/jama.2015.10078.

[109] Aspromonte MC, Bellini M, Gasparini A, Carraro M, Bettella E, Polli R, et al. Characterization of intellectual disability and autism comorbidity through gene

panel sequencing. Human Mutation 2019;40:1346–63. https://doi.org/10.1002/humu.23822.

[110] Tang S, Addis L, Smith A, Topp SD, Pendziwiat M, Mei D, et al. Phenotypic and genetic spectrum of epilepsy with myoclonic atonic seizures. Epilepsia 2020;61:995–1007. https://doi.org/10.1111/epi.16508.

[111]Zhu T, Liang C, Li D, Tian M, Liu S, Gao G, et al. Histone methyltransferase Ash1L mediates activity-dependent repression of neurexin-1α. Sci Rep 2016;6:26597. https://doi.org/10.1038/srep26597.

[112] Liu S, Tian M, He F, Li J, Xie H, Liu W, et al. Mutations in ASH1L confer susceptibility to Tourette syndrome. Mol Psychiatry 2020;25:476–90. https://doi.org/10.1038/s41380-019-0560-8.

[113] Liu W, Xu L, Zhang C, Shen L, Dong J, Zhang H, et al. *ASH1L* may contribute to the risk of Tourette syndrome: Combination of family-based analysis and case–control study. Brain and Behavior 2022;12:e2539. https://doi.org/10.1002/brb3.2539.

[114] C Yuen RK, Merico D, Bookman M, L Howe J, Thiruvahindrapuram B, Patel RV, et al. Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. Nat Neurosci 2017;20:602–11. https://doi.org/10.1038/nn.4524.

[115] Hu C, Wang Y, Li C, Mei L, Zhou B, Li D, et al. Targeted sequencing and clinical strategies in children with autism spectrum disorder: A cohort study. Front Genet 2023;14:1083779. https://doi.org/10.3389/fgene.2023.1083779.

[116] Li D, Choque Olsson N, Becker M, Arora A, Jiao H, Norgren N, et al. Rare variants in the outcome of social skills group training for autism. Autism Research 2022;15:434–46. https://doi.org/10.1002/aur.2666.

[117] Mitani T, Isikay S, Gezdirici A, Gulec EY, Punetha J, Fatih JM, et al. High prevalence of multilocus pathogenic variation in neurodevelopmental disorders in the Turkish population. The American Journal of Human Genetics 2021;108:1981–2005. https://doi.org/10.1016/j.ajhg.2021.08.009.

[118] Wang T, Hoekzema K, Vecchio D, Wu H, Sulovari A, Coe BP, et al. Large-scale targeted sequencing identifies risk genes for neurodevelopmental disorders. Nat Commun 2020;11:4932. https://doi.org/10.1038/s41467-020-18723-y.

[119] Zhang Y, Li Y, Guo R, Xu W, Liu X, Zhao C, et al. Genetic diagnostic yields of 354 Chinese ASD children with rare mutations by a pipeline of genomic tests. Front Genet 2023;14:1108440. https://doi.org/10.3389/fgene.2023.1108440.

[120] Zhou X, Feliciano P, Shu C, Wang T, Astrovskaya I, Hall JB, et al. Integrating de novo and inherited variants in 42,607 autism cases identifies mutations in new

moderate-risk genes. Nat Genet 2022;54:1305–19. https://doi.org/10.1038/s41588-022-01148-2.

[121] Okamoto N, Miya F, Tsunoda T, Kato M, Saitoh S, Yamasaki M, et al. Novel MCA/ID syndrome with *ASH1L* mutation. American J of Med Genetics Pt A 2017;173:1644–8. https://doi.org/10.1002/ajmg.a.38193.

[122] Shen W, Krautscheid P, Rutz AM, Bayrak-Toydemir P, Dugan SL. De novo loss-of-function variants of ASH1L are associated with an emergent neurodevelopmental disorder. European Journal of Medical Genetics 2019;62:55–60. https://doi.org/10.1016/j.ejmg.2018.05.003.

[123] Xi H, Peng Y, Xie W, Pang J, Ma N, Yang S, et al. A chromosome 1q22 microdeletion including ASH1L is associated with intellectual disability in a Chinese family. Mol Cytogenet 2020;13:20. https://doi.org/10.1186/s13039-020-00483-5.

[124] Liu H, Liu D-T, Lan S, Yang Y, Huang J, Huang J, et al. ASH1L mutation caused seizures and intellectual disability in twin sisters. Journal of Clinical Neuroscience 2021;91:69–74. https://doi.org/10.1016/j.jocn.2021.06.038.

[125] Homsy J, Zaidi S, Shen Y, Ware JS, Samocha KE, Karczewski KJ, et al. De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies. Science 2015;350:1262–6. https://doi.org/10.1126/science.aac9396.

[126] Iossifov I, O'Roak BJ, Sanders SJ, Ronemus M, Krumm N, Levy D, et al. The contribution of de novo coding mutations to autism spectrum disorder. Nature 2014;515:216–21. https://doi.org/10.1038/nature13908.

[127] Sanchis-Juan A, Megy K, Stephens J, Armirola Ricaurte C, Dewhurst E, Low K, et al. Genome sequencing and comprehensive rare-variant analysis of 465 families with neurodevelopmental disorders. The American Journal of Human Genetics 2023;110:1343–55. https://doi.org/10.1016/j.ajhg.2023.07.007.

[128] Sheth F, Shah J, Jain D, Shah S, Patel H, Patel K, et al. Comparative yield of molecular diagnostic algorithms for autism spectrum disorder diagnosis in India: evidence supporting whole exome sequencing as first tier test. BMC Neurol 2023;23:292. https://doi.org/10.1186/s12883-023-03341-0.

[129] Willsey AJ, Sanders SJ, Li M, Dong S, Tebbenkamp AT, Muhle RA, et al. Coexpression Networks Implicate Human Midfetal Deep Cortical Projection Neurons in the Pathogenesis of Autism. Cell 2013;155:997–1007. https://doi.org/10.1016/j.cell.2013.10.020.

[130] Schaaf CP, Betancur C, Yuen RKC, Parr JR, Skuse DH, Gallagher L, et al. A framework for an evidence-based gene list relevant to autism spectrum disorder. Nat Rev Genet 2020;21:367–76. https://doi.org/10.1038/s41576-020-0231-2.

[131] ASH1L | gnomAD v4.0.0 | gnomAD n.d. https://gnomad.broadinstitute.org/gene/ENSG00000116539?dataset=gnomad_r4 (accessed November 13, 2023).

[132] Liu W, Xie Y, Ma J, Luo X, Nie P, Zuo Z, et al. IBS: an illustrator for the presentation and visualization of biological sequences. Bioinformatics 2015;31:3359–61. https://doi.org/10.1093/bioinformatics/btv362.

[133] Gao Y, Duque-Wilckens N, Aljazi MB, Wu Y, Moeser AJ, Mias GI, et al. Loss of histone methyltransferase ASH1L in the developing mouse brain causes autistic-like behaviors. Commun Biol 2021;4:756. https://doi.org/10.1038/s42003-021-02282-z.

[134] Gao Y, Aljazi MB, Wu Y, He J. Vorinostat, a histone deacetylase inhibitor, ameliorates the sociability and cognitive memory in an Ash1L-deletion-induced ASD/ID mouse model. Neuroscience Letters 2021;764:136241. https://doi.org/10.1016/j.neulet.2021.136241.

[135] Gao Y, Aljazi MB, He J. Neural Hyperactivity Is a Core Pathophysiological Change Induced by Deletion of a High Autism Risk Gene Ash1L in the Mouse Brain. Front Behav Neurosci 2022;16:873466. https://doi.org/10.3389/fnbeh.2022.873466.

[136] Brinkmeier ML, Geister KA, Jones M, Waqas M, Maillard I, Camper SA. The Histone Methyltransferase Gene Absent, Small, or Homeotic Discs-1 Like Is Required for Normal Hox Gene Expression and Fertility in Mice1. Biology of Reproduction 2015;93. https://doi.org/10.1095/biolreprod.115.131516.

[137] Jones M, Chase J, Brinkmeier M, Xu J, Weinberg DN, Schira J, et al. Ash1l controls quiescence and self-renewal potential in hematopoietic stem cells. J Clin Invest 2015;125:2007–20. https://doi.org/10.1172/JCI78124.

[138] Galichet C, Lovell-Badge R, Rizzoti K. Nestin-Cre Mice Are Affected by Hypopituitarism, Which Is Not Due to Significant Activity of the Transgene in the Pituitary Gland. PLoS ONE 2010;5:e11443. https://doi.org/10.1371/journal.pone.0011443.

[139] Xia M, Liu J, Wu X, Liu S, Li G, Han C, et al. Histone Methyltransferase Ash1l Suppresses Interleukin-6 Production and Inflammatory Autoimmune Diseases by Inducing the Ubiquitin-Editing Enzyme A20. Immunity 2013;39:470–81. https://doi.org/10.1016/j.immuni.2013.08.016.

[140] Yin B, Yu F, Wang C, Li B, Liu M, Ye L. Epigenetic Control of Mesenchymal Stem Cell Fate Decision via Histone Methyltransferase Ash1l. Stem Cells 2019;37:115–27. https://doi.org/10.1002/stem.2918.

[141] Horan GS, Ramírez-Solis R, Featherstone MS, Wolgemuth DJ, Bradley A, Behringer RR. Compound mutants for the paralogous hoxa-4, hoxb-4, and hoxd-4 genes show more complete homeotic transformations and a dose-dependent

increase in the number of vertebrae transformed. Genes Dev 1995;9:1667–77. https://doi.org/10.1101/gad.9.13.1667.

[142] Hanson RD, Hess JL, Yu BD, Ernst P, Van Lohuizen M, Berns A, et al. Mammalian *Trithorax* and *Polycomb* -group homologues are antagonistic regulators of homeotic development. Proc Natl Acad Sci USA 1999;96:14372–7. https://doi.org/10.1073/pnas.96.25.14372.

[143] Favier B. Developmental functions of mammalian Hox genes. Molecular Human Reproduction 1997;3:115–31. https://doi.org/10.1093/molehr/3.2.115.

[144] Buffone MG, Ijiri TW, Cao W, Merdiushev T, Aghajanian HK, Gerton GL. Heads or tails? Structural events and molecular mechanisms that promote mammalian sperm acrosomal exocytosis and motility. Molecular Reproduction Devel 2012;79:4–18. https://doi.org/10.1002/mrd.21393.

[145] Deinhardt K, Chao MV. Shaping neurons: Long and short range effects of mature and proBDNF signalling upon neuronal structure. Neuropharmacology 2014;76:603–9. https://doi.org/10.1016/j.neuropharm.2013.04.054.

[146] Cheon S, Culver AM, Bagnell AM, Ritchie FD, Vacharasin JM, McCord MM, et al. Counteracting epigenetic mechanisms regulate the structural development of neuronal circuitry in human neurons. Mol Psychiatry 2022;27:2291–303. https://doi.org/10.1038/s41380-022-01474-1.

[147] Crawley JN. Exploratory behavior models of anxiety in mice. Neuroscience & Biobehavioral Reviews 1985;9:37–44. https://doi.org/10.1016/0149-7634(85)90030-2.

[148] Seibenhener ML, Wooten MC. Use of the Open Field Maze to Measure Locomotor and Anxiety-like Behavior in Mice. JoVE 2015:52434. https://doi.org/10.3791/52434.

[149] Angoa-Pérez M, Kane MJ, Briggs DI, Francescutti DM, Kuhn DM. Marble Burying and Nestlet Shredding as Tests of Repetitive, Compulsive-like Behaviors in Mice. JoVE 2013:50978. https://doi.org/10.3791/50978.

[150] Crawley JN. Mouse Behavioral Assays Relevant to the Symptoms of Autism*. Brain Pathology 2007;17:448–59. https://doi.org/10.1111/j.1750-3639.2007.00096.x.

[151] Walf AA, Frye CA. The use of the elevated plus maze as an assay of anxiety-related behavior in rodents. Nat Protoc 2007;2:322–8. https://doi.org/10.1038/nprot.2007.44.

[152] Bergado-Acosta JR, Sangha S, Narayanan RT, Obata K, Pape H-C, Stork O. Critical role of the 65-kDa isoform of glutamic acid decarboxylase in consolidation

and generalization of Pavlovian fear memory. Learn Mem 2008;15:163–71. https://doi.org/10.1101/lm.705408.

[153] McHugh TJ, Jones MW, Quinn JJ, Balthasar N, Coppari R, Elmquist JK, et al. Dentate Gyrus NMDA Receptors Mediate Rapid Pattern Separation in the Hippocampal Network. Science 2007;317:94–9. https://doi.org/10.1126/science.1140263.

[154] Han Z, Zhang X, Zhu J, Chen Y, Li CT. High-Throughput Automatic Training System for Odor-Based Learned Behaviors in Head-Fixed Mice. Front Neural Circuits 2018;12:15. https://doi.org/10.3389/fncir.2018.00015.

[155] Lalonde R, Strazielle C. Brain regions and genes affecting limb-clasping responses. Brain Research Reviews 2011;67:252–9. https://doi.org/10.1016/j.brainresrev.2011.02.005.

[156] Pasciuto E, Borrie SC, Kanellopoulos AK, Santos AR, Cappuyns E, D'Andrea L, et al. Autism Spectrum Disorders: Translating human deficits into mouse behavior. Neurobiology of Learning and Memory 2015;124:71–87. https://doi.org/10.1016/j.nlm.2015.07.013.

[157] Li G, Ye Z, Shi C, Sun L, Han M, Zhuang Y, et al. The Histone Methyltransferase Ash1l is Required for Epidermal Homeostasis in Mice. Sci Rep 2017;7:45401. https://doi.org/10.1038/srep45401.

[158] Foss-Feig JH, Adkinson BD, Ji JL, Yang G, Srihari VH, McPartland JC, et al. Searching for Cross-Diagnostic Convergence: Neural Mechanisms Governing Excitation and Inhibition Balance in Schizophrenia and Autism Spectrum Disorders. Biological Psychiatry 2017;81:848–61. https://doi.org/10.1016/j.biopsych.2017.03.005.

[159] Kalueff AV, Stewart AM, Song C, Berridge KC, Graybiel AM, Fentress JC. Neurobiology of rodent self-grooming and its value for translational neuroscience. Nat Rev Neurosci 2016;17:45–59. https://doi.org/10.1038/nrn.2015.8.

[160] Singer HS, Szymanski S, Giuliano J, Yokoi F, Dogan AS, Brasic JR, et al. Elevated Intrasynaptic Dopamine Release in Tourette's Syndrome Measured by PET. AJP 2002;159:1329–36. https://doi.org/10.1176/appi.ajp.159.8.1329.

[161] Stoner R, Chow ML, Boyle MP, Sunkin SM, Mouton PR, Roy S, et al. Patches of Disorganization in the Neocortex of Children with Autism. N Engl J Med 2014;370:1209–19. https://doi.org/10.1056/NEJMoa1307491.

[162] Gregory GD, Vakoc CR, Rozovskaia T, Zheng X, Patel S, Nakamura T, et al. Mammalian ASH1L Is a Histone Methyltransferase That Occupies the Transcribed Region of Active Genes. Molecular and Cellular Biology 2007;27:8466–79. https://doi.org/10.1128/MCB.00993-07.

[163] Zhang T, Ren T, Lin H, Tong Y, Zhang J, Nie J, et al. ASH1L contributes to oocyte apoptosis by regulating DNA damage. American Journal of Physiology-Cell Physiology 2022;323:C1264–73. https://doi.org/10.1152/ajpcell.00196.2022.

[164] Ding X, Liu S, Tian M, Zhang W, Zhu T, Li D, et al. Activity-induced histone modifications govern Neurexin-1 mRNA splicing and memory preservation. Nat Neurosci 2017;20:690–9. https://doi.org/10.1038/nn.4536.

[165] Husmann D, Gozani O. Histone lysine methyltransferases in biology and disease. Nat Struct Mol Biol 2019;26:880–9. https://doi.org/10.1038/s41594-019-0298-7.

[166] Rogawski DS, Grembecka J, Cierpicki T. H3K36 methyltransferases as cancer drug targets: rationale and perspectives for inhibitor development. Future Medicinal Chemistry 2016;8:1589–607. https://doi.org/10.4155/fmc-2016-0071.

[167] Farkas LM, Huttner WB. The cell biology of neural stem and progenitor cells and its significance for their proliferation versus differentiation during mammalian brain development. Current Opinion in Cell Biology 2008;20:707–15. https://doi.org/10.1016/j.ceb.2008.09.008.

[168] Homem CCF, Repic M, Knoblich JA. Proliferation control in neural stem and progenitor cells. Nat Rev Neurosci 2015;16:647–59. https://doi.org/10.1038/nrn4021.

[169] Gage FH. Mammalian Neural Stem Cells. Science 2000;287:1433–8. https://doi.org/10.1126/science.287.5457.1433.

[170] Corbin JG, Gaiano N, Juliano SL, Poluch S, Stancik E, Haydar TF. Regulation of neural progenitor cell development in the nervous system. Journal of Neurochemistry 2008;106:2272–87. https://doi.org/10.1111/j.1471-4159.2008.05522.x.

[171] Berdasco M, Esteller M. Aberrant Epigenetic Landscape in Cancer: How Cellular Identity Goes Awry. Developmental Cell 2010;19:698–711. https://doi.org/10.1016/j.devcel.2010.10.005.

[172] Holmberg J, Perlmann T. Maintaining differentiated cellular identity. Nat Rev Genet 2012;13:429–39. https://doi.org/10.1038/nrg3209.

[173] Zikopoulos B, Barbas H. Altered neural connectivity in excitatory and inhibitory cortical circuits in autism. Front Hum Neurosci 2013;7. https://doi.org/10.3389/fnhum.2013.00609.

[174] Friederici AD. The Brain Basis of Language Processing: From Structure to Function. Physiological Reviews 2011;91:1357–92. https://doi.org/10.1152/physrev.00006.2011.

[175] Sestieri C, Shulman GL, Corbetta M. The contribution of the human posterior parietal cortex to episodic memory. Nat Rev Neurosci 2017;18:183–92. https://doi.org/10.1038/nrn.2017.6.

[176] Hevner RF. Layer-Specific Markers as Probes for Neuron Type Identity in Human Neocortex and Malformations of Cortical Development: Journal of Neuropathology and Experimental Neurology 2007;66:101–9. https://doi.org/10.1097/nen.0b013e3180301c06.

[177] Alcamo EA, Chirivella L, Dautzenberg M, Dobreva G, Fariñas I, Grosschedl R, et al. Satb2 Regulates Callosal Projection Neuron Identity in the Developing Cerebral Cortex. Neuron 2008;57:364–77. https://doi.org/10.1016/j.neuron.2007.12.012.

[178] Arlotta P, Hobert O. Homeotic Transformations of Neuronal Cell Identities. Trends in Neurosciences 2015;38:751–62. https://doi.org/10.1016/j.tins.2015.10.005.

[179] Kala K, Haugas M, Lilleväli K, Guimera J, Wurst W, Salminen M, et al. *Gata2* is a tissue-specific post-mitotic selector gene for midbrain GABAergic neurons. Development 2009;136:253–62. https://doi.org/10.1242/dev.029900.

[180] Grimaud C, Nègre N, Cavalli G. From genetics to epigenetics: the tale of Polycomb group and trithorax group genes. Chromosome Res 2006;14:363–75. https://doi.org/10.1007/s10577-006-1069-y.

[181] Testa G, Schaft J, Van Der Hoeven F, Glaser S, Anastassiadis K, Zhang Y, et al. A reliable lacZ expression reporter cassette for multipurpose, knockout-first alleles. Genesis 2004;38:151–8. https://doi.org/10.1002/gene.20012.

[182] Kranz A, Fu J, Duerschke K, Weidlich S, Naumann R, Stewart AF, et al. An improved Flp deleter mouse in C57Bl/6 based on Flpo recombinase. Genesis 2010;48:512–20. https://doi.org/10.1002/dvg.20641.

[183] Lakso M, Pichel JG, Gorman JR, Sauer B, Okamoto Y, Lee E, et al. Efficient in vivo manipulation of mouse genomic sequences at the zygote stage. Proc Natl Acad Sci USA 1996;93:5860–5. https://doi.org/10.1073/pnas.93.12.5860.

[184] Kessaris N, Fogarty M, Iannarelli P, Grist M, Wegner M, Richardson WD. Competing waves of oligodendrocytes in the forebrain and postnatal elimination of an embryonic lineage. Nat Neurosci 2006;9:173–9. https://doi.org/10.1038/nn1620.

[185] Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucl Acids Res 1988;16:1215–1215. https://doi.org/10.1093/nar/16.3.1215.

[186] Greco TL, Takada S, Newhouse MM, McMahon JA, McMahon AP, Camper SA. Analysis of the vestigial tail mutation demonstrates that Wnt-3a gene dosage

regulates mouse axial development. Genes Dev 1996;10:313–24. https://doi.org/10.1101/gad.10.3.313.

[187] Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods 2012;9:671–5. https://doi.org/10.1038/nmeth.2089.

[188] Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. Nat Methods 2012;9:676–82. https://doi.org/10.1038/nmeth.2019.

[189] Ollion J, Cochennec J, Loll F, Escudé C, Boudier T. TANGO: a generic tool for high-throughput 3D image analysis for studying nuclear organization. Bioinformatics 2013;29:1840–1. https://doi.org/10.1093/bioinformatics/btt276.

[190] McGrath B, Wu P, Salvi S, Girgla N, Chen X, Zhu J, et al. ASXL3 controls cortical neuron fate specification through extrinsic self-renewal pathways. Neuroscience; 2021. https://doi.org/10.1101/2021.07.20.452995.

[191] Gierahn TM, Wadsworth MH, Hughes TK, Bryson BD, Butler A, Satija R, et al. Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput. Nat Methods 2017;14:395–8. https://doi.org/10.1038/nmeth.4179.

[192] Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, et al. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell 2015;161:1202–14. https://doi.org/10.1016/j.cell.2015.05.002.

[193] Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15–21. https://doi.org/10.1093/bioinformatics/bts635.

[194] Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, et al. Comprehensive Integration of Single-Cell Data. Cell 2019;177:1888-1902.e21. https://doi.org/10.1016/j.cell.2019.05.031.

[195] Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. Cell 2021;184:3573-3587.e29. https://doi.org/10.1016/j.cell.2021.04.048.

[196] Lodato S, Arlotta P. Generating Neuronal Diversity in the Mammalian Cerebral Cortex. Annu Rev Cell Dev Biol 2015;31:699–720. https://doi.org/10.1146/annurev-cellbio-100814-125353.

[197] Loo L, Simon JM, Xing L, McCoy ES, Niehaus JK, Guo J, et al. Single-cell transcriptomic analysis of mouse neocortical development. Nat Commun 2019;10:134. https://doi.org/10.1038/s41467-018-08079-9.

[198] Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics 2013;14:128. https://doi.org/10.1186/1471-2105-14-128.

[199] Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res 2016;44:W90–7. https://doi.org/10.1093/nar/gkw377.

[200] Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, et al. The single-cell transcriptional landscape of mammalian organogenesis. Nature 2019;566:496–502. https://doi.org/10.1038/s41586-019-0969-x.

[201] McInnes L, Healy J, Melville J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction 2020.

[202] Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, et al. Reversed graph embedding resolves complex single-cell trajectories. Nat Methods 2017;14:979–82. https://doi.org/10.1038/nmeth.4402.

[203] Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol 2014;32:381–6. https://doi.org/10.1038/nbt.2859.

[204] Ge SX, Jung D, Yao R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. Bioinformatics 2020;36:2628–9. https://doi.org/10.1093/bioinformatics/btz931.

[205] Hubert KA, Wellik DM. Hox genes in development and beyond. Development 2023;150:dev192476. https://doi.org/10.1242/dev.192476.

[206] McCaffery P, Deutsch CK. Macrocephaly and the control of brain growth in autistic disorders. Progress in Neurobiology 2005;77:38–56. https://doi.org/10.1016/j.pneurobio.2005.10.005.

[207] Sacco R, Gabriele S, Persico AM. Head circumference and brain size in autism spectrum disorder: A systematic review and meta-analysis. Psychiatry Research: Neuroimaging 2015;234:239–51. https://doi.org/10.1016/j.pscychresns.2015.08.016.

[208] Hashem S, Nisar S, Bhat AA, Yadav SK, Azeem MW, Bagga P, et al. Genetics of structural and functional brain changes in autism spectrum disorder. Transl Psychiatry 2020;10:229. https://doi.org/10.1038/s41398-020-00921-3.

[209] Minshew NJ, Williams DL. The New Neurobiology of Autism: Cortex, Connectivity, and Neuronal Organization. Arch Neurol 2007;64:945. https://doi.org/10.1001/archneur.64.7.945.

[210] Pucilowska J, Vithayathil J, Tavares EJ, Kelly C, Karlo JC, Landreth GE. The 16p11.2 Deletion Mouse Model of Autism Exhibits Altered Cortical Progenitor Proliferation and Brain Cytoarchitecture Linked to the ERK MAPK Pathway. Journal of Neuroscience 2015;35:3190–200. https://doi.org/10.1523/JNEUROSCI.4864-13.2015.

[211] Matsumura K, Seiriki K, Okada S, Nagase M, Ayabe S, Yamada I, et al. Pathogenic POGZ mutation causes impaired cortical development and reversible autism-like phenotypes. Nat Commun 2020;11:859. https://doi.org/10.1038/s41467-020-14697-z.

[212] Zurkirchen L, Varum S, Giger S, Klug A, Häusel J, Bossart R, et al. Yin Yang 1 sustains biosynthetic demands during brain development in a stage-specific manner. Nat Commun 2019;10:2192. https://doi.org/10.1038/s41467-019-09823-5.

[213] Gao P, Postiglione MP, Krieger TG, Hernandez L, Wang C, Han Z, et al. Deterministic Progenitor Behavior and Unitary Production of Neurons in the Neocortex. Cell 2014;159:775–88. https://doi.org/10.1016/j.cell.2014.10.027.

[214] Clegg JM, Parkin HM, Mason JO, Pratt T. Heparan Sulfate Sulfation by Hs2st Restricts Astroglial Precursor Somal Translocation in Developing Mouse Forebrain by a Non-Cell-Autonomous Mechanism. J Neurosci 2019;39:1386–404. https://doi.org/10.1523/JNEUROSCI.1747-17.2018.

[215] Young KM, Fogarty M, Kessaris N, Richardson WD. Subventricular Zone Stem Cells Are Heterogeneous with Respect to Their Embryonic Origins and Neurogenic Fates in the Adult Olfactory Bulb. J Neurosci 2007;27:8286–96. https://doi.org/10.1523/JNEUROSCI.0476-07.2007.

[216] Lv X, Ren S-Q, Zhang X-J, Shen Z, Ghosh T, Xianyu A, et al. TBR2 coordinates neurogenesis expansion and precise microcircuit organization via Protocadherin 19 in the mammalian cortex. Nat Commun 2019;10:3946. https://doi.org/10.1038/s41467-019-11854-x.

[217] Georgala PA, Manuel M, Price DJ. The Generation of Superficial Cortical Layers Is Regulated by Levels of the Transcription Factor Pax6. Cerebral Cortex 2011;21:81–94. https://doi.org/10.1093/cercor/bhq061.

[218] Quintana-Urzainqui I, Kozić Z, Mitra S, Tian T, Manuel M, Mason JO, et al. Tissue-Specific Actions of Pax6 on Proliferation and Differentiation Balance in Developing Forebrain Are Foxg1 Dependent. iScience 2018;10:171–91. https://doi.org/10.1016/j.isci.2018.11.031.

[219] Agirman G, Broix L, Nguyen L. Cerebral cortex development: an outside-in perspective. FEBS Lett 2017;591:3978–92. https://doi.org/10.1002/1873-3468.12924.

[220] Britanova O, De Juan Romero C, Cheung A, Kwan KY, Schwark M, Gyorgy A, et al. Satb2 Is a Postmitotic Determinant for Upper-Layer Neuron Specification in the Neocortex. Neuron 2008;57:378–92. https://doi.org/10.1016/j.neuron.2007.12.028.

[221] McKenna WL, Ortiz-Londono CF, Mathew TK, Hoang K, Katzman S, Chen B. Mutual regulation between *Satb2* and *Fezf2* promotes subcerebral projection neuron identity in the developing cerebral cortex. Proc Natl Acad Sci USA 2015;112:11702–7. https://doi.org/10.1073/pnas.1504144112.

[222] Kwan KY, Šestan N, Anton ES. Transcriptional co-regulation of neuronal migration and laminar identity in the neocortex. Development 2012;139:1535–46. https://doi.org/10.1242/dev.069963.

[223] Arlotta P, Molyneaux BJ, Chen J, Inoue J, Kominami R, Macklis JD. Neuronal Subtype-Specific Genes that Control Corticospinal Motor Neuron Development In Vivo. Neuron 2005;45:207–21. https://doi.org/10.1016/j.neuron.2004.12.036.

[224] Chen B, Wang SS, Hattox AM, Rayburn H, Nelson SB, McConnell SK. The *Fezf2*–*Ctip2* genetic pathway regulates the fate choice of subcortical projection neurons in the developing cerebral cortex. Proc Natl Acad Sci USA 2008;105:11382–7. https://doi.org/10.1073/pnas.0804918105.

[225] Tamada K, Fukumoto K, Toya T, Nakai N, Awasthi JR, Tanaka S, et al. Genetic dissection identifies Necdin as a driver gene in a mouse model of paternal 15q duplications. Nat Commun 2021;12:4056. https://doi.org/10.1038/s41467-021-24359-3.

[226] Li Q, Wang L, Ma Y, Yue W, Zhang D, Li J. P-Rex1 Overexpression Results in Aberrant Neuronal Polarity and Psychosis-Related Behaviors. Neurosci Bull 2019;35:1011–23. https://doi.org/10.1007/s12264-019-00408-2.

[227] Sun Y, Ma L. New Insights into Long Non-Coding RNA MALAT1 in Cancer and Metastasis. Cancers 2019;11:216. https://doi.org/10.3390/cancers11020216.

[228] Renaudineau S, Poucet B, Laroche S, Davis S, Save E. Impaired long-term stability of CA1 place cell representation in mice lacking the transcription factor *zif268* / *egr1*. Proc Natl Acad Sci USA 2009;106:11771–5. https://doi.org/10.1073/pnas.0900484106.

[229] Veyrac A, Gros A, Bruel-Jungerman E, Rochefort C, Kleine Borgmann FB, Jessberger S, et al. *Zif268* / *egr1* gene controls the selection, maturation and functional integration of adult hippocampal newborn neurons by learning. Proc Natl Acad Sci USA 2013;110:7062–7. https://doi.org/10.1073/pnas.1220558110.

[230] Zhang C, Xu L, Zheng X, Liu S, Che F. Role of *Ash1l* in Tourette syndrome and other neurodevelopmental disorders. Developmental Neurobiology 2021;81:79–91. https://doi.org/10.1002/dneu.22795.

[231] Achilleos A, Trainor PA. Mouse Models of Rare Craniofacial Disorders. Current Topics in Developmental Biology, vol. 115, Elsevier; 2015, p. 413–58. https://doi.org/10.1016/bs.ctdb.2015.07.011.

[232] Sakai D, Dixon J, Achilleos A, Dixon M, Trainor PA. Prevention of Treacher Collins syndrome craniofacial anomalies in mouse models via maternal antioxidant supplementation. Nat Commun 2016;7:10328. https://doi.org/10.1038/ncomms10328.

[233] Li Z, Klein JA, Rampam S, Kurzion R, Campbell NB, Patel Y, et al. Asynchronous excitatory neuron development in an isogenic cortical spheroid model of Down syndrome. Front Neurosci 2022;16:932384. https://doi.org/10.3389/fnins.2022.932384.

[234] Tuorto F, Liebers R, Musch T, Schaefer M, Hofmann S, Kellner S, et al. RNA cytosine methylation by Dnmt2 and NSun2 promotes tRNA stability and protein synthesis. Nat Struct Mol Biol 2012;19:900–5. https://doi.org/10.1038/nsmb.2357.

[235] Ponti G, Obernier K, Guinto C, Jose L, Bonfanti L, Alvarez-Buylla A. Cell cycle and lineage progression of neural progenitors in the ventricular-subventricular zones of adult mice. Proc Natl Acad Sci USA 2013;110. https://doi.org/10.1073/pnas.1219563110.

[236] Mitchell-Dick A, Chalem A, Pilaz L-J, Silver DL. Acute Lengthening of Progenitor Mitosis Influences Progeny Fate during Cortical Development in vivo. Dev Neurosci 2019;41:300–17. https://doi.org/10.1159/000507113.

[237] Paulsen B, Velasco S, Kedaigle AJ, Pigoni M, Quadrato G, Deo AJ, et al. Autism genes converge on asynchronous development of shared neuron classes. Nature 2022;602:268–73. https://doi.org/10.1038/s41586-021-04358-6.

[238] Thomas LA, Akins MR, Biederer T. Expression and adhesion profiles of SynCAM molecules indicate distinct neuronal functions. J of Comparative Neurology 2008;510:47–67. https://doi.org/10.1002/cne.21773.

[239] Worley P, Zeng W, Huang G, Kim J, Shin D, Kim M, et al. Homer proteins in Ca2+ signaling by excitable and non-excitable cells. Cell Calcium 2007;42:363–71. https://doi.org/10.1016/j.ceca.2007.05.007.

[240] Williams CG, Lee HJ, Asatsuma T, Vento-Tormo R, Haque A. An introduction to spatial transcriptomics for biomedical research. Genome Med 2022;14:68. https://doi.org/10.1186/s13073-022-01075-1.

[241] Rodriques SG, Stickels RR, Goeva A, Martin CA, Murray E, Vanderburg CR, et al. Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution 2019.

[242] Stickels RR, Murray E, Kumar P, Li J, Marshall JL, Di Bella DJ, et al. Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2. Nat Biotechnol 2021;39:313–9. https://doi.org/10.1038/s41587-020-0739-1.

[243] Bhattacherjee A, Zhang C, Watson BR, Djekidel MN, Moffitt JR, Zhang Y. Spatial transcriptomics reveals the distinct organization of mouse prefrontal cortex and neuronal subtypes regulating chronic pain. Nat Neurosci 2023;26:1880–93. https://doi.org/10.1038/s41593-023-01455-9.

[244] Bayraktar OA, Bartels T, Holmqvist S, Kleshchevnikov V, Martirosyan A, Polioudakis D, et al. Astrocyte layers in the mammalian cerebral cortex revealed by a single-cell in situ transcriptomic map. Nat Neurosci 2020;23:500–9. https://doi.org/10.1038/s41593-020-0602-1.

[245] Panday A, Elango R, Willis NA, Scully R. A modified CUT&RUN-seq technique for qPCR analysis of chromatin-protein interactions. STAR Protocols 2022;3:101529. https://doi.org/10.1016/j.xpro.2022.101529.

[246] Buenrostro JD, Wu B, Litzenburger UM, Ruff D, Gonzales ML, Snyder MP, et al. Single-cell chromatin accessibility reveals principles of regulatory variation. Nature 2015;523:486–90. https://doi.org/10.1038/nature14590.

[247] Cusanovich DA, Daza R, Adey A, Pliner HA, Christiansen L, Gunderson KL, et al. Multiplex single-cell profiling of chromatin accessibility by combinatorial cellular indexing. Science 2015;348:910–4. https://doi.org/10.1126/science.aab1601.

[248] Dehay C, Kennedy H. Cell-cycle control and cortical development. Nat Rev Neurosci 2007;8:438–50. https://doi.org/10.1038/nrn2097.

[249] Livet J, Weissman TA, Kang H, Draft RW, Lu J, Bennis RA, et al. Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. Nature 2007;450:56–62. https://doi.org/10.1038/nature06293.

[250] Weissman TA, Pan YA. Brainbow: New Resources and Emerging Biological Applications for Multicolor Genetic Labeling and Analysis. Genetics 2015;199:293–306. https://doi.org/10.1534/genetics.114.172510.

[251] Snippert HJ, Van Der Flier LG, Sato T, Van Es JH, Van Den Born M, Kroon-Veenboer C, et al. Intestinal Crypt Homeostasis Results from Neutral Competition between Symmetrically Dividing Lgr5 Stem Cells. Cell 2010;143:134–44. https://doi.org/10.1016/j.cell.2010.09.016.

[252] Pilz G-A, Shitamukai A, Reillo I, Pacary E, Schwausch J, Stahl R, et al. Amplification of progenitors in the mammalian telencephalon includes a new radial glial cell type. Nat Commun 2013;4:2125. https://doi.org/10.1038/ncomms3125.

[253] Zhu J, Liu C, Huang X, Van De Leemput J, Lee H, Han Z. H3K36 Di-Methylation Marks, Mediated by Ash1 in Complex with Caf1-55 and MRG15, Are Required

during Drosophila Heart Development. JCDD 2023;10:307. https://doi.org/10.3390/jcdd10070307.

[254] Sohal DS, Nghiem M, Crackower MA, Witt SA, Kimball TR, Tymitz KM, et al. Temporally Regulated and Tissue-Specific Gene Manipulations in the Adult and Embryonic Heart Using a Tamoxifen-Inducible Cre Protein. Circulation Research 2001;89:20–5. https://doi.org/10.1161/hh1301.092687.

[255] Namiki T, Kamoshita M, Kageyama A, Terakawa J, Ito J, Kashiwazaki N. Utility of progesterone receptor-ires-Cre to generate conditional knockout mice for uterine study. Animal Science Journal 2021;92:e13615. https://doi.org/10.1111/asj.13615.

[256] Hill TP, Später D, Taketo MM, Birchmeier W, Hartmann C. Canonical Wnt/β-Catenin Signaling Prevents Osteoblasts from Differentiating into Chondrocytes. Developmental Cell 2005;8:727–38. https://doi.org/10.1016/j.devcel.2005.02.013.

[257] Rodda SJ, McMahon AP. Distinct roles for Hedgehog and canonical Wnt signaling in specification,differentiation and maintenance of osteoblast progenitors. Development 2006;133:3231–44. https://doi.org/10.1242/dev.02480.

[258] Trissal MC, Wong TN, Yao J-C, Ramaswamy R, Kuo I, Baty J, et al. *MIR142* Loss-of-Function Mutations Derepress ASH1L to Increase *HOXA* Gene Expression and Promote Leukemogenesis. Cancer Research 2018;78:3510–21. https://doi.org/10.1158/0008-5472.CAN-17-3592.

[259] Zhu L, Li Q, Wong SHK, Huang M, Klein BJ, Shen J, et al. ASH1L Links Histone H3 Lysine 36 Dimethylation to MLL Leukemia. Cancer Discovery 2016;6:770–83. https://doi.org/10.1158/2159-8290.CD-16-0058.