Epigenomic and transcriptomic analysis of developing, adult and aging brain: mechanisms of brain folding, neuronal function and finding novel therapy for dementia

CUMULATIVE DOCTORAL THESIS

for the award of the degree

"Doctor rerum naturalium" (Dr.rer.nat.)

of the Georg-August-Universität Göttingen

within the doctoral program,

International Max Planck Research School (IMPRS) of

Neuroscience

of the Georg-August University School of Science (GAUSS)

submitted by

M. Sadman Sakib

Born in Dhaka, Bangladesh

Göttingen 2020

Thesis Committee

1. Prof. Dr. André Fischer, Dept. for Psychiatry and Psychotherapy, German Center for Neurodegenerative Diseases (DZNE), University Medical Center Göttingen.
2. Prof. Dr. Tiago Fleming Outeiro, Experimental Neurodegeneration, University Medical Center, Göttingen.
3. Dr. Camin Dean, Trans-synaptic Signaling, European Neuroscience Institute, Göttingen (Currently group leader, Synaptic Dysfunction, at DZNE, Berlin)

Extended Thesis Committee

- Prof. Dr. Alexander Flügel, Institute for Neuroimmunology and Multiple Sclerosis Research, University Medical Center Göttingen
- Dr. Dr. Oliver Schlüter, Psychiatry and Psychotherapy, University Medical Center Göttingen
- Prof. Dr. Argyris Papantonis, Institute of Pathology, University Medical Center Göttingen

Date of oral examination:
30 November 2020

Declaration

Herewith I declare, that I prepared this doctoral thesis "Cell-type specific gene-expression and epigenome analysis in the developing, adult and aging brain" on my own and with no other sources and aids than quoted.

Göttingen, 23.10.2020   M. Sadman Sakib

_____

I would like to dedicate this work to my lovely wife Ferdous Sultana Samantha, my beautiful parents (Md. Nuruzzaman, Farhat Negar) and my little sister (Fariha Jahin Retee).

"If I have seen further than others, it is by standing upon the shoulders of giants."

-Isaac Newton (1675)

Table of content

General Introduction

1.General Introduction

I.   Chromatin plasticity and gene expression

• Epigenetics and histone modifications

Eukaryotic cellular DNA being tightly packed with histone proteins in the nucleus is termed as chromatin. While gene expression was thought to be controlled only by the DNA sequences, decades of research have established epigenetic control being equal or even more important in transcriptional control of a typical cellular event, more importantly, as a sensor to communicate between gene and environment (Marsit 2015). We now know that the DNA of a given eukaryotic cell is tightly packed in a chromatin environment, of which, nucleosome is the fundamental subunit consisting of 147 base pairs of DNA, wrapped twice around octameric histone proteins.
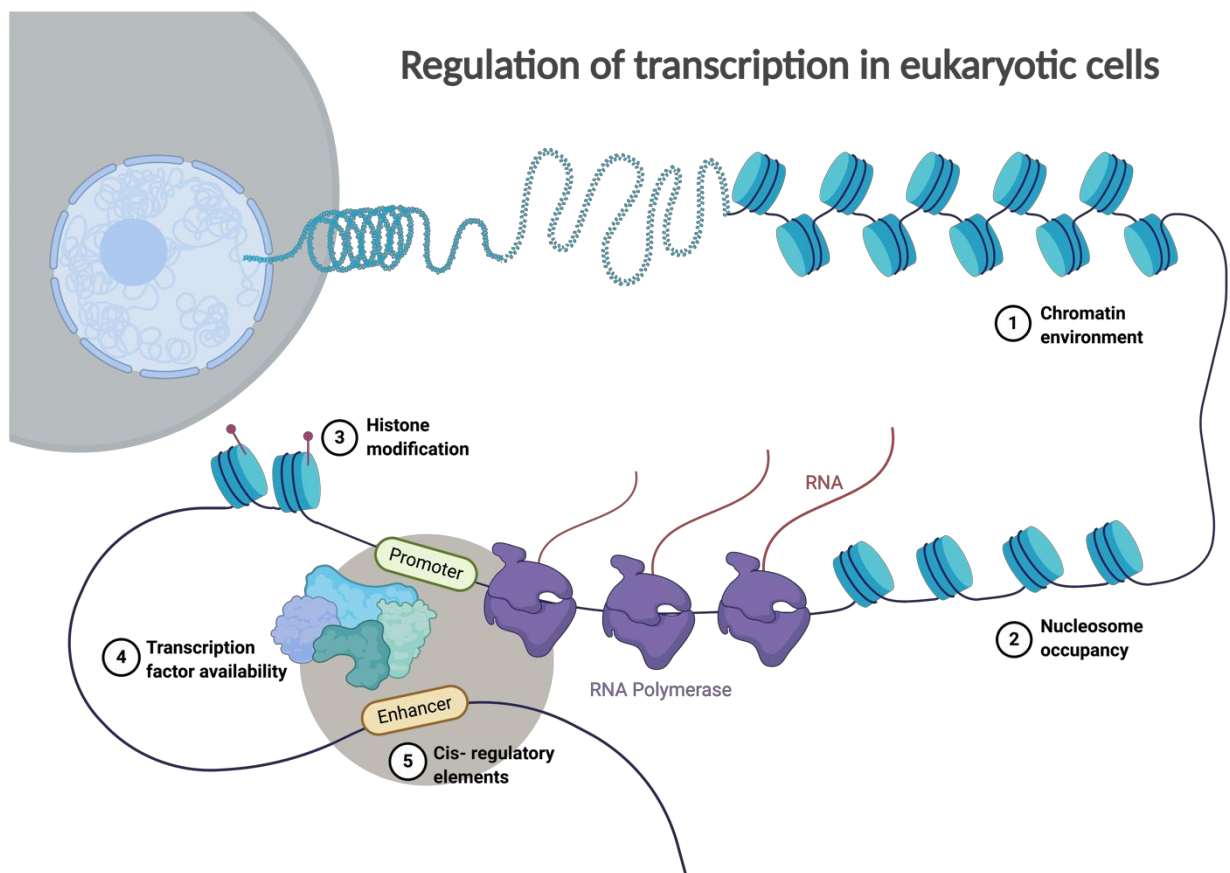
Diagram 1: Depiction of gene-chromatin environment inside a eukaryotic cell nucleus. Numbering refers to the layers of epigenetic and transcriptional controls for gene expression.

Epigenetic regulation comes in different modalities. The term "Epigenetics" originally implies changes in phenotypes without the change of the genotype (Waddington 1942). Among the most studied epigenetic mark is the cytosine-5 methylation in the DNA itself, enriched at CpG (Cytosine-Guanine rich) islands, that inversely correlates with gene expression(Jones 2012, Jones, Goodman et al. 2015, Kim and Costello 2017). Other key epigenetic regulations discovered till today are histone modifications, histone variants, microRNA, circular RNA, long noncoding RNA and RNA methylation (Fischer 2014, Chen, Li et al. 2017, Wei, Huang et al. 2017, Meng, Zhou et al. 2019, Mongan, Emes et al. 2019). However, discussion for all the other types of epigenetic mechanisms is beyond the scope of this thesis. Therefore, I will focus only on describing histone modification, which is part of this thesis work and probably one of the most widely studied epigenetic mechanisms till date.
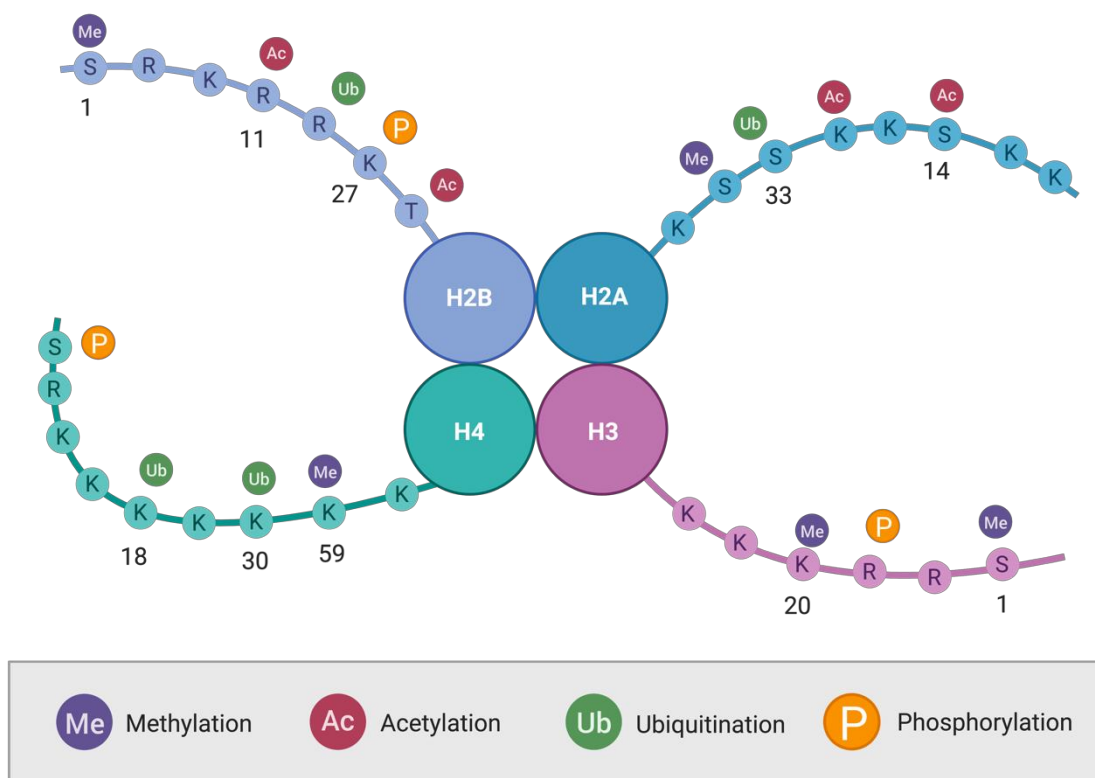
Diagram 2: Diverse histone modifications at its N-terminal tail. Major modifications are methylation, acetylation, ubiquitination and phosphorylation.

Histone proteins are divided into 5 major classes. H1 (linker histone), H2A, H2B, H3 and H4 (core histones) are some of the most long lived proteins in a cellular system (Bhasin, Reinherz et al. 2006, Toyama, Savas et al. 2013). They are positively charged due to their

abundance in arginine and lysine contents, by which they form dimers of each core histones and produces octameric nucleosome to form chromatin with the negatively charged DNA.

• Histone modifications mediate gene expression control

Nucleosome core histones can bear hundreds of different chemical modifications including acetylation, methylation, phosphorylation, ubiquitination etc. Earliest evidence of histone modification was shown from the study by Allfrey et. Al, where they showed for the first time that histones can be acetylated and methylated using radio isotope labelling and proposed its role in gene expression regulation around 55 years ago(Allfrey, Faulkner et al. 1964). Apart from these modifications, there are several new modifications discovered till today where the authors found that histones can be tagged with dopamine or serotonin, two key neurotransmitters of the brain. Hence they are called histone dopaminylation and serotonylation, which have been shown to be involved in transcriptional plasticity (Farrelly, Thompson et al. 2019, Lepack, Werner et al. 2020). These new found modifications suggest that there still might be undiscovered histone modifications to be found in future.
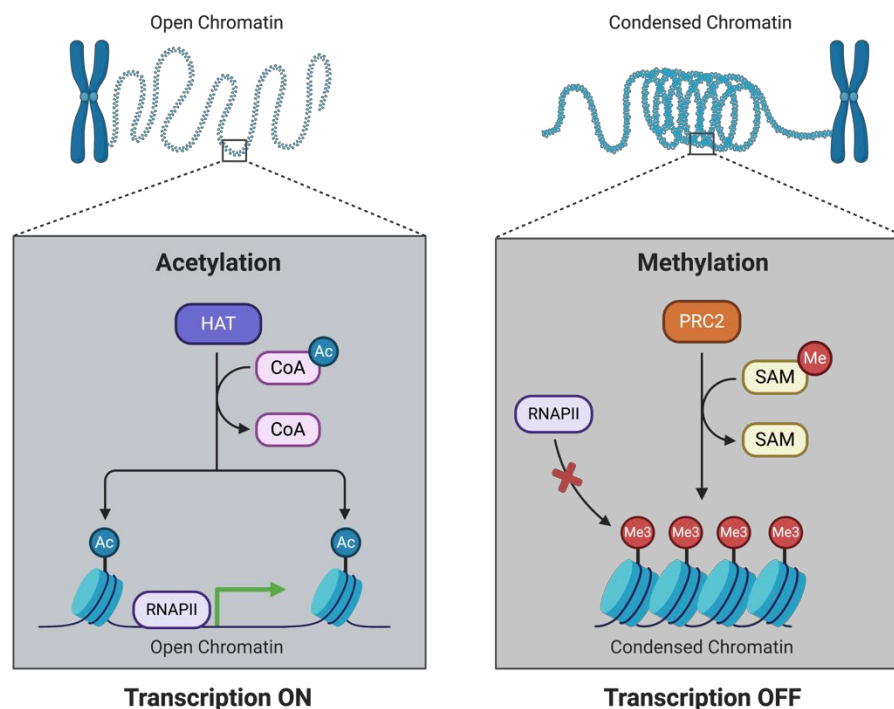
Diagram 3: Example of enzymatic histone modifications resulting in open or closed chromatin, thereby inducing or inhibiting gene transcription. For acetylation, Acetyl-CoA

3

and for methylation, S-adenosyl methionine (SAM) moiety is being used to chemically modify histones.

While histones can be chemically modified by its "writer", it is not permanent, and can be reversed by specific enzymes termed as "erasers". Based on their dynamic interplay, histone modifications form the "Histone code" that regulates gene transcription. The two most common marks, histone acetylation and methylation are chemically added or removed via vast numbers of histone acetyltransferases (HAT) and histone deacetylases (HDAC) or histone methyltransferases (HMT) and histone demethylases (HDM), respectively (Hyun, Jeon et al. 2017, Xu, Zhang et al. 2017).

Histone modifications effect chromatin conformation and/or gene expression mainly via two different processes(Bannister and Kouzarides 2011);

    (i)   Influencing chromatin structure by opening with an activatory or closing them with an inhibitory histone mark (*Diagram 3*), and

    (ii)   They act as molecular beacons to attract effector molecule binding, be it transcriptional activator or repressor.

Although many of these mechanisms are currently being heavily studied, we are still on the way of understanding and deciphering more. However, it is evident that histone modifications are strong predictors of gene expression (Karlic, Chung et al. 2010).

• Bivalent chromatin region: pathway to neurodegeneration?

"Bivalent domain" or "Bivalent region" were first coined in 2006 when the authors found activatory H3K4me3 and inhibitory H3K27me3 marks at essentially the same genomic regions while studying mouse embryonic stem cells (Bernstein, Mikkelsen et al. 2006). They were found at silent promoter regions which were supposed to code for developmentally expressed transcription factors. But they were not expressed since they were in "poised" state due to having bivalent promoters, but were later expressed during cell differentiation. Interestingly, genes were activated by losing the inhibitory H3K27me3 or inhibited by losing the activatory H3K4me3 mark from the bivalent regions. These histones of bivalent chromatin regions are modified by two groups of proteins, Trithorax group (TrxG, responsible for H3K4me3) and Polycomb group (PcG, responsible for H3K27me3) proteins (Geisler and Paro 2015, Harikumar and Meshorer 2015, Blanco, Gonzalez-Ramirez et al. 2020).

4

Although bivalent regions are extremely important in cell developmental stages and differentiation, recent publication have implicated role of bivalent promoter regions to be crucial in adult neurons, where the authors have shown that bivalent promoters suppress developmental and non-neuronal gene expression in adult neurons and genetic knockout of PcG proteins led to de-repression and activation of those silent promoters, allowing neurons to have aberrant gene transcription and eventually neurodegeneration in mice cerebellum (von Schimmelmann, Feinberg et al. 2016). All these evidences suggest that further investigation is needed to untangle the role of bivalent chromatin regions in aspect of aging, dementia and hippocampal memory impairment.

II. Nuclei instead of cells as next generation approach for brain cell type specific investigations

Our brain is an ensemble of billions of neurons and non-neuronal cells, encompassed by uncountable number of synapses that is rapidly firing every second, which is the basis of sensory-motor control, cognition, learning, memory formation and retrieval of memory. It is of paramount importance to fully understand individual cell type specific gene expression and epigenome to fully comprehend the concurrent mechanisms and discover novel therapy options where the origin of any diseases is thought to be potentially starting from few deregulated, distinct cell types. While utilizing fluorescence-activated cell sorting (FACS) have been generally utilized to produce cell type specific data from the brain (Saxena, Wagatsuma et al. 2012, Abruzzi, Chen et al. 2015, Okaty, Freret et al. 2015), it is not a viable method as it is difficult to isolate whole cells from the brain and using protease based cell dissociation methods led to unwanted gene and protein expression changes during sample processing(Huang, Hsing et al. 2010).

The very first attempt of isolating and FACS sorting distinct cell type from the brain was performed in 1977 (Campbell, Schanchner et al. 1977), where the authors, for the first time, did density gradient centrifugation and subsequent FACS sorting for glia enriched cell populations from the mouse cerebellar cortex. The discovery of pan-neuronal nuclei

marker protein and monoclonal antibody for its detection(Mullen, Buck et al. 1992) led towards utilizing this antibody to FACS sort neuronal nuclei specifically and study DNA or chromatin modifications for the first time (Siegmund, Connor et al. 2007, Jiang, Matevossian et al. 2008). Throughout last decade, nuclei-based methods to study cell type specific genomics have been on exponential level, with its wide ranges of application in RNAseq, ChIPseq, ATACseq, single nuclei RNAseq and so on (Nishioka, Shimada et al. 2013, Dincer, Gavin et al. 2015, Krishnaswami, Grindberg et al. 2016, Lacar, Linker et al. 2016, Fullard, Giambartolomei et al. 2017, Habib, Avraham-Davidi et al. 2017, Reddy, O'Brien et al. 2017).
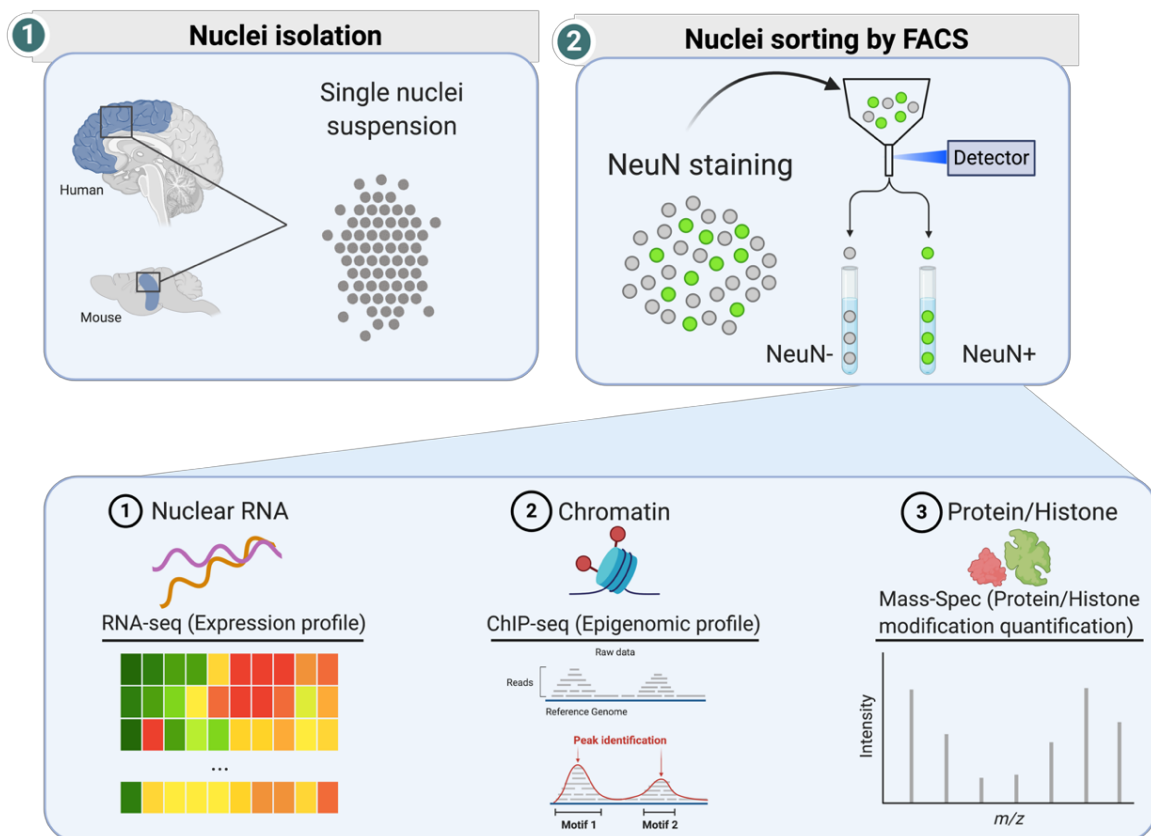
Diagram 4: Utility of brain nuclei isolation for multi-omics studies. Isolated nuclei can be subjected to cell type specific sorting and subsequent analysis of gene expression, epigenome or even proteome.

Apart from RNA and DNA based studies, nuclei sorting has also been used to profile protein mass-spec based studies where neuronal and non-neuronal nuclear protein half-lives were measured (Fornasiero, Mandad et al. 2018, I am an author in this study). Major utility of this method is unbiased isolation of intact cell nuclei that can be used to either FACS sort cell types of interest using antibody targeting nuclear protein markers, or use

this nuclei suspension to perform single nuclei RNA sequencing. This avoids harsh treatments applied to isolate whole cells that perturbs gene expression. Studies have compared nuclear and cellular gene expressions and concluded that nuclear transcriptome can substitute whole cell based transcriptome given the high correlation (Grindberg, Yee-Greenbaum et al. 2013). Moreover, nuclei-based methods enable to utilize frozen tissues from archived patient brain samples that can unlock plethora of new data for more precise measurements of genome information.

III.   Cortical folding: tale of the progenitors

Cerebral cortex regulates complex social behaviors in the mammalian brain. Mammalian species have great variation with cortical dimensions, even though body-size **:** brain-weight is not correlated with intelligence (Striedter 2005, Zilles, Palomero-Gallagher et al. 2013). In pathological terms, having both smaller or larger brains then usual (termed as micro- or macro- cephaly, respectively) can be prognosed with intellectual disabilities, autism and cognitive disorders (Woods, Bond et al. 2005, Olney 2007). Cortical development relies upon augmented neural progenitor and neural stem cell division during embryogenesis, and concomitant differentiation into mature neurons. This in turn leads to subsequent folding of the cortex. Evolutionarily, small mammals, primarily rodents and small primates have smooth-surfaced brain (lissencephalic), while most primate brains and ferrets have folded cortices (gyrencephalic)(Kelava, Lewitus et al. 2013).

(continued to next page….)
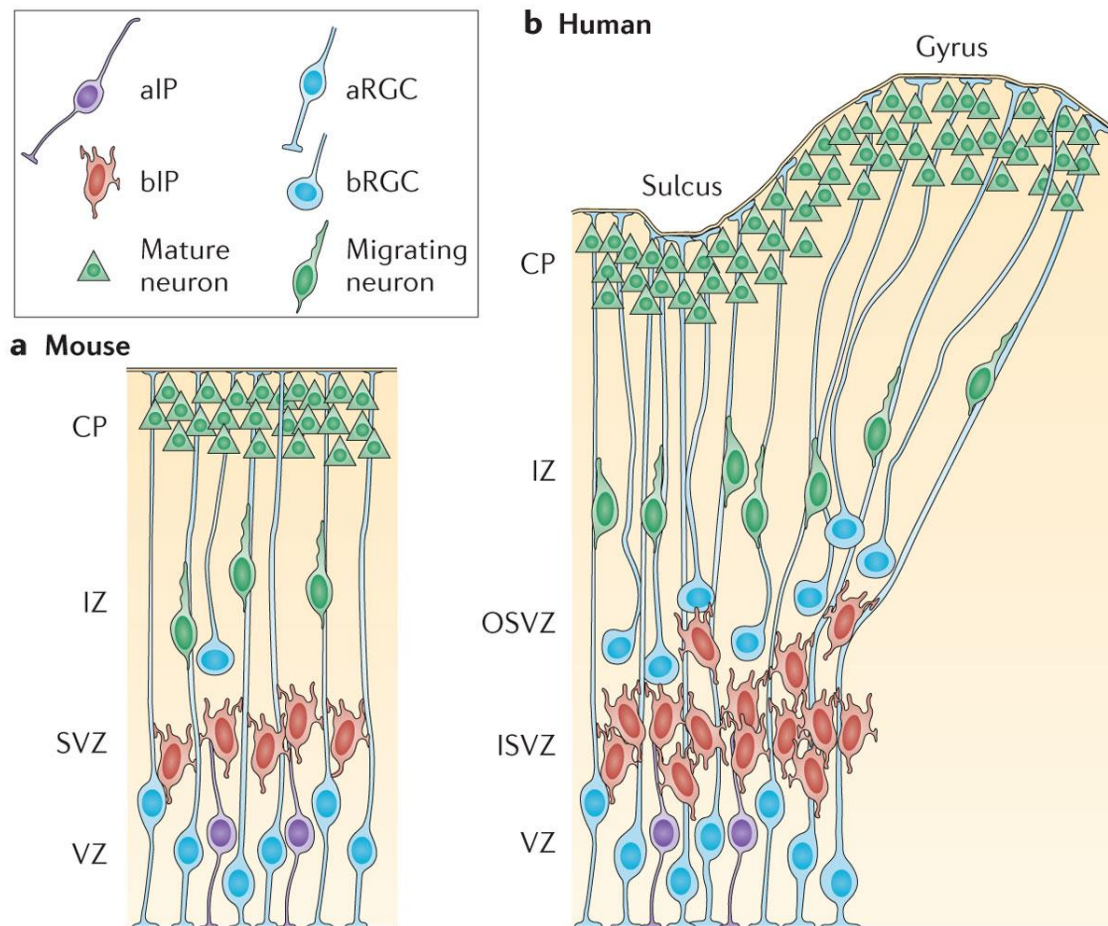
Diagram 5: Difference of basal intermediate progenitor (bIP) cell numbers could potentially give rise towards gyration of primate(human) brains. aIPs = apical intermediate progenitors, aRGCs = Apical radial glial cells, bIPs = Basal intermediate progenitors, bRGCs = Basal radial glial cells, CP = cortical plate, IZ = intermediate zone, SVZ = subventricular zone, ISVZ = inner SVZ, OSVZ = outer SVZ, VZ = ventricular zone. Adapted from Sun and Hevner 2014.

Previous studies have identified several key molecular underpinnings of cortical folding, ranging from role of cell cycle related genes (Estivill-Torrus, Pearson et al. 2002, Chen, Melendez et al. 2009, Glickstein, Monaghan et al. 2009, Mi, Carr et al. 2013), neuronal apoptosis (Haydar, Kuan et al. 1999), microRNAs (De Pietri Tonelli, Pulvers et al. 2008, Hong, Zhang et al. 2013, Saurat, Andersson et al. 2013) and asymmetric-symmetric division (Cappello, Attardo et al. 2006). All these mechanisms point towards core mechanistic role of neural progenitors, specifically basal intermediate progenitor cells (bIPs), controlling the gyrification across mammals, as well as in human abnormal cortical size related diseases (Kriegstein, Noctor et al. 2006, Lange, Huttner et al. 2009, Gruber, Zhou et al. 2011). Since the number of bIPs vary across lissencephalic and gyrencephalic mammals evolutionarily (Sun and Hevner 2014), the "Intermediate progenitor hypothesis" suggests that differential proliferation rate of bIPs contributes

towards cortical folding (Kriegstein, Noctor et al. 2006). Typically, bIPs expresses Eomesodermin/Tbr2(T-box brain protein 2) gene, a nuclear protein. Antibody against it can be used to visualize bIPs in brain sections via immunohistochemistry (Mutch, Schulte et al. 2010, Narayanan, Pham et al. 2018). As no epigenetic mechanisms has been discovered before controlling this process of gyrification, it will be highly interesting to utilize nuclei isolation, staining and sorting of Tbr2 nuclei specifically to study epigenome and gene expression in context of cortical folding in different mammalian species (demonstrated in manuscript 1).

IV. Histone H3K4 methylation, learning and memory

As discussed in the first section, histone modifications are crucial epigenetic processes regulating gene expression, thereby protein synthesis. Among all the marks, Histone 3 Lysine 4 (H3K4me) methylation is a prominent and conserved epigenetic mark that denotes transcriptional activation, either being present at mono-methylation form (H3K4me31) at the enhancers or tri-methylation form (H3K4me3) at the promoter elements of a gene (Diagram 6). While memory formation in the brain, i.e. encoding, consolidation and recall happens due to trillion neuronal synapses, gene transcription and protein translation seems to be an absolute requirement due to classical neurobiological experiments linking neuronal long-term potentiation (LTP) and de novo protein synthesis (reviewed in Fischer 2014). In the last decade, functional studies of lysine methyl transferases and demethylases has strengthened the fact that H3K4 methylation is a fundamental neuroepigenetic factor for learning and memory formation related transcriptional programs (Gupta, Kim et al. 2010, Kerimoglu, Agis-Balboa et al. 2013, Aguilar-Valles, Vaissiere et al. 2014, Jakovcevski, Ruan et al. 2015, Shen, Jiang et al. 2016, Kerimoglu, Sakib et al. 2017, Scandaglia, Lopez-Atalaya et al. 2017, Webb, Sanchez et al. 2017). Mutations in writers or erasers of H3K4 methylation have been implicated in severe cognitive disorders such as autism, schizophrenia, intellectual disability and other rare diseases affecting the brain (Jensen, Amende et al. 2005, Ng, Bigham et al. 2010, Jones, Dafou et al. 2012, Dong, Walker et al. 2014, Takata, Xu et al. 2014, Tunovic, Barkovich et al. 2014, Zech, Boesch et al. 2016).
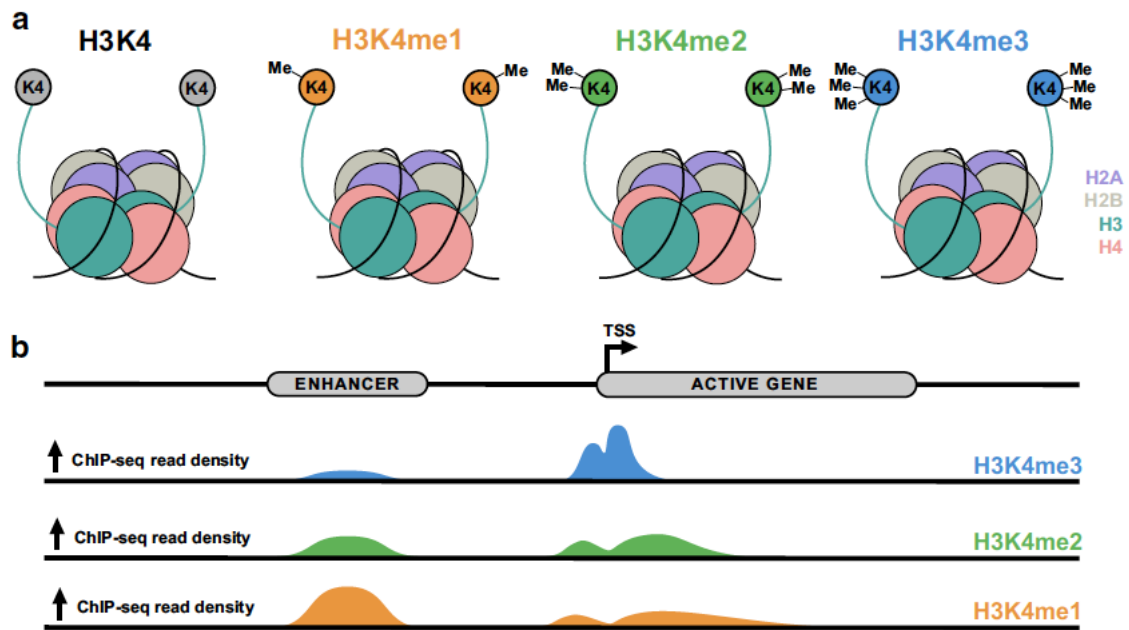
Diagram 6: Stepwise H3K4 methylation and enrichment patters in genomic elements such as promoters or enhancers. While H3K4me1 is mostly present at enhancers, H3K4me3 is exclusively present at transcription start sites (TSS). Adapted from (Collins, Greer et al. 2019)

Writers of H3K4 methylation have been functionally conserved from yeast to drosophila till mammals (Shilatifard 2012). For any given histone methyltransferase, *SET* domain (Su(var)3-9, Enhancer-of-zester and Trithorax) performs the catalysis of modifying H3K4 with methyl groups, sourced from S-adenosyl methionine (SAM) moiety (Shilatifard 2012). There are currently six histone methyltransferases and another six histone demethylases identified and functionally characterized in mammals (Allis, Berger et al. 2007). Our lab is particularly interested into histone methylation machineries and we have previously shown two histone methyltransferases, namely *Kmt2a* and *Kmt2b* to be involved in learning and memory formation (Kerimoglu, Agis-Balboa et al. 2013, Kerimoglu, Sakib et al. 2017). While different histone methylation machineries have been implicated into completely different human diseases, it is of utmost importance to functionally characterize all other HMTs and HDMs in aspect of learning and memory formation to gain deeper insight of the H3K4 methylation dynamic ranges with respect to neuronal functions. In this aspect, in this thesis, I have characterized one of the key histone methyltransferases implicated in intellectual disability and autism, Setd1b and its molecular imprints due to loss of function specifically from the forebrain excitatory neurons (demonstrated in manuscript 2).
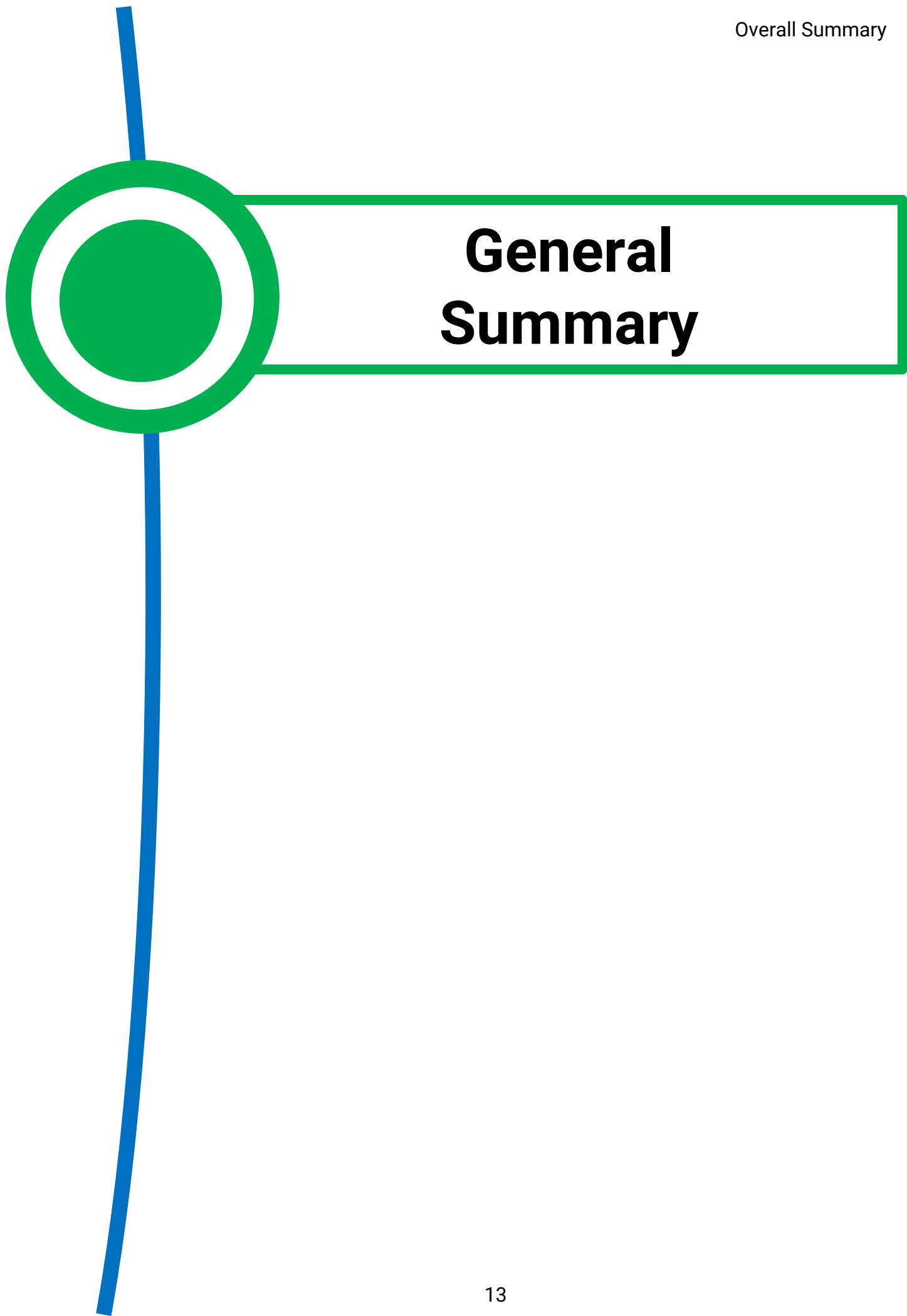
V. Aging related dementia and epigenetic therapy

Dementia and cognitive decline are parts of normal aging, manifested by progressive worsening of declarative memory, namely the ability to retrieve certain experiences and facts(Albert and Moss 1988). Age related cognitive decline starts during middle of the lifespan and since not everybody ages equally, human age itself is not sufficient to predict the start of this incident. However, human cognitive decline generally starts at around 50 years of age (Hedden and Gabrieli 2004).

Molecular mechanisms related to terminally aged brain has been linked to upregulated gene expression related to inflammation, calcium signaling, oxidative stress, mitochondrial dysfunction, dysregulated splicing and downregulated genes related to neuronal synapses (Prolla 2002, Blalock, Chen et al. 2003, Stilling, Benito et al. 2014, Benito, Urbanke et al. 2015, Ianov, Rani et al. 2016, Ianov, De Both et al. 2017). Very recently, large scale multi tissue bulk and single cell RNA sequencing study from around 20 organ types and 10 different age groups in mice revealed systemic manifestation of aging into overall physiology (Schaum, Lehallier et al. 2020, Tabula Muris 2020). These two studies are by far the most comprehensive transcriptome based longitudinal approach to find systemic gene expression deregulation due to aging in mice. The authors of those studies have concluded that genome instability, senescence and immune dysregulation are the key mechanisms driving age related systemic failure. Recent cross species epigenetic studies for H3K4me3 and H3K27ac (histone mark for enhancer) from heart, liver, brain cerebellum and olfactory bulb have also been linked to age related genome wide histone modification changes and induction of inflammatory responses (Benayoun, Pollina et al. 2019). However they did not perform cell type specific, not even hippocampus specific epigenetic profiling, which is crucial to link epigenetic changes with cognitive function(Halder, Hennion et al. 2016).

While the onset of age-related cognitive decline showed little to no gene expression changes, our group has previously shown that the molecular basis of early cognitive decline can be connected to decreased histone acetylation mark in mice (Peleg, Sananbenesi et al. 2010). By using an epigenetic drug to increase histone acetylation via inhibition of histone deacetylases, this age induced cognitive decline can be restored in aged mice as well. This was the very first evidence that by altering one of the histone modifications, age related dementia can be rescued.

As discussed in the previous section (4. Histone H3K4 methylation, learning and memory), histone methylation is an important neuroepigenetic marker that has not been characterized intensively in a genome wide fashion in age related cognitive decline till date. Evidences from western blot of mass-spectrometry-based analysis detected changes in histone methylation in aged mouse brains (Wang, Tsai et al. 2010, Gong, Qian et al. 2015). Given the vast implications of histone methylation changes in diverse cellular and tissue aging (Sen, Shah et al. 2016), urgent investigation of hippocampal cell type specific histone methylation changes and age-related cognitive decline is needed (demonstrated in manuscript 3). Afterwards, appropriate epigenetic drugs can be utilized to rescue age induced cognitive decline by reversing the aging brain histone methylome to a healthy state.

General Summary

2. General Summary

Histone modifications and gene expression are tightly regulated processes in the brain that has been shown to play crucial role from the beginning of brain development, learning-memory formation and aging. While brain comprises of numerous types of neurons and non-neuronal cells, this regulation is highly cell type specific. To gain more mechanistic insights on cell type specific epigenetic and transcriptomic processes, in this thesis, I demonstrated brain nuclei isolation, cell nuclei specific antibody staining and FACS sorting can be successfully utilized to perform cell type specific genome wide histone mark characterization, gene expression and single nuclei RNA sequencing. I have applied these tools to gain valuable mechanistic insights of the causal epigenetic mechanism for cortical folding, functional role of a histone methyltransferase in memory impairment, and multi omics-based characterization of aged induced cognitive decline model.

In the first manuscript, we found that embryonic mice treated with histone deacetylase inhibitors (therefore, increasing histone acetylation) led to higher amounts of basal progenitor (BP) cells in their cortex. This resulted into higher number of mature neurons, thereby producing cortical gyration phenotypes in lissencephalic rodent brains. To understand causal mechanisms, I established and performed for the first time, BP nuclei specific gene expression and histone 3 lysine 9(H3K9) acetylation dataset from embryonic mice cortex. This cell type specific analysis led to discovering distinct increased H3K9ac induced gene expression signature, that contained key regulatory transcription factor, resulting into higher amount of BP proliferation. Further validation experiments via epigenome editing confirmed the epigenetic basis of cortical gyrification in a lissencephalic brain via increasing histone acetylation.

For the second manuscript, I investigated the molecular role of a histone methyltransferase (HMT), Setd1b in mature neurons. Forebrain excitatory neuron specific Setd1B conditional knockout (cKO) resulted into severe memory impairment which required further characterization of neuron specific epigenetic and transcriptomic perturbation due to this cKO. To understand molecular function of Setd1b cKO in neurons, I isolated neuron specific nuclei from WT vs cKO mice hippocampal CA region and performed 4 different histone modification ChIPseq (H3K4me3, H3K4me1, H3K9ac, H3K27ac) and neuron specific nuclear RNA seq. Bioinformatic data analysis revealed promoter specific alteration of all 4 marks and significant down regulation of memory

14

forming genes. Comparison with other two previously studied HMT revealed Setd1b to be having broadest H3K4me3 peaks and regulating distinct sets of genes, which manifested to the severe most behavioral deficit. To understand expression pattern of those three HMTs, I performed single nuclei RNA sequencing of sorted neurons from wild type mice and found, even though Setd1b is expressed in a small subset of neurons, those neurons had the highest level of neuronal function and memory forming gene expression, compared to other two HMT expressing neurons studied previously by our group. Overall, our work shows neuron specific role of Setd1b and its contribution towards hippocampal memory formation.

In the third manuscript, I applied neuronal and non-neuronal epigenome and transcriptome data generation and analysis of 3 vs 16 months old mice. As it is well known that memory impairment starts during the middle of life, and previous gene expression studies in mice showed very little to no changes while having cognitive deficit, I utilized nuclei based cell sorting method to study two promoter epigenetic marks(H3K4me3, H3K27me3) and RNA expression (including coding and non-coding) in neuronal and non-neuronal cells separately. Due to the novelty of the data, I first characterized the basal activatory H3K4me3, inhibitory H3K27me3, bivalent regions and gene expression in neuronal and non-neuronal nuclei. These epigenomic and transcriptomic datasets would be a valuable resource to the community to compare cell type specific gene expression and epigenomes with their datasets. Moreover, profiling epigenetic marks in old hippocampal CA1 neurons and non-neurons revealed massive decrease of epigenetic marks mostly in the non-neurons, while neurons only had decreased inhibitory H3K27me3 mark. Mechanistically, these epigenome changes correspond to probable non-neuronal dysfunction and neuronal upregulation of aberrant developmental pathways. Surprisingly, nuclear RNAseq revealed significant number of genes deregulated in non-neuronal cells, compared to neurons. By integrating transcriptome and epigenome, I found decreased H3K4me3 leading to decreased gene expression in non-neuronal cells, that resulted into probably downregulated neuronal support function and downregulated important glial metabolic pathways related to extra cellular matrix.

Therefore, in this thesis, I have described cell type specific neurodevelopmental, neuronal and cognitive decline related epigenetic and transcriptional pathways that would add valuable knowledge and resources to the neuroscientific community.

Manuscript 1
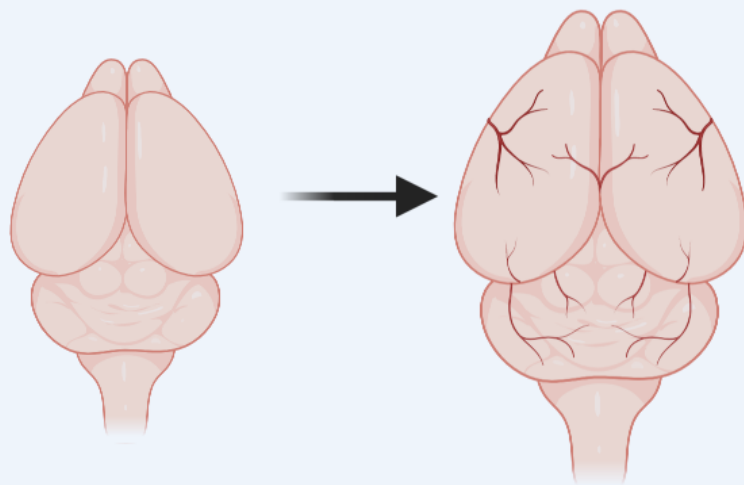
Mouse brain folding (like primates) with HDAC-inhibitor

3. Manuscript 1

H3 acetylation selectively promotes basal progenitor proliferation and neocortex expansion by activating TRNP1 expression

Cemil Kerimoglu[1,2,#], Linh Pham[1,4,#], Anton B. Tonchev[3,5,#], **M. Sadman Sakib**[1,2,#], Yuanbin Xie[1], Godwin Sokpor[1,4], Pauline Antonie Ulmke[1], Lalit Kaurani[1,2], Eman Abbas[1], Joachim Rosenbusch[1], Alexandra Michurina[1], Vincenzo Capece[1], Meglena Angelova[5], Miriam Esgleas[6], Mareike Albert[7], Radoslav Minkov[8], Emil Kovachev[8], Ulrike Teichmann[3], Rho H. Seong[9], Wieland Huttner[7], Magdalena Götz[6], Huu Phuc Nguyen[4], Anastassia Stoykova[3,10], Jochen F. Staiger[1,10], Andre Fischer[1,2,10,*], and Tran Tuoc[1,4,10,*]

My contribution:
Shared first author for major contribution to the experimental part.

- Establishing basal neuron progenitor cell nuclei (Tbr2 expressing) sorting from embryonic mouse brain samples.
- Sorted Tbr2 specific nuclear RNAseq and ChIPseq data generation.
- Contribution to validating inactivated Cas9 mediated epigenome editing in cell culture system.

Manuscript publication status:

Ongoing

Submitted to
"Science Advances":
17 May 2020

Reviewers'
comments received:
22 August 2020

Responses to
Reviewers being
sent

H3 acetylation selectively promotes basal progenitor proliferation and neocortex expansion by activating TRNP1 expression

Cemil Kerimoglu[1,2,#], Linh Pham[1,4,#], Anton B. Tonchev[3,5,#], **M. Sadman Sakib**[1,2,#], Yuanbin Xie[1], Godwin Sokpor[1,4], Pauline Antonie Ulmke[1], Lalit Kaurani[1,2], Eman Abbas[1], Joachim Rosenbusch[1], Alexandra Michurina[1], Vincenzo Capece[1], Meglena Angelova[5], Miriam Esgleas[6], Mareike Albert[7], Radoslav Minkov[8], Emil Kovachev[8], Ulrike Teichmann[3], Rho H. Seong[9], Wieland Huttner[7], Magdalena Götz[6], Huu Phuc Nguyen[4], Anastassia Stoykova[3,10], Jochen F. Staiger[1,10], Andre Fischer[1,2,10,*], and Tran Tuoc[1,4,10,*]

[1]University Medical Center, Georg-August-University Goettingen, 37075 Goettingen, Germany.

[2]German Center for Neurodegenerative Diseases, 37077 Goettingen, Germany.

[3]Max-Planck-Institute for Biophysical Chemistry, 37077 Goettingen, Germany.

[4]Ruhr University of Bochum, 44791 Bochum, Germany

[5]Department of Anatomy and Cell Biology, Medical University-Varna, Varna 9002, Bulgaria.

[6]Institute of Stem Cell Research, Helmholtz Center Munich, German Research Center for Environmental Health, Munich, Germany.

[7]Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany

[8]Specialized Hospital for Obstetrics and Gynecology "Prof. Dimitar Stamatov" –Varna, Medical University – Varna, Varna 9002, Bulgaria.

[9]Research Center for Functional Cellulomics, Seoul National University, Seoul, Korea

[10]DFG Center for Nanoscale Microscopy & Molecular Physiology of the Brain (CNMPB), 37075 Goettingen, Germany

[#]**Contributed equally**

[*] Corresponding author:
Andre Fischer (andre.fischer@dzne.de)
Tran Tuoc (Tran.Tuoc@ruhr-uni-bochum.de)

ABSTRACT

Evolutionarily, the expansion of the human neocortex accounts for many of the unique cognitive abilities of humans. This expansion appears to reflect the increased proliferative potential of basal progenitors (BPs) in mammalian evolution, which may have evolved through epigenetic alterations in BPs. However, whether or how the epigenome in BPs differs between humans and other species is unknown. Here, we report that histone H3 acetylation is a key epigenetic regulation in BP amplification and cortical expansion. Through epigenetic profiling of sorted BPs, we show that H3K9 acetylation is low in murine BPs and high in human BPs. Elevated H3K9ac preferentially increases BP proliferation, increasing the size and folding of the normally smooth mouse neocortex. Mechanistically, H3K9ac drives BP amplification by increasing expression of the evolutionarily regulated gene, TRNP1, in the developing cortex. Our findings demonstrate a previously unknown mechanism that controls cortical architecture.

One Sentence Summary: H3 lysine 9 acetylation promotes basal progenitor amplification, neocortex expansion and gyrification by activating TRNP1 expression in evolution.

INTRODUCTION

The neocortex of the mammalian brain is radially structured into six neuronal layers and multiple functional domains that form the structural basis for human sensorimotor processing and intellectual ability. During embryogenesis, most cortical neurons are born from the successive division of neural progenitors cells (NPCs) located in the forebrain germinal zones (i.e., the ventricular (VZ) and subventricular (SVZ) zones). The various types of NPCs can be distinguished by their cell morphology, polarity, ability to generate a given cell lineage, and the site at which they undergo mitosis (Lui, Hansen et al. 2011, Borrell and Gotz 2014, Taverna, Gotz et al. 2014, Dehay, Kennedy et al. 2015). The two main types of NPCs in the developing cortex are the apical progenitors (APs) and basal progenitors (BPs). APs include the apical/ventricular radial glia cells (a/vRGs), which divide at the surface of the apical VZ. BPs, which are derived from APs, include the basal (or outer) radial glia (bRGs) and the basal intermediate progenitors (bIPs); the latter lack apical contact and have defined mitotic activities in the inner and outer subventricular zones (iSVZ and oSVZ, respectively) (Lui, Hansen et al. 2011, Borrell and Gotz 2014, Taverna, Gotz et al. 2014, Dehay, Kennedy et al. 2015). The aRGs and bRGs are capable of asymmetric division to self-renew, and can directly or indirectly (via bIPs) produce neurons (Lui, Hansen et al. 2011, Borrell and Gotz 2014, Taverna, Gotz et al. 2014, Dehay, Kennedy et al. 2015).

In rodents, which have a lissencephalic cortex, most BPs are neurogenic bIPs. In gyrencephalic species, such as primates, BPs (including bIPs and bRGs) are capable of undergoing self-amplification through symmetric proliferative divisions before they terminally divide to generate neurons (Lui, Hansen et al. 2011). The intricate folding (gyrification) of the human neocortex is considered to be an evolutionary adaptation to the massive expansion of neuronal populations arising from the high proliferative competence of human NPCs, especially BPs (Lui, Hansen et al. 2011, Nonaka-Kinoshita, Reillo et al. 2013, Stahl, Walcher et al. 2013, Dehay, Kennedy et al. 2015).

Recent cell sorting- and single cell-based transcriptional profiling analyses have identified a number of factors that are important for BP proliferation, cortical expansion, and folding (Nonaka-Kinoshita, Reillo et al. 2013, Stahl, Walcher et al. 2013, Bae, Tietjen et al. 2014, Florio, Albert et al. 2015, Johnson, Wang et al. 2015, Pollen, Nowakowski et al. 2015, Ju, Hou et al. 2016, Wang, Hou et al. 2016, Del Toro, Ruff et al. 2017, Liu, Liu et al. 2017, Fiddes, Lodewijk et al. 2018, Suzuki, Gacquer et al. 2018). Epigenomic methods have been

recently established to unravel epigenetic landscapes at the single-cell level, but such strategies are limited by their low coverage of the genome and tend to cluster cells in a manner that is biased toward easily profiled genomic regions. Thus, epigenetic profiling of BPs is still challenging, and the epigenetic mechanisms that are thought to coordinate the expression/repression of gene sets during the evolutionary expansion of the neocortex remain unknown.

Here, we used cell sorting and a new mass spectrometry-based epigenetic profiling to identify H3K9 acetylation as a key epigenetic regulation in BP proliferation, cortical expansion, and cortical folding. We found that although the levels of H3K9ac are comparable in murine and human APs, species-specific differences exist in the histone H3 acetylation of BPs, which is low in mouse BPs and high in human BPs. Interestingly, elevation of H3K9ac in the developing mouse cortex led to BP-specific increases in the promoter H3K9 acetylation and expression of TRNP1, which is a well-known regulator of NPC proliferation and cortical expansion (Stahl, Walcher et al. 2013, Martinez-Martinez, De Juan Romero et al. 2016). The experimental enhancement of H3K9ac also dramatically augmented the TRNP1 expression-dependent proliferative capacity of BPs, leading to enlarged cortical size and formation of gyri in the developing mouse cortex. The use of an epigenome editing-based approach to increase H3K9ac specifically at the TRNP1 promoter resulted in increased TRNP1 expression and BP proliferation. Notably, the promoter H3K9 acetylation and expression of TRNP1 were both higher in human BPs compared to mouse BPs, underscoring the relevance of H3K9ac-associated TRNP1 expression changes in neocortical evolution. Our findings demonstrate a novel mechanism of cortical expansion during evolution and suggest that it may contribute to the formation of neocortical gyri in higher primate/human brain.

RESULTS

Assessing the epigenetic changes in BPs during cortical evolution

Fig. 1
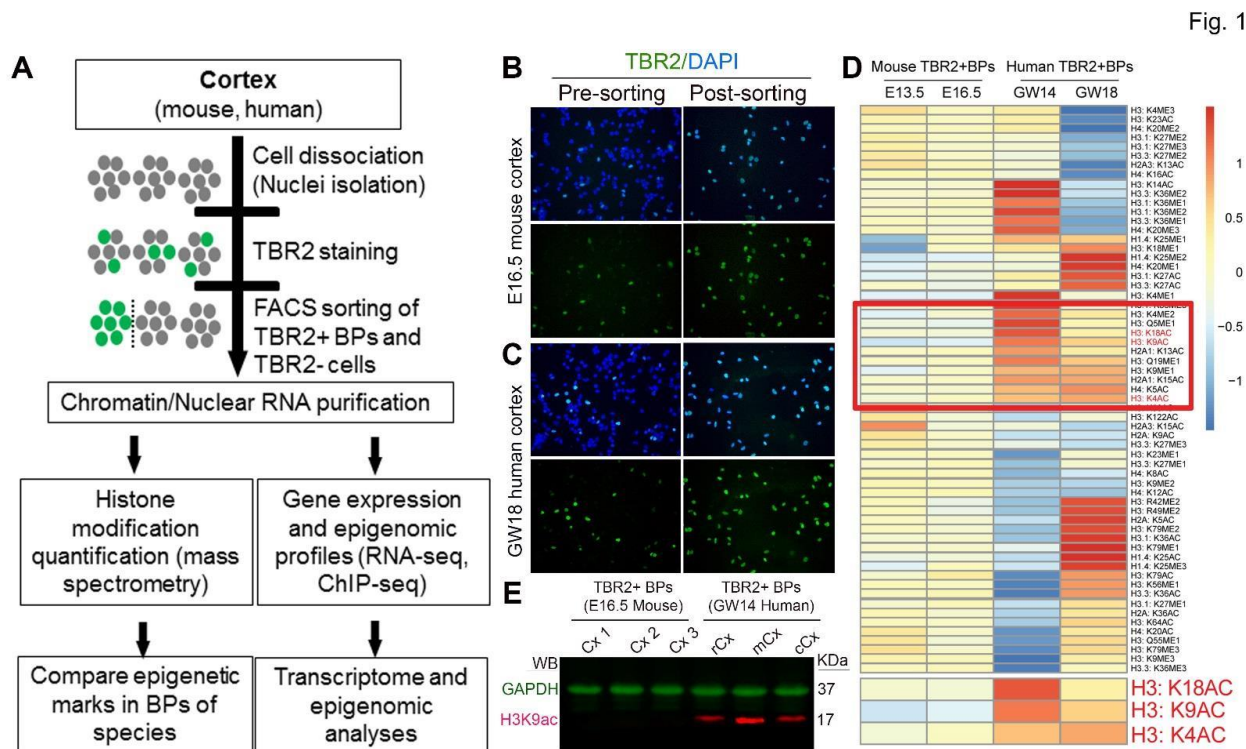
Figure 1: *Systematic screening for epigenetic marks showing differential levels between mouse and human basal progenitors (BPs)* (A) A scheme of the experimental design used to compare the transcriptomes and epigenomes of mouse and human BPs. (B, C) Purification of TBR2+ BPs in developing mouse and human cortex. Representative images of cell suspensions from mouse cortex (B) and human cortex (C) stained with DAPI and a TBR2 antibody. (D) The data for epigenetic marks is presented as a heat map. Bottom: enlarged pattern showing that the levels of H3K18ac, H3K9ac, and H3K4ac are low in mouse BPs and high in human BPs. (E) Western blot analysis of protein extracts from TBR2+ BPs of whole mouse cortex (Cx) and from human cortex at rostral (r), medial (m) and caudal (c) areas at indicated stages with H3K9ac antibody (in red) and with GAPDH antibody (in green, as loading control).

To test whether cortical expansion in evolution is correlated with alteration of the epigenetic landscape, we first investigated whether histone post-translational modifications (PTMs) differ between TBR2 positive (+) BPs from mouse and human cortices (Fig. 1A). To purify TBR2+ BPs from mouse and human developing cortices, we adapted a previously reported intracellular immunofluorescent staining and FACS protocol. We used an antibody to label intracellular TBR2 in single-cell suspensions isolated from E13.5 and E16.5 mouse cortices and gestational week (GW)14 and 18

human cortices, and then performed cell sorting of TBR2+ and TBR2 negative (-) cells (Fig. 1, fig. S1). Previous studies showed that TBR2 is expressed in bIPs (Englund, Fink et al. 2005) and in a subset of PAX6+, HOPX+ bRGs (Wang, Tsai et al. 2011, Vaid, Camp et al. 2018) in the lissencephalic rodent brain. In the gyrencephalic brains of ferret and primates, TBR2 labeling was seen for bIPs and almost half of the SOX2+, PAX6+ bRG population (Fietz, Kelava et al. 2010, Hansen, Lui et al. 2010). The expression of TBR2 however was not found in TNC+, PTPRZ1+ bRG subpopulation (Pollen, Nowakowski et al. 2015). Thus, the sorted TBR2+ cells from mouse and human developing cortices actually represent the majority of mouse BPs and human BPs, with the latter including human bIPs and a subset of human bRGs.

A newly-established mass spectrometry-based method (see Supplemental methods) was applied to quantify peptides containing methylated (me) or acetylated (ac) amino acid residues (Lysine, K; Glutamine, Q; Arginine, R) on the core histones (H2, H3, H4) and the linker histone H1 (Fig. 1D). Intriguingly, the levels of several epigenetic marks, predominantly modified H3 and its variants, appeared to be higher in human BPs at GW14 and GW18 than in mouse BPs at E13.5 and E16.5 (Fig. 1D, in selected frame). Some of the relevant H3 acetylation marks (Fig. 1D; red-labeled in the frame with a higher magnification shown at the bottom), including H3K9ac, H3K18ac, and H3K4ac, were previously shown to be enriched at promoters and enhancers, and to activate transcription. These findings suggest that epigenetic landscapes, particularly H3 acetylation, reflect extensive changes in BPs during cortical evolution, from rodents to humans.

Differential levels of acetylated histone H3 in basal progenitors of developing mouse and human cortex

From among the histone acetylation marks found to differ between murine and human BPs, we selected H3K9ac for further in-depth analysis because it exhibited the highest difference between BPs from these species (Fig. 1D) and has been implicated in neuronal function, development, and plasticity. To validate the screening result, we first performed immunohistochemical (IHC) analyses with antibodies against H3K9ac and progenitor subtype markers (as indicated in the legend of Fig. 2). In E15.5 and E16.5 mouse cortices, high levels of H3K9ac (H3K9ac^{high}) were associated with cells in the cortical plate (CP) and the majority of PAX6+ APs in the VZ, whereas relatively low levels of H3K9ac

(H3K9ac^low) were observed in most cells of the intermediate zone (IZ) and many cells of the VZ/SVZ (Fig. 2A/C). Remarkably, all TBR2+ BPs and many PAX6+ BPs in the SVZ and IZ were H3K9ac^low cells (Fig. 2A/C/D).
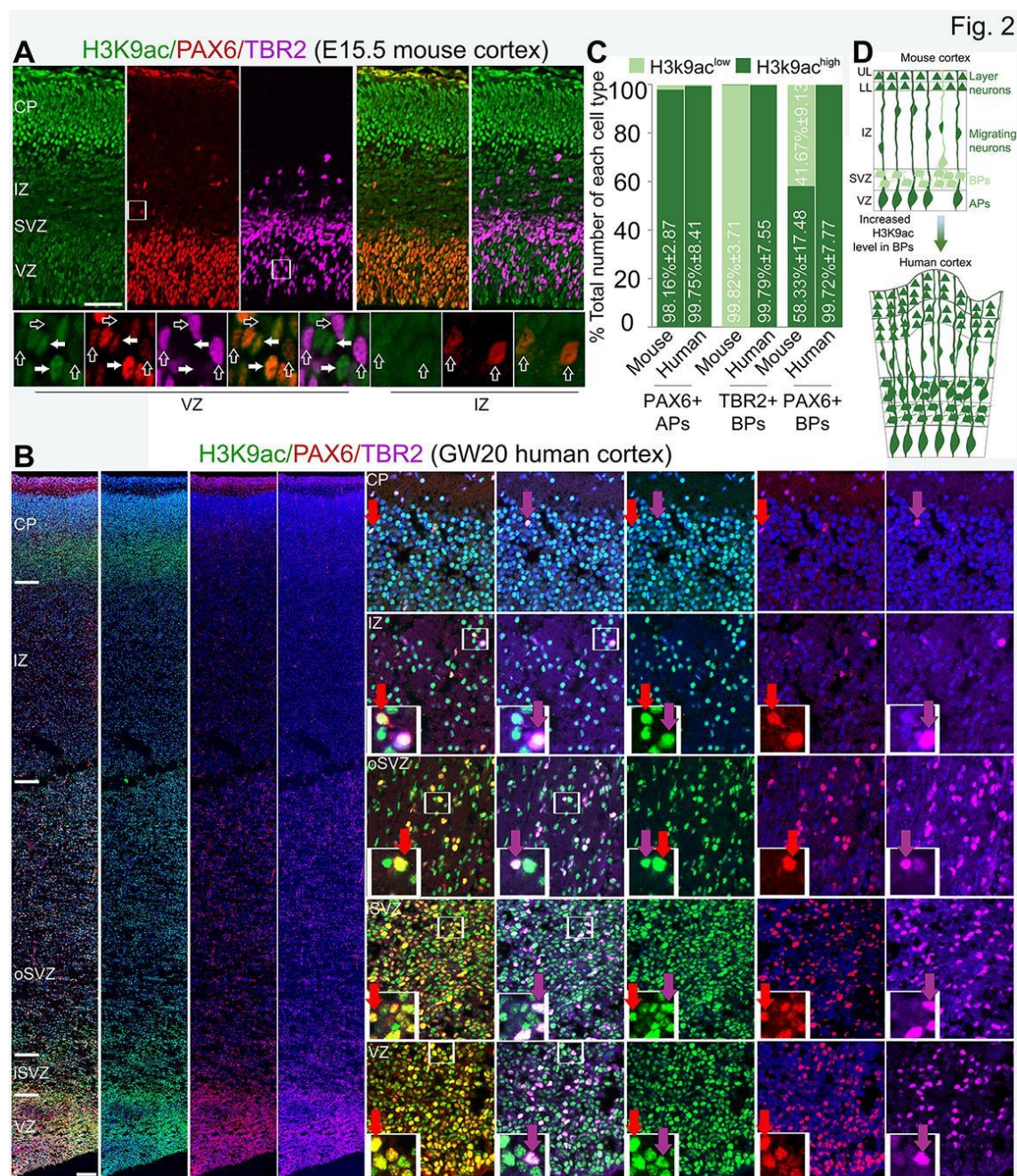
Figure 2: *Histone H3K9 is acetylated differently in BPs in the murine and human developing cortex.* (A) Images of cortical sections that were obtained from E15.5 mouse embryos and subjected to triple IHC of using antibodies against H3K9ac and PAX6 (to label APs in the VZ and BPs in the IZ) and TBR2 (to mark BPs in the SVZ). The lower panels show higher-magnification images of the boxed areas within the VZ and IZ; they reveal that PAX6^high+, TBR2^low+ APs are H3K9ac^high+ cells (filled arrows), whereas PAX6^high+ BPs in the IZ and TBR2^high+ BPs are H3K9ac^low+ cells (hollow arrows). (B) Images of cortical tissue from human embryos obtained at GW20 and subjected to triple immunolabeling with antibodies against H3K9ac, PAX6, and TBR2. Most of the PAX6+

APs and BPs (red arrows) and the TBR2+ BPs (magenta arrows) are highly immunoreactive for H3K9ac. (C) Statistical analyses of IHC results (shown in A and B) comparing the levels of H3k9ac (H3k9ac^{high} and H3k9ac^{low}) in progenitor subtypes in developing mouse and human cortex. The following were compared: PAX6+ APs in the VZ, TBR2+ BPs in the VZ/SVZ, and PAX6+ BPs in the IZ of the mouse cortex; and PAX6+ APs in the VZ, TBR2+ BPs in the VZ/iSVZ/oSVZ, and PAX6+ BPs in the iSVZ/oSVZ of the human cortex. (D) Schema illustrating the higher level of H3 acetylation found in human versus mouse BPs. Abbreviations: APs, apical progenitors; BPs, basal progenitors; VZ, ventricular zone; SVZ, subventricular zone; iSVZ, inner subventricular zone; oSVZ, outer subventricular zone; IZ, intermediate zone; CP, cortical plate. Scale bars = 50 μm.

Likewise, in the human cortex at GW20, a high level of H3K9ac was detected in the CP and in PAX6+ APs (Fig. 2B−D). In striking contrast to the mouse cortex, most PAX6+ BPs and TBR2+ BPs in the human iSVZ and oSVZ were H3K9ac^{high} cells (Fig. 2B−D). The difference in the H3K9ac levels of mouse and human TBR2+ BPs was also confirmed by Western blot analysis with lysates from FACS-collected TBR2+ cells (Fig. 1E).

To determine whether the detected difference in the H3 acetylation levels of the two species was restricted to lysine (K) 9 or also present on other amino acid residues, we examined the pan-acetylation of H3 (H3ac) (fig. S2). Triple IHC analysis of cortical tissues from E15.5 mouse and GW20 human embryos revealed that in the mouse cortex, H3ac is highly expressed in PAX6+ APs, but not in TBR2+ BPs. In the human cortex, however, PAX6+ APs in the VZ and PAX62+/TBR2+ BPs in the iSVZ and oSVZ showed high-level H3ac expression (fig. S2A-C). Notably, numerous H3ac- and H3K9ac- progenitors were found among the KI67+ cycling cells, suggesting that the elevations of H3ac and H3K9ac are not a common general trait of proliferating progenitors (fig. S2D/E). As histone H3 is acetylated more highly in human BPs than in mouse BPs, we designed experiments to elucidate the *in vivo* outcome of increased H3ac levels, using HDAC inhibition and overexpression of the H3K9 acetyltransferase, KAT2A/GCN5, during mouse cortical development.

Increased H3 acetylation promotes the generation of basal progenitors

We first tested whether increased acetylation of H3 could increase the genesis and proliferation of BPs. To elevate H3ac, we administered Trichostatin A (TSA) which is a selective class I/II histone deacetylase inhibitor (HDACi), to embryos from I (WT) mice (Fig. 3A). We previously showed that the cortex-specific loss of BAF155, as seen in BAF155cKO embryos, promoted delamination of APs, increasing the population of PAX6+, TBR2+ BPs in the IZ, and diminished the pool of PAX6+ APs in the VZ (Narayanan, Pham et al. 2018) (see also Fig. 3B/C, panels: Control + Veh and BAF155cKO + Veh). As PAX6+ BPs are relatively rare in the WT mouse cortex (Wang, Tsai et al. 2011, Vaid, Camp et al. 2018), we used the BAF155cKO mutant as a mouse model to investigate the effect of HDAC inhibition on the proliferation of BPs. TSA was injected beginning at 12.5 days post coitum (d.p.c.), and WT and BAF155cKO embryos were examined at E16.5–E18.5 (Fig. 3A). TSA treatment had no major effect on the pool of PAX6+/AP2γ+/SOX2+ APs in the VZs of WT and BAF155cKO mutants when compared to the corresponding vehicle (Veh)-treated controls (Fig. 3B/C). Interestingly, TSA injection increased the numbers of PAX6+/AP2γ+/SOX2+/TBR2+ BPs in the SVZ/IZ (Fig. 3B/C, left panels) of WT mice and more pronouncedly in BAF155cKO embryos (Fig. 3B/C, right panels).
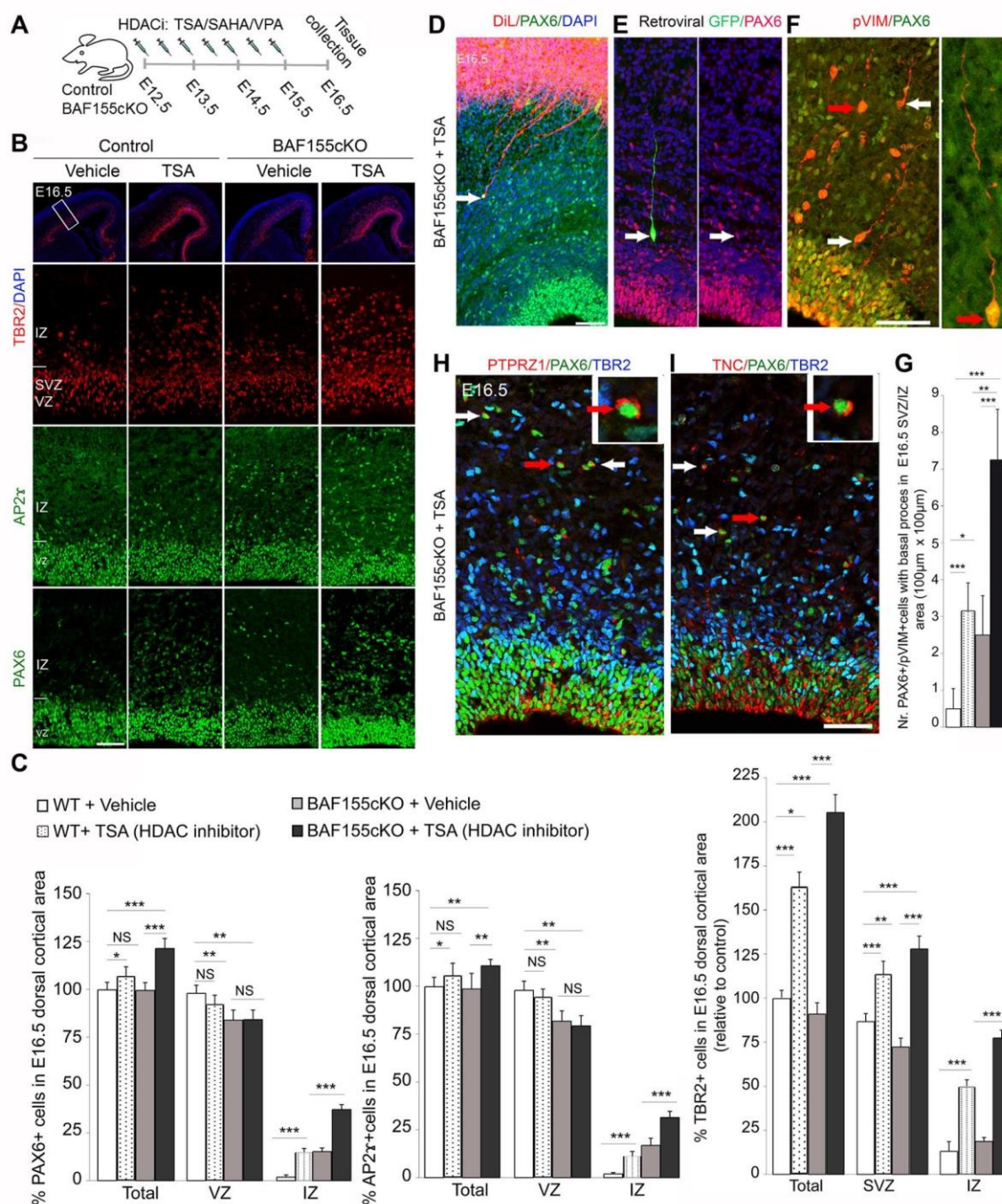
Figure 3: ***Enhanced acetylation of H3 increases the number of basal progenitors.***(A) Experimental paradigm in which BAF155cKO and control embryos were treated with HDAC inhibitors (HDACi: TSA, SAHA, VPA) at the indicated developmental stages. (B) IHC images showing TBR2 (BPs), PAX6 (APs, BPs), and AP2γ (APs, BPs) in sections from control and BAF155cKO embryos with or without TSA treatment. (C) Statistical comparisons indicate that increased H3 acetylation enhanced the number of BPs (TBR2+, PAX6+, AP2γ+ cells of the IZ) in TSA-treated WT and BAF155cKO embryos compared to vehicle (Veh)-treated controls. (D-F) Confirmation of BPs/bRGs (arrows) based on their morphology. BPs/bRGs exhibit long basal processes but no apical processes in TSA-treated cortex, as revealed by double labeling of Pax6 with either DiI (D), or retroviral GFP (E), or pVIM (F). Notably, fanned fibers of PAX6+ BPs/bRGs were observed when DiI

labeling was applied at the pia (D). (F) High-magnification image of a PAX6+/pVIM+ BPs/bRG, indicated by a red arrow. (G) Statistical analysis comparing number of PAX6+/pVIM+ BPs/bRGs with basal processes in the SVZ/IZ of the indicated embryos. (H, I) IHC analysis shows the expression of the human-enriched bRG markers, PTPRZ1 (H), and TNC (I), in PAX6+/TBR2- cells (arrows) of TSA-treated BAF155cKO cortex at E16.5 (see also fig. S8 for E18.5 cortex). Values are presented as means ± SEMs (*$p < 0.05$, ***$p < 0.01$, ***$p < 0.005$). Abbreviations: VZ, ventricular zone; SVZ, subventricular zone; IZ, intermediate zone. Scale bars = 50 μm.

Triple IHC analysis of PAX6, TBR2, and KI67 at E16.5 and E18.5 revealed that inhibition of HDAC leads to regionally restricted increases in TBR2+, PAX6+, KI67+ BPs in sections taken from the rostral, middle, and caudal dorsolateral cortex (d/lCx), but not the medial cortex (mCx) (fig. S3). HDAC inhibition exerted a dose-dependent effect, as more TBR2+/PAX6+ BPs were found in E18.5 cortex treated with TSA for 6 days (E12.5−E17.5) compared to those of embryos treated for only 3 days (E12.5−E15.5) (fig. S4A/B). Most TBR2+ BPs of the WT cortex were previously reported to be negative for PAX6 immunostaining (Englund, Fink et al. 2005). Interestingly, the proportion of cortical progenitors expressing both PAX6 and TBR2 was high in TSA-treated cortex (fig. S4C), as also reported in the developing gyrencephalic neocortex (Fietz, Kelava et al. 2010, Hansen, Lui et al. 2010).

Abundant PAX6+/DiI+ or PAX6+/retroviral GFP+ or PAX6+/pVIM+ cells with basal processes characteristic of BPs/bRGs were found in TSA-treated mouse cortex (Fig. 3D-G). Notably, TSA treatment resulted in increased numbers of PAX6+/pVIM+ BPs/bRGs, with the highest numbers found in TSA-treated BAF155cKO cortex (Fig. 3G). In addition, cell-surface proteins enriched in human bRG cells, such as TNC and PTPRZ1 (Pollen, Nowakowski et al. 2015), were also highly expressed in a subset of PAX6+/TBR2- a/bRG cells from TSA-treated mouse cortex (Fig. 3H/I, see also fig. S8).

To further confirm that the genesis of BPs is increased upon HDAC inhibition, pregnant mice were treated with agents whose properties are similar to those of TSA; these included valproic acid (VPA) and suberoylanilide hydroxamic acid (SAHA), which also inhibit class I/II HDACs. Both VPA- and SAHA-treated developing cortices had more BPs than vehicle-treated cortex (fig. S5; see also Fig. 6H and fig. S10E/F for cortical phenotype of the H3K9 acetyltransferase KAT2A overexpression), supporting the idea that the HDAC inhibition-induced elevation of H3 acetylation directly contributes to increasing the generation of BPs.

Increased H3 acetylation preferentially enhances the proliferation of basal progenitors but not apical progenitors

Fig. 4
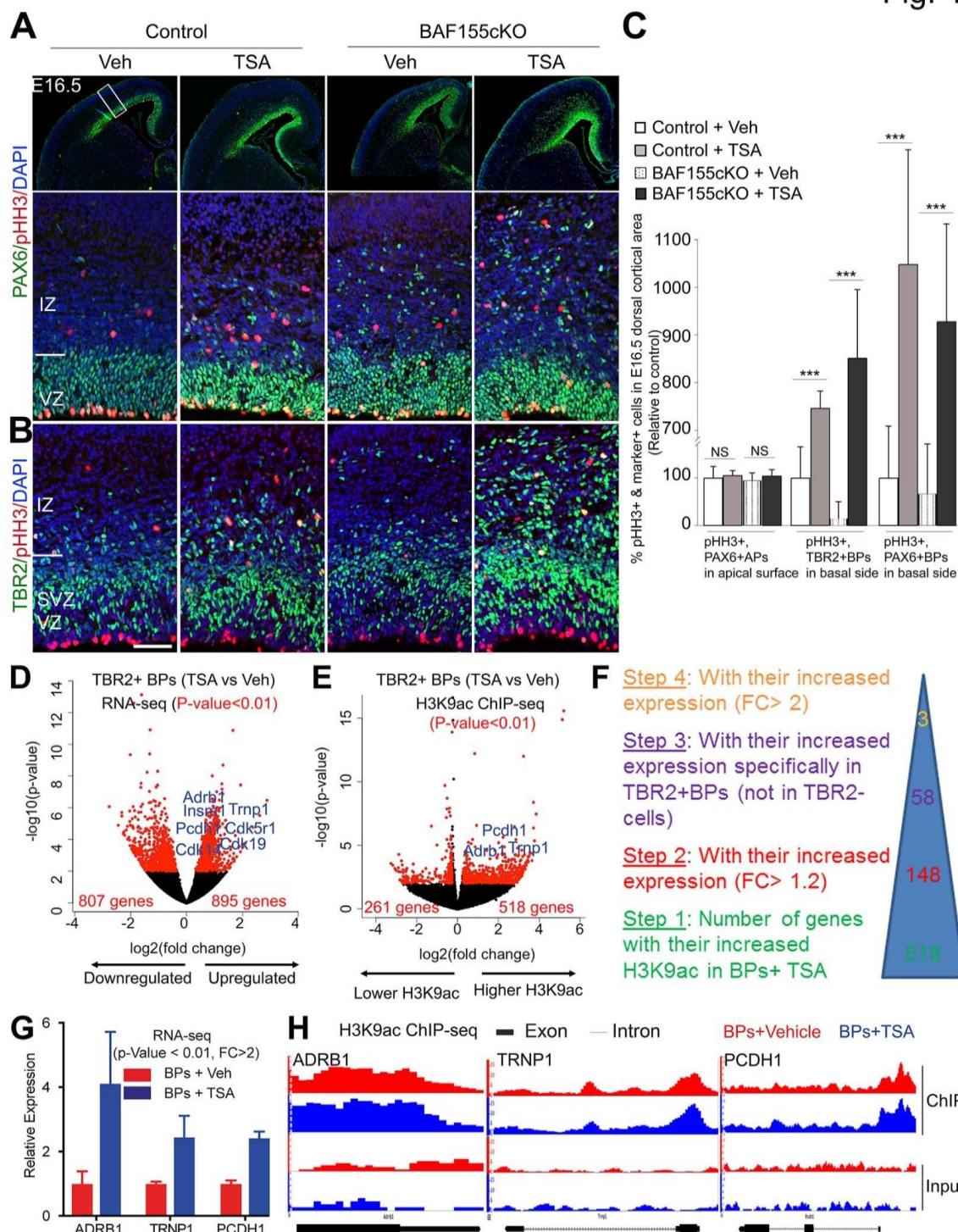
Figure 4: *H3 acetylation specifically promotes proliferation of basal progenitors and directly activates expression of ADRB1, TRNP1, and PCDH1 in TBR2+ basal progenitors.* (A, B) Images showing double IHC of PAX6/pHH3 and TBR2/pHH3 in control or BAF155-deficient cortex treated with Veh- or HDACi, as assessed at E16.5. Lower panels show higher magnifications of the areas indicated by the white box. (C) Statistical analyses reveal that elevated acetylation of H3 causes a significant increase in the number of TBR2+/pHH3+ BPs and PAX6+/pHH3+ BPs, whereas the number of apical

PAX6+/pHH3+ APs cells is not affected. (D, E) Volcano plots showing statistically significant changes (Paired Student's t-test < 0.01, FC > 1.2) visualized by our RNA-Seq (D) and H3K9ac ChIP-Seq (E) analyses of TBR2+ BPs in TSA vs. Vehicle experiments (see also fig. S7 for E16.5 BAF155cKO cortex and TBR2- cells). (F) Stepwise exclusion parameters were applied to the H3K9ac ChIP-Seq and RNA-Seq data sets to search for genes specifically activated by H3K9ac in TBR2+ BPs. (G) The expression levels of ADRB1, TRNP1 and PCDH1 are upregulated in TSA-treated BPs. (H) Distribution of H3K9ac along the gene bodies of ADRB1, TRNP1, and PCDH1 in TSA-treated BPs (red) and Vehicle-treated BPs (blue). Input (bottom row) and distributions after immunoprecipitation (upper two rows) are depicted. Values are presented as mean ± SEM (*p < 0.05, ***p < 0.01, ***p < 0.005). Abbreviations: VZ, ventricular zone; SVZ, subventricular zone; IZ, intermediate zone. Scale bars = 50 μm.

The presence of more BPs in TSA-treated BAF155cKO cortex (Fig. 3A-C) suggested that the delaminated progenitors undergo self-amplification in response to enhanced H3 acetylation. To ascertain if H3 acetylation is important for the proliferation of cortical progenitors (APs, BPs), we examined their proliferative capacity by performing IHC with antibodies against PAX6, TBR2, and phosphorylated histone H3 (pHH3) (Fig. 4A/B). Based on the expression of PAX6 in the VZ and pHH3 at the apical VZ surface (Fig. 4A/C), our data suggest that TSA injection did not influence the proliferation of APs, which already have a high endogenous level of acetylated H3. Strikingly, upon HDAC inhibitor treatment, both control and BAF155cKO cortices presented more non-apical proliferating (pHH3+ BPs) cells, along with higher ratios of proliferating BPs (TBR2+/pHH3+) to total TBR2+ BPs and proliferating BPs (basally located PAX6+/pHH3+) to total PAX6+ BPs (Fig. 4A–C). Consistent with these data, the numbers of actively cycling BPs (KI67+/TBR2+) and BPs (KI67+/PAX6+) in the SVZ/IZ were considerably increased in TSA-treated BAF155cKO cortex (fig. S6A/B). In addition to the slower cell cycle in neural progenitors from primate than that from rodent (Dehay, Kennedy et al. 2015), both proliferative APs and BPs exhibit a substantially longer S-phase than neurogenic progenitors (Arai, Pulvers et al. 2011). We therefore focused on examining the effect of TSA on progression between S-phase to G2+M- phases in APs and BPs. Given that, the average lengths at mid-gestation stage of the S- and G2+M- cell cycle phases are about 3.5 hours and 2 hours, respectively (Arai, Pulvers et al. 2011), we then examined whether the TSA treatment affect the progression within S-G2-M phases in APs and BPs. A thymidine analog injection paradigm (4 hours- IdU pulse-labeling) was used to mark cortical progenitors within these phases. Double immune-staining with antibodies

against IdU and pHH3 to label APs (apical surface- located IdU+/pHH3+) and BPs (basally located IdU+/pHH3+), which already entered into G+M- phases (Fig. S6C/D). Remarkably, less BPs (not APs) reached to late G2-M phases in TSA- treated cortices compared to Veh-treated cortices. This implicated that TSA treatment resulted in lengthening of cell cycle progression in S-G2-M phases in BPs, specifically.

Together, these data indicate that the proliferative capability of murine BPs is enhanced by elevated acetylation of H3.

Identification of H3K9ac target genes in TBR2+ BPs

Given that TSA treatment had a much stronger effect on BP genesis in BAF155cKO cortex, we firstly compared the gene expression and genome-wide H3K9ac profiles of TSA- and vehicle-treated BAF155cKO cortices by RNA-Seq and ChIP-Seq (fig. S7A/B; Table S1, S2). We found that TSA treatment yielded upregulation of 1961 genes and downregulation of 1799 genes (p-value < 0.01 & |fold change| > 1.2) (fig. S7A, Table S1). Examination of H3K9 acetylation on the TSA-regulated genes revealed that the upregulated genes had clear increases in the level of H3K9ac (fig. S7C), whereas the downregulated genes did not show any significant change in H3K9ac (fig. S7D). This result is in accordance with data showing that loss of H3K9 acetyltransferases can trigger activation of a set of genes via secondary effects.

In accordance with our IHC data showing an increase in the number of TBR2+/PAX6+ BPs, many neurogenic BP/bIP genes (e.g., NEUROG1, NEUROG2, EOMES/TBR2, NEUROD2, NEUROD6) were upregulated in the cortex of TSA-injected BAF155cKO mutants (fig. S7A; Table S1). TSA treatment also yielded increased expression of genes previously shown to be enriched in BPs/bRGs (e.g. TNC, PTPRZ1, PAQR8, GIGYF2, PDLIM3, ZC3HAV1; fig. S8A; Table S1) (Pollen, Nowakowski et al. 2015). The increased expression of many BP-enriched genes in response to HDAC inhibition in both BAF155cKO and WT cortex was also confirmed by qPCR and IHC analyses (fig. S7C, S8). Furthermore, the results from our ChIP-Seq and ChIP/qPCR analyses also confirmed that many of the BP genes (e.g., TNC, PAQR8, NEUROG2, EOMES, NEUROD1, NEUROD6; fig. S7D; Table S2) exhibited increases in their H3K9ac levels in the cortices of TSA-treated BAF155cKO and WT embryos. These findings indicated that H3 acetylation positively regulates the expression of BP genes.

Next, we evaluated the TSA treatment response of TBR2+ BPs (Fig. 4D/E; Table S3, S4) and TBR2- cells (fig. S7E/F; Table S5, S6). For this purpose, TBR2+ and TBR2- nuclei were sorted by FACS (Fig. 1; fig. S1). TSA treatment yielded expressional upregulation of 895 genes in TBR2+ BPs. To gain an evidence that H3 acetylation induces expression of BP genes during evolution, we compared the upregulated genes in TSA-treated TBR2+ BP genes (Fig. 5c) and BP/IP genes, which were recently identified specifically for macaque and human (Pollen, Bhaduri et al. 2019). Remarkably, the TSA treatment provoked expression of 40.1% (166 out of 414) macaque- specific BP genes and 13.3% (62 out of 467) human- specific BP genes in TBR2+ BPs in the developing mouse cortex (fig. 7G/H; Table S7), suggesting H3 acetylation has an evolutionary relevance in BP genesis during mammalian evolution. Among the upregulated genes in TSA-treated TBR2+ BP genes, 37 showed correlation between their upregulation and an increased level of H3K9ac upon TSA treatment (fig. S7I; Table S4; p-value < 0.01). Surprisingly, neither the H3K9ac level nor the expression levels of typical BP genes (e.g., EOMES, NEUROG1, NEUROG2, NEUROG1, NEUROD4) were found to be altered in TSA-treated BPs (fig. S7I; Table S3-S6). Thus, the expressional upregulation of most typical BP genes in TSA-treated TBR2+ cells seems to be a secondary effect.

The increased proliferation of BPs, which was visualized as increases in TBR2+/pHH3+ and TBR2+/KI67+ cells (Fig. 4A/B, fig. S6), suggested that the expression of proliferation-related genes would be increased. Indeed, we found that the expression levels of cell cycle genes (e.g., CDK19, CDK5R1, CDK14) and INSM1, which is known to promote BP proliferation, were upregulated in TSA-treated BPs (Fig. 4D; Table S3). Remarkably, however, the levels of H3K9ac at the promoters of these genes were not increased following TSA treatment (Fig. 4E; Table S4). These findings suggested that TSA treatment might influence the expression of factors that act upstream of the aforementioned proliferation-regulating genes.

To identify candidates for functional analysis, we undertook an unbiased approach. From among all genes that exhibited increased H3K9ac at their promoters in TBR2+ BP cells after TSA injection (|fold change| > 1.2; Fig. 4F), we selected those that also showed a trend for increased expression (|fold change| > 1.2; Fig. 4F). This screen netted 148 genes (Fig. 4F). From them, we filtered out those that exhibited increased expression in TBR2- cells, yielding 58 genes that had increased H3K9ac and were uniquely upregulated in TBR2+ BP cells following TSA treatment (Fig. 4F). These genes were further filtered

according to their basal expression levels (baseMean > 20) and selection of those that increased by more than 2-fold. This pipeline yielded three candidates: ADRB1, TRNP1, and PCDH1 (Fig. 4G/H, fig. S7K). This suggests that H3 acetylation promotes BP amplification by directly activating the expression of a small set of genes.

H3 acetylation controls the amplification of basal progenitors by regulating TRNP1 expression in the developing cortex
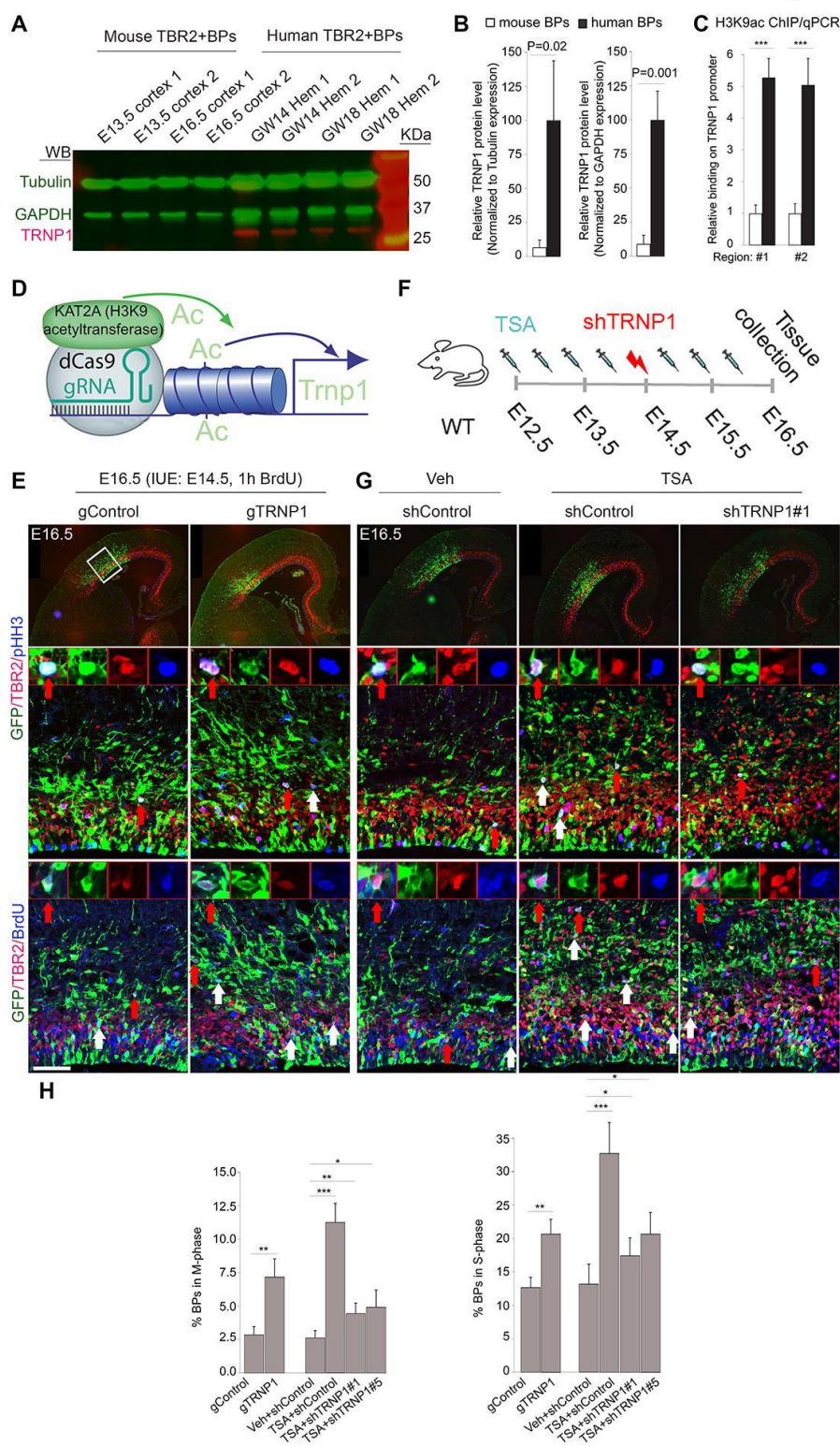
Fig. 5

Figure 5: *H3 acetylation controls the proliferation of BPs by activating TRNP1 expression.* (A, B) Western blot analysis of protein extracts from TBR2+ BPs obtained from mouse and human cortices at the indicated stages, as assessed with antibodies against TRNP1

(red), Tubulin, and GAPDH (green, as loading controls). (B) Relative levels of TRNP1 protein are presented in the diagram. Little to no TRNP1 protein is expressed in mouse BPs, whereas this expression is relatively high in human BPs. (C) ChIP-qPCR comparing the H3K9ac levels at the TRNP1 promoters in mouse and human TBR2+ BPs. (D) Schematic overview of the CRISPR/dCas9-based deposition of H3K9ac at the TRNP1 promoter, which was used to activate its expression. (E) Mouse E14.5 dorsolateral cortex was in utero electroporated with a gTRNP1-dCas9-KAT2A-T2A-eGFP plasmid (gTRNP1) or gControl-dCas9-KAT2A-T2A-eGFP plasmid (gControl), and IHC analysis of GFP, TBR2, pHH3, or BrdU was performed at E16.5. Images represent triple optical sections. White box indicates areas shown at higher magnification and red arrows point to examples of cells that are immunoreactive for GFP, TBR2, pHH3, or BrdU. (F) Descriptive scheme of the rescue experiment, in which mouse cortex was treated with the HDAC inhibitor, TSA, and gTRNP1 or shControl constructs at the indicated stages. (G) Triple IHC analysis for the markers listed in panel E. (H) Statistical quantification of the results shown in panels E and G is shown. Values are presented as mean ± SEM (*p < 0.05, ***p < 0.01, ***p < 0.005). Abbreviations: VZ, ventricular zone; SVZ, subventricular zone; IZ, intermediate zone. Scale bars = 50 μm.

In the developing mouse cortex, TRNP1 was exclusively localized in the nuclei of APs in the VZ and newborn neurons in the CP, with little to no TRNP1 expression seen among TBR2+ BPs in the SVZ (Fig. 5A) (Stahl, Walcher et al. 2013); it thus exhibits a pattern similar to that of H3K9ac (Fig. 1E, Fig. 2). In line with this, the expression and promoter H3K9ac levels of TRNP1 were lower in TBR2+ BPs than in TBR2- cells of the developing mouse cortex, as revealed by our RNA-Seq and ChIP-Seq experiments (fig. S9A/B). Notably the expression (Fig. 5A/B) and promoter H3K9ac (Fig. 5C, fig. S9C) levels of TRNP1 were strikingly higher in human TBR2+ BPs than in mouse TBR2+ BPs. These data suggest that the H3K9ac level directly controls the expression of TRNP1 (also see the model in Fig. 6I). This proposal was further corroborated by our observation that TRNP1 expression is increased specifically in TBR2+ cells upon TSA treatment (fig. S9D/E).

To more directly investigate whether a higher level of H3K9ac at the TRNP1 locus might enhance the gene expression and proliferation of BPs in the mouse developing neocortex, we sought to increase H3K9ac at the TRNP1 locus and examine the associated phenotype. To increase deposition of the H3K9ac mark, we adapted a CRISPR/Cas9-based system (Albert, Kalebic et al. 2017) that allows targeted editing of H3K9ac in the developing cortex (Fig. 5D). Given that the levels of H3K9ac at the TRNP1 promoter differed between mouse TBR2+ BPs and human TBR2+ BPs (Fig. 5C) and between mouse TBR2+ BPs and TBR2- cells (fig. S9B), and TSA treatment increased H3K9ac at the TRNP1

promoter specifically in BPs (fig. S9D/E), we selected the TRNP1 promoter for targeted H3K9ac editing. Several guide RNAs (gRNAs) targeting the TRNP1 promoter were designed and tested (fig. S9F/G). To edit histone acetylation at the TRNP1 locus, we generated plasmid constructs (gTRNP1-dCas9-KAT2A-T2A-eGFP) harboring DNA sequences that encoded a gRNA, dCas9 (a nuclease-deficient Cas9) fused with an H3K9ac writer KAT2A, and a fluorescent reporter (GFP) (Fig. 5D). For targeting of the TRNP1 locus, two constructs, each expressing one gRNA (gRNA#2 or gRNA#4), were used; these were designated gTRNP1 (fig. S9F/G).

To validate that H3K9ac levels were altered upon introduction of gTRNP1, we performed ChIP-qPCR in FACS-purified transfected (GFP+) Neuro2A cells. H3K9ac levels were found to be significantly increased at the TRNP1 promoter region upon gTRNP1 transfection, compared to control conditions (fig. S9H). In addition, qPCR analysis of the sorted cells revealed that TRNP1 expression was also increased following transfection with the generated editing constructs (fig. S9I). These results confirm that the CRISPR/Cas9-based epigenome-editing system established here achieved targeted deposition of H3K9ac at the TRNP1 promoter and increased the expression of TRNP1 in neural cells.

To achieve epigenome editing-based modulation of H3K9ac levels in the developing cortex, we delivered the constructs into APs by *in utero* electroporation (IUE) at E14.5. At E16.5, we analyzed the progeny of the electroporated APs, including TBR2+ BPs (Fig. 5D/E). By comparing the proportion of TBR2+/pHH3+ (BPs in M-phase of the cell cycle) and TBR2+/BrdU+ (BPs in S-phase) cells among the targeted TBR2+ BPs (revealed by GFP) in control- and editing construct-injected cortices, we examined whether the altered levels of H3K9ac might influence the proliferation of BPs (Fig. 5E/F). Similar to TSA treatment, epigenome editing-based augmentation of the promoter H3K9ac level of TRNP1 increased the percentage of basal mitosis (GFP+/TBR2+/pHH3+ or BrdU+) among targeted BPs (GFP+/TBR2+; Fig. 5E/H).

Because TRNP1 overexpression was shown to promote the proliferation of NSCs in the developing mouse and ferret cortex (Stahl, Walcher et al. 2013, Martinez-Martinez, De Juan Romero et al. 2016), we performed rescue expression to examine whether a specific increase in TRNP1 expression caused the aberrantly enhanced proliferation of BPs upon TSA treatment. To this end, we electroporated TSA-treated cortex with shTRNP1 constructs (Stahl, Walcher et al. 2013) to downregulate TRNP1 expression (Fig. 5F/G). Remarkably, the observed TSA treatment-induced proliferation phenotype of TBR2+ BPs

was largely rescued by TRNP1 knockdown in the cortex (Fig. 5G/H). These findings suggest that the elevated level of H3K9ac at the TRNP1 promoter is directly involved in facilitating the proliferation of BPs (also see the model in Fig. 6I).

Elevated H3ac leads to radial expansion and induces gyrification of the developing mouse cortex
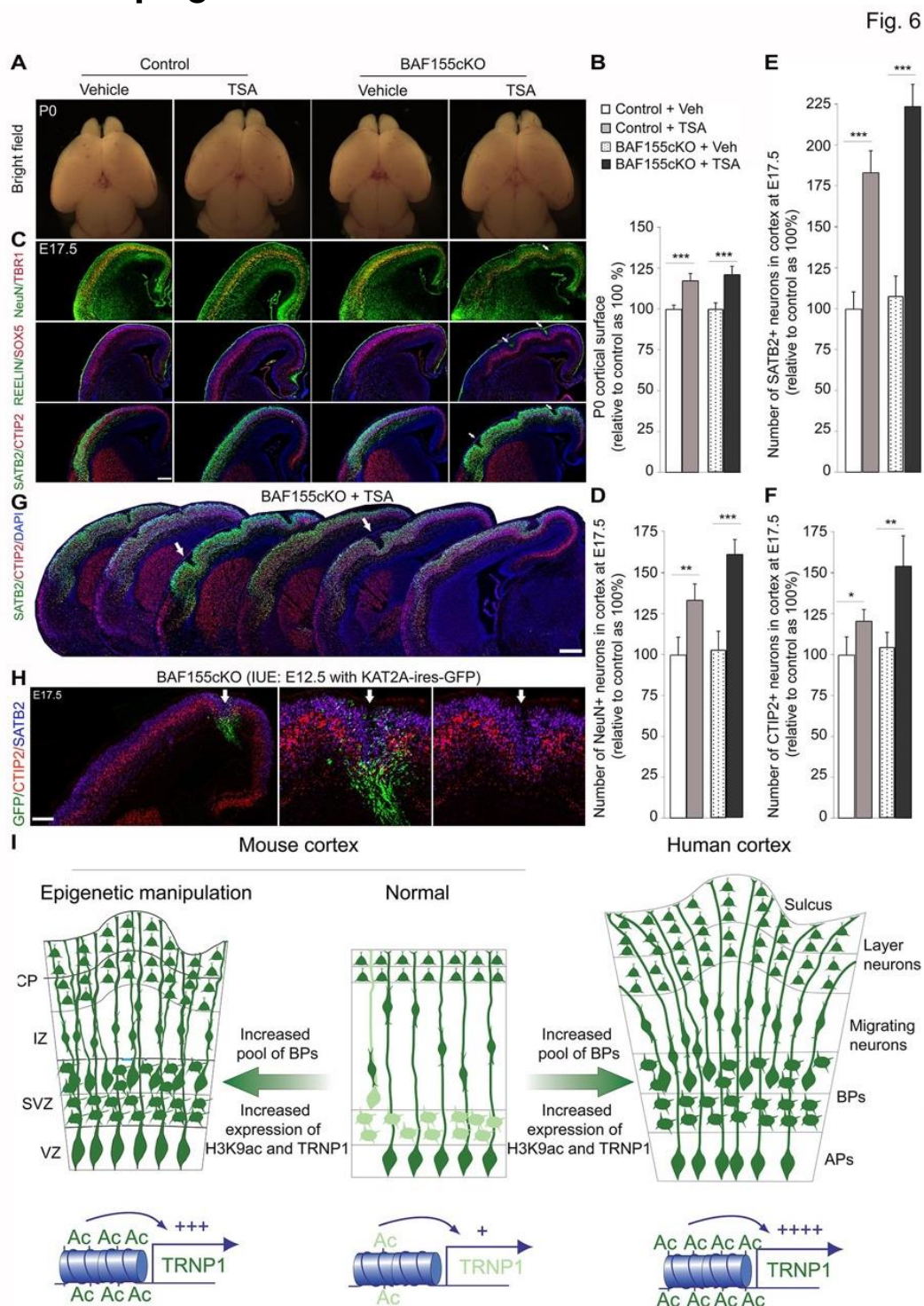
Figure 6: *Elevated H3ac levels promote cortical expansion and folding.* (A) Dorsal-view images of brains from control and BAF155cKO mice with and without TSA treatment (B) and quantification of their cortical surfaces at P0 (C, see also Supplemental methods). (C-F) IHC (C) and quantitation (D-F) of the number of neuronal subsets labeled by the neuronal markers, SATB2, CTIP2, TBR1, and NeuN, in control and BAF155cKO cortices with and without HDAC inhibitor treatment at E17.5. Elevated H3 acetylation increases neurogenesis and cortical thickness and induces gyrification (indicated by arrows). (G)

Serial coronal sections at E17.5 marked by CTIP2 and SATB2 expression showing the folded cortex in different cortical areas. (H) Images of GFP, SATB2, and CTIP2 immunofluorescence of a coronal section of an E17.5 cortical hemisphere from a mouse embryo that had been in utero electroporated at E12.5 with KAT2A-ires-eGFP expression plasmids. Lower images show higher-power magnification of a sulcus-like structure, which is indicated by white arrows. (I) Hypothetical model proposing how the changes in H3K9ac levels in evolution and in mouse models with epigenetic manipulation affect the H3K9ac level at the TRNP1 promoter, the expression of TRNP1, the proliferation of BPs, and the expansion and folding of the cortex. Symbols: +, +++, and ++++ indicate weak, moderate, and strong relative expression levels, respectively. Abbreviations: VZ, ventricular zone; SVZ, subventricular zone; IZ, intermediate zone; CP, cortical plate; APs, apical progenitors; BPs, basal progenitors; Ac, H3K9ac. Values are presented as mean ± SEM (*$p < 0.05$, ***$p < 0.01$, ***$p < 0.005$). Scale bars = 100 μm.

To address the role of H3 acetylation in late corticogenesis *in vivo*, mouse embryos were treated with TSA for a prolonged period. Remarkably, TSA-treated mice exhibited cortices with larger surface areas compared with vehicle (Veh)-injected controls (Fig. 6A/B). To evaluate the effect of the increased level of H3 acetylation on neuron production in the developing cortex, we examined the expression of laminar and neuronal subtype-specific genes at E17.5 (Fig. 6C). Immunostaining for the pan-neuronal marker, NeuN, indicated that there was a considerable increase in cortical radial thickness after TSA treatment (Fig. 6D). The number of SATB2+ callosal and CTIP2+ subcerebral projection neurons was also significantly increased in the cortex of TSA-treated embryos compared with controls (Fig. 6E/F).

Previous evidence suggested that an increased number of bRG cells offers extra scaffolding for the radial migration of neurons and causes divergence (fanning out) of radial glial processes in gyrated cortices (Lui, Hansen et al. 2011, Nonaka-Kinoshita, Reillo et al. 2013, Stahl, Walcher et al. 2013, Borrell and Gotz 2014). Indeed, RC2 and NESTIN immunostaining, which reveal the layout of RG processes, showed that in the TSA-treated cerebral cortex radial fibers spread out divergently when traversing the CP (fig. S10A, white dashed lines). Fanned-out fibers of PAX6+ bRGs were observed by labeling with the lipophilic dye, DiI, which was applied at the pia (Fig. 3D). At E17.5, most early-born neurons (L6, L5) had completed their migration. Strikingly, TSA treatment often led to mild folding of the cortex in WT embryos (fig. S10B-D), but more profound folding in BAF155cKO mutants in distinct cortical areas, as indicated by L6 (TBR1) and L5 (CTIP2)

immunostaining (Fig. 6G, S10B−D). Although HDAC inhibitor-treated mice died soon after birth, increased folding of TSA-treated cortex was observed at all examined stages (E16.5−P0) (fig. S10D). Our findings also provide further link between abundance level of BPs and folding feature in mammalian cortex as the inhibition of HDAC leads to regionally restricted increases in BPs (fig. S3) and cortical folding (fig. S10B) in sections taken from the rostral, middle, and caudal dorsolateral cortex (d/lCx), but not the medial cortex (mCx). We distinguished cortical folding by the presence of an intact basement membrane and rostro-caudal continuity of both sulci and gyri; this allowed us to distinguish our findings from the defining features of classical neuronal ectopias, such as that seen in lissencephaly type II (also known as cobblestone lissencephaly).

Finally, we investigated whether increasing the level of H3K9ac by overexpressing *KAT2A*, which encodes H3K9 acetyltransferase, might induce proliferation of BPs and cortical folding. E13.5 WT or *BAF155cKO* brains were electroporated with a *KAT2A*-expression plasmid. Compared with control (GFP-injected cortex) at E15.5, overexpression of *KAT2A* increased the percentage of proliferating BPs (GFP+/Ki67+/TBR2+ or PAX6+) among targeted BPs (GFP+/TBR2+ or PAX6+) in WT (fig. S10E) and BAF155cKO cortex (fig. S10F). Furthermore, five out of six BAF155cKO and three out of six WT cortices injected with *KAT2A* expression plasmids at E12.5 displayed focal neocortex folding at E17.5 (Fig. 6H).

Collectively, these findings suggest that increased H3 acetylation at the promoter region of the evolution-related gene, TRNP1, selectively promotes the proliferative capacity of BPs, leading to enhanced neuronal output, increased cortical expansion, and gyrification during mammalian evolution (Fig. 6I).

DISCUSSION

In this study, we describe H3 acetylation as a novel epigenetic mechanism that regulates neocortical expansion. We present evidence that human BPs have higher H3 acetylation levels than mouse BPs, and that elevated levels of H3ac preferentially promote BP self-amplification and augment neuronal output, leading to enlarged size and folding of the mouse neocortex (summarized in Fig. 7H). Mechanistically, this process involves the epigenetic and gene expression controls of the evolutionarily regulated gene, TRNP1, in cortical development.

Contribution of BP proliferation to cortical expansion and folding

In lissencephalic rodents, APs in the VZ and TBR2+ BPs in the SVZ are largely responsible for neuron production during cortical development. Previous works showed that β-catenin overexpression-induced VZ progenitor amplification in the mouse brain yielded folding of the ventricular surface, not the cortical surface (Chenn and Walsh 2002), and that increasing the pool of BPs/bIPs enlarged brain size but did not induce cortical folding (Nonaka-Kinoshita, Reillo et al. 2013, Tuoc, Boretius et al. 2013).

On the other hand, published evidence suggests that there is a link between the proportion of BPs/bRGs and the extent of cortical gyrification in various species (Hansen, Lui et al. 2010, Reillo, de Juan Romero et al. 2011, Shitamukai, Konno et al. 2011, Wang, Tsai et al. 2011). bRG progenitors are derivatives of APs that presumably after becoming delaminated from their apical anchorage (Borrell and Gotz, 2014). However, when adherens junction proteins are lost (Narayanan, Pham et al. 2018) or the linkages to cytoskeletal belts are broken by downregulation of the small GTPase, RhoA, apically anchored progenitor cells are delaminated without significant change in bRG generation or cortical folding. Likewise, our data and that of others indicate that PAX6 and BAF155 control the delamination of RGs at least partly by regulating the expressional programs of genes encoding adherens junction proteins and RhoA (Narayanan, Pham et al. 2018). Similar to many other mouse mutants that exhibit delamination of APs during corticogenesis, BAF155cKO (Narayanan, Pham et al. 2018, Xie, Castro-Hernandez et al. 2019) and PAX6cKO mice (Tuoc, Radyushkin et al. 2009, Tuoc, Boretius et al. 2013) do not exhibit an enlarged and folded cortex, suggesting that the ectopic localization of RGs outside of the VZ may be insufficient to trigger cortical expansion and/or folding. Recent

observations in the cortex of marmoset (a lissencephalic species) or agouti (a rodent with a moderately gyrencephalic cortex) revealed that the presence of an augmented population of bRGs in the expanded oSVZ of some mammalian subclasses may not correlate with the occurrence of cortical folding and gyrencephaly (Garcia-Moreno, Vasistha et al. 2012, Kelava, Reillo et al. 2012). In gyrencephalic ferrets, for example, the oSVZ initially appears as a massing of numerous bRGs that are directly produced by the transient delamination of APs during early cortical development (Martinez-Martinez, De Juan Romero et al. 2016). Later, however, the generation of bRGs becomes completely independent of AP delamination, instead relying on their self-expansion (Martinez-Martinez, De Juan Romero et al. 2016). Along the same lines, functional blockade of TBR2 in ferrets was shown to cause premature neuronal differentiation of SVZ progenitors, diminishing the numbers of bIPs and bRGs and impairing gyrification (Toda, Shinmyo et al. 2016). These findings support the idea that BP proliferation is a key driver of cortical folding (Toda, Shinmyo et al. 2016). Thus, the simple delamination of VZ progenitors is insufficient to generate BPs with a high proliferative capacity; instead, this process seems to critically involve altered expression of regulatory factors that boost BP amplification.

H3 acetylation promotes BP proliferation and cortical expansion

BPs from rodents and primates have distinct characteristics. In the lissencephalic rodent cortex, BPs are mainly neurogenic progenitors that express the transcription factor (TF), TBR2. In gyrencephalic species (such as ferret, primate, and human), in contrast, most of the BPs are proliferative progenitors and almost half of them co-express the TFs, TBR2 and PAX6 (Lui, Hansen et al. 2011). In the developing ferret cortex, TBR2 is required for the production of both types of BPs (bIPs and bRGs) (Toda, Shinmyo et al. 2016). A prerequisite for identifying the gene expression and epigenetic programs of BPs during evolution is thus the analysis of BP-specific signatures in distinct species. Recent single-cell transcriptomic and genetic studies have sought to explain the phylogenic expansion and gyrification of the mammalian cortex (Arai, Pulvers et al. 2011, Fietz, Lachmann et al. 2012, de Juan Romero, Bruder et al. 2015, Florio, Albert et al. 2015, Johnson, Wang et al. 2015, Pollen, Nowakowski et al. 2015, Thomsen, Mich et al. 2016). A few factors (such as TRNP1, ARHGAP11B, and TBC1D3) have been found to control BP genesis, BP proliferation, cortical expansion, and folding (Nonaka-Kinoshita, Reillo et al. 2013, Stahl, Walcher et al. 2013, Tuoc, Boretius et al. 2013, Bae, Tietjen et al. 2014, Florio, Albert et al.

2015, Ju, Hou et al. 2016, Wang, Hou et al. 2016, Del Toro, Ruff et al. 2017). Recently, remodeling of chromatin via histone post-translational modifications has also been implicated in cortical neurogenesis (Albert and Huttner 2018). However, the epigenetic mechanisms that coordinate the expression or repression of genes that are essential for the developmental events underlying cortical expansion and folding during evolution are still largely unknown.

In this study, we purified TBR2+ BPs from mouse and human cortices. To compare the bulk level of epigenetic marks between mouse and human BPs, we quantified the PTMs of histones by simultaneously measuring the relative abundance of most known histone methylation and acetylation marks. This method requires a moderate number of cells that can be obtained by FACS, and allows for the systematic screening of epigenetic changes during evolution. We identified a few epigenetic marks that display different levels in TBR2+ BPs isolated from mouse versus human cortex. As PAX6+, SOX2+ BPs/bRGs are relatively rare in the WT mouse cortex (Wang, Tsai et al. 2011, Vaid, Camp et al. 2018), we used the BAF155cKO mutant as a mouse model to investigate the effect of HDAC inhibition on the proliferation of BPs.

Based on the expression of PAX6, AP2γ, and SOX2 in the VZ and pHH3 at the apical VZ surface (Fig. 4A/C), our data suggest that the elevation of H3ac in developing mouse cortex did not influence the proliferation and delamination of APs, which were consistent with a relatively high level of H3K9ac in both mouse and human APs. Intriguingly, we found that the level of H3K9ac is higher in human BPs than in mouse BPs. Moreover, we found that elevated H3K9ac specifically expands the pool of BPs, but not APs, and increases the number of both lower-layer CTIP2+ neurons and upper-layer SATB2+ neurons.

The gyrification formation involves in generating a big population of neurons and the sufficient scaffold for neuronal migration in lateral dispersion manner into cortical layers. In gyrencephalic species, the cortical expansion is achieved by an extraordinary increase in neurogenic bIPs and the neurogenic/scaffold bRG in the oSVZ, which are critical to lead radially migrating neurons into lateral dispersion (Lui, Hansen et al. 2011, Borrell and Gotz 2014, Taverna, Gotz et al. 2014, Dehay, Kennedy et al. 2015). The generation of more bIP and bRGs in TSA- treated BAF155cKO cortex than in TSA- treated WT cortex could explain the corresponding difference in cortical size (Fig. 6A-F) and level of cortical folding (fig.

S10B) in two mouse models. In agreement with previous studies (Garcia-Moreno, Vasistha et al. 2012, Kelava, Reillo et al. 2012, Liu, Liu et al. 2017), our findings suggest that the formation of cortical folding requires a great abundance of both bIPs and bRGs. Together, our findings indicate that H3 acetylation has prominent impacts on the cerebral cortex, in terms of both radial and tangential expansion (Figure 7H), by increasing the population of highly proliferative BPs and thus the degree of cortical gyrification (as seen in the primate brain).

TRNP1 is a main target gene of H3 acetylation in BP amplification

Our RNA-Seq and H3K9ac ChIP-Seq analyses for many genes in the sorted TBR2+ BPs indicate that HDAC inhibition either increases the H3K9ac level or upregulates the expression of the studied genes. Interestingly, the correlation of these data indicates that H3K9ac directly activates the expression of only a few genes (Fig. 4F-H). Such genes include TRNP1, whose manipulated expression in the mouse and ferret cerebral cortex was previously reported to affect various cortical expansion-related features, including the proliferation of cortical progenitors and the induction of cortical folding (Pilz, Shitamukai et al. 2013, Stahl, Walcher et al. 2013, Martinez-Martinez, De Juan Romero et al. 2016). We herein show that H3K9ac positively regulates the expression of TRNP1 in BPs to orchestrate BP proliferation in the developing cortex. Expressional analyses indicate that the expression patterns of H3K9ac (this study) and TRNP1 (Stahl, Walcher et al. 2013) are similar in developing mouse and human cortex. In particular, H3K9ac and TRNP1 exhibit low or undetectable levels in mouse BPs, whereas their expression is much higher in human BPs. Furthermore, we show that the overexpression of TRNP1 by means of H3K9ac epigenome editing increases BP proliferation, as also seen in HDAC inhibitor-treated cortex, and that this phenotype is largely reverted to control levels upon TRNP1 knockdown. These findings bring to light the role of H3 acetylation as a mechanism upstream of TRNP1 function for BP proliferation in cortical development during evolution. In summary, we herein identify a mechanism whereby BP amplification is epigenetically controlled through H3 acetylation, which specifically affects the expression level of TRNP1. Our findings show that the temporally dynamic regulation of H3 acetylation and its downstream effector, TRNP1, are part of an evolutionary epigenetic mechanism aimed at enlarging the diversity of neocortical phenotypes during evolution.

MATERIALS AND METHODS

Plasmids

Plasmids used in this study: pCIG2-KAT2A-ires-eGFP, gTRNP1-dCas9-KAT2A-T2A-eGFP (this study), shTRNP1 constructs (Stahl, Walcher et al. 2013).

Antibodies

The following polyclonal (pAb) and monoclonal (mAb) primary antibodies used in this study were obtained from the indicated commercial sources: H3K9Ac rabbit pAb (Abcam), SOX2 mouse mAb (1:100; R&D Systems), SOX2 goat (1:100; Santa Cruz), PAX6 mouse mAb (1:100; Developmental Studies Hybridoma Bank), PAX6 rabbit pAb (1:200; Covance), TBR2 rabbit pAb (1:200; Abcam), CidU rat pAb (1:100; Accurate), KI67 rabbit pAb (1:50; Vector), GFP rabbit pAb (1:1000; Abcam), GFP chick pAb (1:1000; Abcam), AP2γ mouse mAb (1:100; Abcam), phospho-H3 rat pAb (1:300; Abcam), pVIM mouse mAb (1:500; MBL), TNC rabbit pAb (Abcam), PTPRZ1 rabbit pAb (Sigma), BAF170 rabbit pAb (Bethyl), BAF170 rabbit pAb (Sigma), BAF155 rabbit pAb (1:20; Santa Cruz), BAF155 mouse mAb (1:100; Santa Cruz), CASP3 rabbit pAb (1:100; Cell Signaling), CTIP2 rat pAb (1:200; Abcam), HuCD mouse mAb (1:20; Invitrogen), TBR1 rabbit pAb (1:300; Chemicon), SATB2 mouse mAb (1:200; Abcam), CUX1 rabbit pAb (1:100; Santa Cruz), β-actin rabbit pAb (Sigma), H3ac rabbit pAb (Upstate), NeuN mouse mAb (Chemicon), REELIN mouse mAb (Gift from Prof. Goffine), SOX5 rabbit pAb (Santa Cruz), KAT2A rabbit pAb (Abcam), NESTIN mouse mAb (BD), RC2 mouse mAb (Developmental Studies Hybridoma Bank), Secondary antibodies used were horseradish peroxidase (HRP)-conjugated goat anti-rabbit IgG (1:10000; Covance), HRP-conjugated goat anti-mouse IgG (1:5000; Covance), HRP-conjugated goat anti-rat IgG (1:10000; Covance), and Alexa 488-, Alexa 568-, Alexa 594- and Alexa 647-conjugated IgG (various species, 1:400; Molecular Probes).

Human fetal brain collection and processing

Human fetal brain tissue was obtained from spontaneous abortions that occurred in the Hospital for Obstetrics and Gynecology "Prof. Dimitar Stamatov", Medical University – Varna, Bulgaria, after an informed written maternal consent, and with the approval of the local Ethics Committee (Protocol No 19/April 2012; Protocol No 55/June 2016). The gestation age in weeks (GW) has been determined based on the history of last

menstruation as reported by the patients. The heads were placed in 4% PFA solution in PBS 7.5 pH for 24h, and then the brains were dissected and postfixed in PFA for 5-7 days, cryoprotected in sucrose and frozen in OCT medium prior to cryosectioning at 20 μm. Cortical tissues from additional cases were used for FACS as described for mouse tissues in the below section.

Animal care, generation of transgenic mice, *in utero* electroporation

Floxed BAF155 (Choi, Ko et al. 2012), Emx1-Cre (Gorski, Talley et al. 2002) mice were maintained in a C57BL6/J background. *In utero* electroporation was performed as described previously (Tuoc and Stoykova 2008, Xie, Castro-Hernandez et al. 2019). Animals were handled in accordance with the German Animal Protection Law and with the permission of the Bezirksregierung Braunschweig.

Mice with treatment of HDAC inhibitors (HDACi)

Trichostatin A (TSA) (Sigma-Aldrich, Cat. T8552-1MG) was dissolved in vehicle (8% ethanol in 1xPBS) at a concentration of 100μg/ml. Suberoylanilide hydroxamic acid (SAHA) (Biomol, Cat. CAS 149647-78-9) was dissolved in Vehicle (DMSO) (10 mg/ml). Sodium salt Valproic acid (VPA) (Sigma, Cat. P4543) was dissolved in Vehicle (saline) at concentration of 100 mg/ml. E12.5 d.p.c. pregnant females were injected intraperitoneally twice on the first day, then once from second day with either vehicle or 150μl of 100 μg/ml TSA solution or 20μl of 10 mg/ml SAHA solution plus 110 μl Saline or 120μl of 100 mg/ml VPA solution. Treated mice were sacrificed at different developmental stages as indicated in the text.

TBR2+ nuclei sorting from embryonic cortex

Since typical cell sorting methods comprise of protease treatment for cell dissociation that can lead to unwanted effects, we opted for unbiased approach to minimize biasness due to sample processing. In this regard, we opted for nuclei sorting, instead of cell sorting, as this is a well-established method of getting the cell types of interest from the brain without the biasness of the sample preparation (Halder, Hennion et al. 2016). Since we were interested to obtain epigenetic and transcriptomic data from TBR2+ and TBR2- nuclei, two different protocols were followed, respectively.

TBR2+ nuclei sorting protocol from embryonic mouse brain for ChIP-Seq

The protocol were adapted from (Halder et al., 2016) with modifications. Freshly prepared embryonic cortices from 5 CD1 pups were pooled for each replicates and homogenized briefly in low-sucrose buffer(320 mM Surcrose, 5mM CaCl2, 5 mM MgAc2, 0.1 mM EDTA, 10mM HEPES pH 8, 0.1% Triton X-100, 1mM DTT, supplemented with Roche protease inhibitor cocktail) with plastic pestles in 1.5mL tubes. 1% Formaldehyde was added as a final concentration and incubated at room temperature for 10 minutes on a rotating wheel. 125mM Glycine was added as final concentration and incubated for 5 minutes at room temperature for quenching remaining formaldehyde. The homogenate was centrifuged at 2000g for 3 minutes at 4°C and supernatant was discarded. The remaining crude nuclear pellet was resuspended in additional low-sucrose buffer and further homogenized with a mechanical homogenizer (IKA Ultraturax T10, with S10N-5G tool). The solution was carefully layered onto 6mL high-sucrose buffer(1000mM Sucrose, 3mM MgAc2, 10mM HEPES pH 8, 1 mM DTT, Roche protease inhibitor) in oak-ridge tubes and centrifuged at 3220g for 10 minutes at 4°C in a swinging bucket rotor centrifuge to get rid of myelin debris. The resulting nuclear pellet were resuspended into left over buffer, transferred into 2mL microfuge tubes(DNA-low bind), centrifuged for 3 minutes at 2000g at 4°C to recover the nuclear pellet. It was resuspended into 500uL PBTB buffer(PBS-Tween-BSA buffer, 1% BSA, 0.2% Tween-20, protease inhibitor in 1X PBS) and 5uL Anti-TBR2(Eomes)-Alexa488 conjugated antibody(IC8889G-025) was added. After incubating for 1 hour at 4°C, the samples were washed with PBTB once and proceeded with the sorting in FACSAria III with 85µm nozzle. Sorted nuclei were collected into PBTB coated falcon tubes, pelleted with brief centrifugation and pellets were flash frozen and saved at -80°C until further processing for ChIP-seq. ChIP was performed according to (Halder et al., 2016) with desired antibodies and the library preparation was performed using NEBNext Ultra II DNA library preparation according to manufacturer's protocol.

TBR2+ nuclei sorting protocol from embryonic mouse brain for RNA isolation and RNA-Seq

Freshly dissected mouse embryonic cortices from 5 CD1 pups were pooled and immediately submerged into enough RNAlater solution in a microfuge tube and kept at 4°C for at least 24 hours. The excess RNAlater solution were pipetted out and washed twice with 1x RNAse free PBS. After last wash, all the steps were done on ice. The tissues were immediately submerged into 500 uL nuclei EZ prep lysis buffer(Sigme, NUC101), dounce homogenized with plastic pestles for 30-45 times and additional lysis buffer was

added to make the volume upto 2 mL. Homogenates were incubated on ice for 7 minutes, then centrifuged for 5 minutes at 500g, 4°C. After removing the supernatant, pellet was resuspended into 2mL lysis buffer and incubated on ice for 7 minutes. Lysates were filtered through 40µm filter into a new 2mL tube and centrifuged for 5 minutes at 500g, 4°C. The pellet was resuspended into 1800uL nuclei suspension buffer(NSB, 1x RNAse-free PBS, 0.5% RNAse-free BSA, 1:200 RNaseIN plus inhibitor, Roche protease inhibitor), centrifuged at 500g for 5 minutes at 4°C and the pellet was resuspended in 500uL NSB. 5uL Anti-TBR2(Eomes)-Alexa488 conjugated antibody(IC8889G-025) was added and incubated for 1 hour at 4°C. The stained nuclei were washed once and resuspended into NSB and sorted with FACSAria III with 85µm nozzle. Sorted nuclei were collected into NSB coated falcon tubes, pelleted with brief centrifugation and the RNA was isolated using Trizol LS(Invitrogen) protocol along with aqueous phase cleanup using Zymo RNA clean & concentrator-5 kit. RNAseq libraries were prepared using Takara SMART-Seq v4 Ultra Low Input RNA kit using 1ng of RNA according to the manufacturer's protocol.

Mass Spectrometry analysis of epigenetic marks.

Histone extraction and preparation for mass spectrometry

Bulk histones were acid-extracted from the sorted TBR2+ nuclei pellets, propionylated and subjected to trypsin digestion as described previously (Zheng, Thomas et al. 2013). Briefly, histones were extracted by incubating samples at room temperature for 1 hour in 0.2M sulfuric acid with intermittent vortexing. Histones were then precipitated by the addition of trichloroacetic acid (TCA) on ice, and recovered by centrifugation at 10,000 x g for 5 minutes at 4°C. The pellet was then washed once with 1mL cold acetone/0.1% HCl and twice with 100% acetone, and then air dried in a clean hood. The histones were propionylated by adding 1:3 v/v propionic anhydride/2-propanol and incrementally adding ammonium hydroxide to keep the pH around 8, and subsequently dried in a SpeedVac concentrator. The pellet was then resuspended in 100mM ammonium bicarbonate and adjusted to pH 7-8 with ammonium hydroxide. The histones were then digested with trypsin resuspended in 100mM ammonium bicarbonate overnight at 37°C, and dried in a SpeedVac concentrator. The pellet was resuspended in 100mM ammonium bicarbonate and propionylated a second time by adding 1:3 v/v propionic anhydride/2-propanol and incrementally adding ammonium hydroxide to keep the pH around 8, and subsequently

48

dried in a SpeedVac concentrator. Histone peptides were resuspended in 0.1% TFA in H2O for mass spectrometry analysis.

Mass spectrometry

Samples were analyzed on a triple quadrupole (QqQ) mass spectrometer (Thermo Fisher Scientific TSQ Quantiva) directly coupled with an UltiMate 3000 Dionex nano-liquid chromatography system. Peptides were first loaded onto an in-house packed trapping column (3cm×150μm) and then separated on a New Objectives PicoChip analytical column (10 cm×75 μm). Both columns were packed with New Objectives ProntoSIL C18-AQ, 3μm, 200Å resin. The chromatography gradient was achieved by increasing percentage of buffer B from 0 to 35% at a flow rate of 0.30 μl/min over 45 minutes. Solvent A: 0.1% formic acid in water, and B: 0.1% formic acid in 95% acetonitrile. The QqQ settings were as follows: collision gas pressure of 1.5 mTorr; Q1 peak width of 0.7 (FWHM); cycle time of 2 s; skimmer offset of 10 V; electrospray voltage of 2.5 kV. Targeted analysis of unmodified and various modified histone peptides was performed. This entire process was repeated three separate times for each sample.

Data analysis

Raw MS files were imported and analyzed in Skyline with Savitzky-Golay smoothing (MacLean, Tomazela et al. 2010). All Skyline peak area assignments for monitored peptide transitions were manually confirmed. A minimum of 3 peptide transitions were quantified for each modification. For each monitored amino acid residue, each modified (and unmodified) form was quantified by first calculating the sum of peak areas of corresponding peptide transitions; the sum of all modified forms was then calculated for each amino acid to represent the total pool of modifications for that residue. Finally, each modification is then represented as a percentage of the total pool of modifications. This process was carried out for each of the three separate mass spec runs, and the raw data provided in the data delivery spreadsheet corresponds to the mean and standard deviation of the resulting three values from this analysis for each modified and unmodified form of the corresponding amino acid residue.

To generate the heatmaps, for each individual modification, the data for each sample was converted to the fraction of the sum across all samples to display relative abundances

across the sample group, and then conditionally formatted in Excel using the default red/white/blue color scheme. The hierarchical clustering heatmap was generated in R using the "pheatmap" function with default settings.

Chromatin immunoprecipitation (ChIP)

The ChIP protocol was described previously (Narayanan et al., 2015). Mouse embryonic cortices or TBR2+ BPs, TBR2- cells were homogenized in sucrose solution (0.32 M sucrose, 5 mM $CaCl_2$, 5 mM $Mg(Ac)_2$, 0.1 mM EDTA, 50 mM HEPES pH 8, 1 mM DTT, 0.1% Triton X-100), fixed in 37% formaldehyde, and then treated with 1.25 M glycine. The samples were then centrifuged, and the pellet (i.e., nuclei) was washed with Nelson buffer (140 mM NaCl, 20 mM EDTA pH 8, 50 mM Tris pH 8, 0.5% NP-40, 1% Triton X-100). Thereafter, nuclei were re-suspended in RIPA-SDS buffer (140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 10 mM Tris pH 8, 1% SDS) and incubated on a rotating wheel for 10 min at 4°C. The samples were then transferred to sonication tubes (Diagenode) and sonicated in a Bioruptor Plus NGS (Diagenode) at a setting of "High" at 4°C. Samples were spun down after every fifth cycle to ensure homogeneous shearing. After sonication, samples were centrifuged at 18,000 × g for 5 min and the supernatant was transferred to a fresh tube. The sheared chromatin was snap frozen and kept at -80°C prior to any further treatment. A small amount was kept in order to test shearing efficiency and total chromatin amount using an Agilent 2100 Bioanalyzer (Agilent Technologies) and QuBit fluorometer (Life Technologies), respectively, after reverse-crosslinking. To that end, samples were treated with 0.1 µg/µl RNase A (Qiagen) and Proteinase K (Roth).

Protein A-coated Dynabeads (Life Technologies) were blocked in 500 µl of 0.5% bovine serum albumin (BSA) for 2 h at 4°C on a rotating wheel. The chromatin was diluted 10-fold in immunoprecipitation (IP) buffer (140 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 50 mM Tris pH 8, 20 mM EDTA, 0.1% SDS) and pre-cleared by incubating with 20 µl of blocked Dynabeads for 1 h at 4°C on a rotating wheel. For immunoprecipitation for the purpose of ChIP sequencing, 1000 ng of pre-cleared chromatin was incubated overnight at 4°C on a rotating wheel with 2 µg of anti-H3K9ac antibody (Millipore). For ChIP-qPCR, 500 ng of pre-cleared chromatin was used with 1 µg of the same antibody. The next day, the samples were incubated with 15 µl of blocked Dynabeads for 2 h at 4°C on a rotating wheel. Beads were then washed with IP buffer and

wash buffer (100 mM Tris pH 8, 500 mM LiCl, 1% NP-40, 1% sodium deoxycholate, 20 mM EDTA), and collected on magnetic stands (Invitrogen). DNA was eluted with 20 µl elution buffer (10 mM Tris pH 8) containing 0.1 µg/µl RNase A. Samples were incubated for 30 min at 37°C in a thermomixer (Eppendorf) with gentle agitation, then diluted 2-fold with Wiemann Buffer (100 mM Tris pH 8, 20 mM EDTA, 2% SDS) containing 1 µl of Proteinase K (20 µg/µl) and incubated overnight at 65°C in a thermomixer (Eppendorf) with agitation. The next day, 3 µl of linear polyacrylamide (Bioline) and 60 µl of SureClean (Bioline) were added to the samples. Samples were then transferred to magnetic stands, the beads were washed with 80% ethanol, and DNA was eluted with 30 µl elution buffer for ChIP-Seq experiment or with 60 µl elution buffer for ChIP-qPCR. Inputs were handled in parallel with immunoprecipitated samples. Finally, DNA quality and quantity were measured on an Agilent 2100 Bioanalyzer and QuBit fluorometer, respectively.

In ChIP/qPCR experiment (Fig. 6C, Fig. S6C) to compare the H3K9ac level of mouse and human TBR2+ BPs, identical conditions (ChIP, PCR) were performed.

ChIP-Seq

Libraries were prepared with NEBNext Library Prep Kit for Illumina (NEB). Input DNA (100 ng) was isolated from all the samples and was pooled for each group separately. The quality and quantity of the libraries were measured with an Agilent 2100 Bioanalyzer and QuBit fluorometer, respectively.

Base calling and conversion to fastq format were performed using Illumina pipeline scripts. Afterwards, quality control on raw data was conducted for each library (FastQC, www.bioinformatics.babraham.ac.uk/projects/fastqc). The following control measurements and information were obtained: per base sequence quality, per sequence quality scores, per base sequence content, per base GC content, per sequence GC content, per base N content, sequence length distribution, sequence duplication levels, overrepresented sequences, Kmer content.

The reads were mapped to a mouse reference genome (mm10) using STAR aligner v2.3.0 (Djebali, Davis et al. 2012). *rmdup* function of samtools (Li, Handsaker et al. 2009) was used to remove PCR duplicates from each BAM file. *merge* function of samtools was used for merging BAM files with unique reads (i.e., with duplicates removed) belonging to replicates from the same group into a single BAM file. All the downstream analyses were performed on BAM files with only unique reads.

Profile plots of H3K9ac were created with NGSPlot (Shen, Shao et al. 2014) using merged BAM files from immunoprecipitated samples and inputs. H3K9ac enrichment at different gene loci was visualized through the Integrated Genome Browser (Nicol, Helt et al. 2009) using wiggle files that were created from the merged BAM files with the script from the MEDIPS package of Bioconductor (Lienhard, Grimm et al. 2014).

Peaks were called on individual with MACS2 with q-value < 0.1(Feng, Liu et al. 2011). Differential binding was assessed with DiffBind package of Bioconductor (Ross-Innes, Stark et al. 2012) with in-built DESEQ2 option implemented in differential analysis. In differential binding analyses promoters were defined as +/- 2000 bp from transcription start site (TSS). Peak annotation was done with HOMER (Heinz, Benner et al. 2010).

RNA-Sequencing

RNA was extracted (RNeasy kit; Qiagen) from the embryonic cortex of E15.5 embryos. cDNA libraries were prepared using the TruSeq RNA Sample Preparation v2 Kit. DNA was quantified using a Nanodrop spectrophotometer, and its quality was assessed using an Agilent 2100 Bioanalyzer.

Base calling, fastq conversion, quality control, and read alignments were all performed as outlined for ChIP-Seq. Reads were aligned to mouse genome mm10 and counted using FeaturesCount (http://bioinf.wehi.edu.au/featureCounts/). Differential expression was assessed using DESeq2 from Bioconductor (Love, Huber et al. 2014). Functional GO enrichment analyses were performed using ToppGene (Chen, Bardes et al. 2009).

Epigenome editing

Constructs for epigenome editing

The backbone vector encoding sgRNAs and 3xFlag-dCas9 (nuclease dead-Cas9)-T2A-GFP expression cassette was generated from the vector pSpCas9n(BB)-2A-GFP (PX461 in addgene): H840A was introduced into Cas9n to produce dCas9 through PCR-mediated mutagenesis cloning, and a polylinker sequence 5'-TCCGGACGGGGATCCACTAGTGTCGACACCGGTCCTAGG-3' was inserted right after the Cas9 coding sequence. The vectors for the mouse TRNP1 promoter epigenome editing were derived from the above generated backbone vector. gRNA sequences targeting specific mouse TRNP1 promoter region (Chr4: 133497514-133498529) was designed using the software 'Genious', and cloned into the vector between BbsI sites after in vitro

test, then the coding sequence for mouse KAT2A was amplified from the vector pCMV-sport2-mGCN5 (Martínez-Balbás, Bauer et al. 2000) (gift from Sharon Dent, Addgene plasmid # 23098), and cloned into the backbone vector between SpeI and AvrII sites, in frame with that of dsCas9 and T2A-GFP. The vectors were confirmed by sequencing.

In vitro test for TRNP1 sgRNAs, quantification of H3K9ac level at TRNP1 promoter and expression of TRNP1 in gTRNP1-dCas9-KAT2A- transfected Neuro2A

For in vitro test, sgRNAs were synthesized through in vitro transcription from a PCR-produced sgRNA template, which habors the T7 promoter abided by the gRNA sequence and the sharped gRNA scaffold (Chen, Gilbert et al. 2013) using MEGAscript T7 Transcription Kit (Invitrogen) according the manufacturer's protocols. Cas9 protein was purchased from IDT (#1074182). In vitro test of Cas9/gRNAs complexes cutting efficiency was performed using a 1016 bp PCR product from TRNP1 promoter according the protocol described in (Kalebic, Taverna et al. 2016).

Based on the in vitro testing result, the gRNA2 and gRNA4 were selected and cloned into above mentioned vector to generate vectors TRNP1-sg2-Flag-dCas9-2A-GFP (TRNP1-sg2), and TRNP1-sg4-Flag-dCas9-2A-GFP (TRNP1-sg4). 12 µg of parental vector Flag-dCas9-2A-GFP or TRNP1-sg2, or/and TRNP1-sg4 were transfected into Neuro2A cells cultured on 10 cm dishes using lipofectamine 2000 reagent (thermofisher) following the manufacture's protocol, and cells were harvested for FACS analysis 3 days post-transfection for qPCR and ChIP/qPCR analyses.

Cell cultures

Cortical primary cells were dissociated from mouse embryos and cultured as described previously (Conti, Pollard et al. 2005, Tuoc and Stoykova 2008). For cell culture assays, plasmids were transfected into Neuro2A cells using Lipofectamine 2000, or were electroporated into primary cortical cells using a mouse neural cell Nucleofector kit and a nucleotransfection device (Amaxa).

IHC and cell-cycle parameter experiments

IHC and determination of cell-cycle index were performed as previously described (Tuoc, Radyushkin et al. 2009).

DiI labeling

Mouse brains were dissected out in cold PBS and fixed in cold 4% PFA in PBS at 4°C overnight, followed by removal of the meninges that covers the cortical surface. The DiI crystals (Life Technologies) were dissolved in 100% ethanol to a final concentration of 1 mg/ml in 1 ml final volume. Then each brain sample was transferred to 1 ml fresh 4% PFA in PBS added with 30 ml of 1 mg/ml DiI solution and incubated at 37°C for another 24 hr. After DiI labeling, brains were washed with PBS, sectioned on the vibratome (Leica VT1200S) into 100 mm slices, and stained with PAX6 antibody and DAPI before mounting.

Retrovirus production and viral infection in slices

GFP expressing retrovirus was produced from the stably transfected packaging cells (293gp NIT-GFP) by transfection with pCMV-VSV-G plasmid (Addgene #8454) (Stewart, Dykxhoorn et al. 2003). Viral particles with their low titer of about 10^6 PFU/ml were used for infection. For the viral infection, cortical slices from brain of TSA- treated BAF155cKO embryos were prepared and cultured as described previously (Wang, Tsai et al. 2011). To label the bRGs, GFP-expressing retrovirus (106 PFU/ml) was microinjected into intermediate zone of the cultured brain slice, using a beveled and calibrated glass micropipette (Drummond Scientific). After 60 hours post-infection, brain slides were washed with PBS and stained with antibodies against GFP and PAX6.

qRT-PCR and WB analyses

qRT-PCR and WB analyses were performed as described previously (Tuoc and Stoykova 2008) using primers described in Supplementary Table S7 and the RT² Profiler PCR Array profiles (Qiagen).

Relative quantification of cortical surface

The cortical surface was measured and analysed as described by O'Leary et al (Bishop, Rubenstein et al. 2002, Bishop, Garel et al. 2003, Sahara and O'Leary 2009, Tuoc, Boretius et al. 2013, Narayanan, Pham et al. 2018). Shortly, forebrains of mutants and controls at

P0 with their presented dorsal view were photographed under a dissection microscope. The cortical surface from the digitized images were processed to compare this parameter between mutants and controls by using NIH ImageJ software.

Cell counts and quantitative analysis of immunohistochemical signal intensity

For the quantitative analyses of immunohistochemical signal intensity of H3K9ac and H3ac, confocal fluorescent images of sections of mouse and human cortex were used. The IHC Images with corresponding IgG isotopes were used as immune-staining controls. The color images of cortex were converted to gray scale to eliminate background. The pixel values of the fluorescent signal intensity were measured by using Analyze/Analyze Particles function (ImageJ software) as previously described (Tuoc and Stoykova 2008, Tuoc, Boretius et al. 2013, Narayanan, Pirouz et al. 2015, Bachmann, Nguyen et al. 2016). The measured value was then subtracted from immune-staining controls (without primary antibody).

Similarly, the relative amount of protein from developed films in WB experiment was quantified densitometrically using ImageJ software as described previously (Tuoc and Stoykova 2008, Tuoc, Boretius et al. 2013, Narayanan, Pirouz et al. 2015, Bachmann, Nguyen et al. 2016).

Immunostaining in IHC images was quantified using anatomically matched forebrain sections from control and mutants or HDACi-treated embryos. Nucleus-marker−positive cells within the cortex of confocal images were counted for comparison. In most cases, cell counts of six matched sections were averaged from three biological replicates. Statistical analyses of histological data were performed using Student's t-test. All bar graphs are plotted as means ± SEM. All statistical tests are two-tailed, and P-values are considered to be significant for α = 0.05. Details of statistical analyses of histological experiments are presented in Table S8.

Image acquisition and statistical analysis

Images were acquired with epifluorescence (Leica DM 6000) and confocal microscopes (Leica TCS SP5). Images were further analyzed with Adobe Photoshop. Statistical significance was determined by Student's t-test or Mann-Whitney test. All graphs are plotted as mean ± SEM. The statistic quantification was carried out as average from at

least three biological replicates. All details of statistical analyses for histological experiments are presented in Table S8.

Data availability

All RNA-Seq and ChIP-Seq data have been deposited in GEO and will be released to public upon acceptance of the manuscript.

SUPPLEMENTAL DATA

Fig. S1
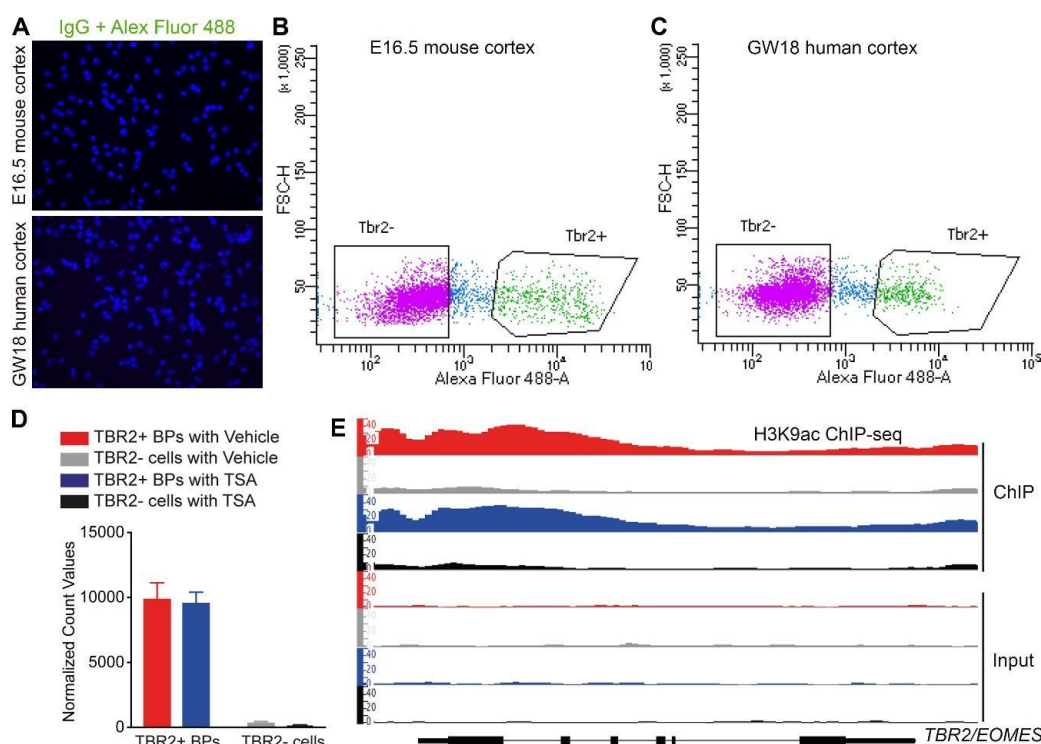
Fig. S1. Purification and characterization of TBR2+ BPs. (A) Images showed stained samples as isotype control (see also Fig. 1B/C). (B, C) Representative plot showing sorting gates for TBR2+ cells and TBR2- cells from mouse cortex at E16.5 (B) and human cortex at GW18 (C). (D, E) Purity of the sorted TBR2+ BPs and TBR2- cells was further proven by RNA-Seq (D) and H3K9ac ChIP-Seq (E). TBR2+ BPs had much higher levels of TBR2 RNA expression (D) and much higher levels of H3K9ac at TBR2 gene locus (E) compared to TBR2- cells.

Fig. S2
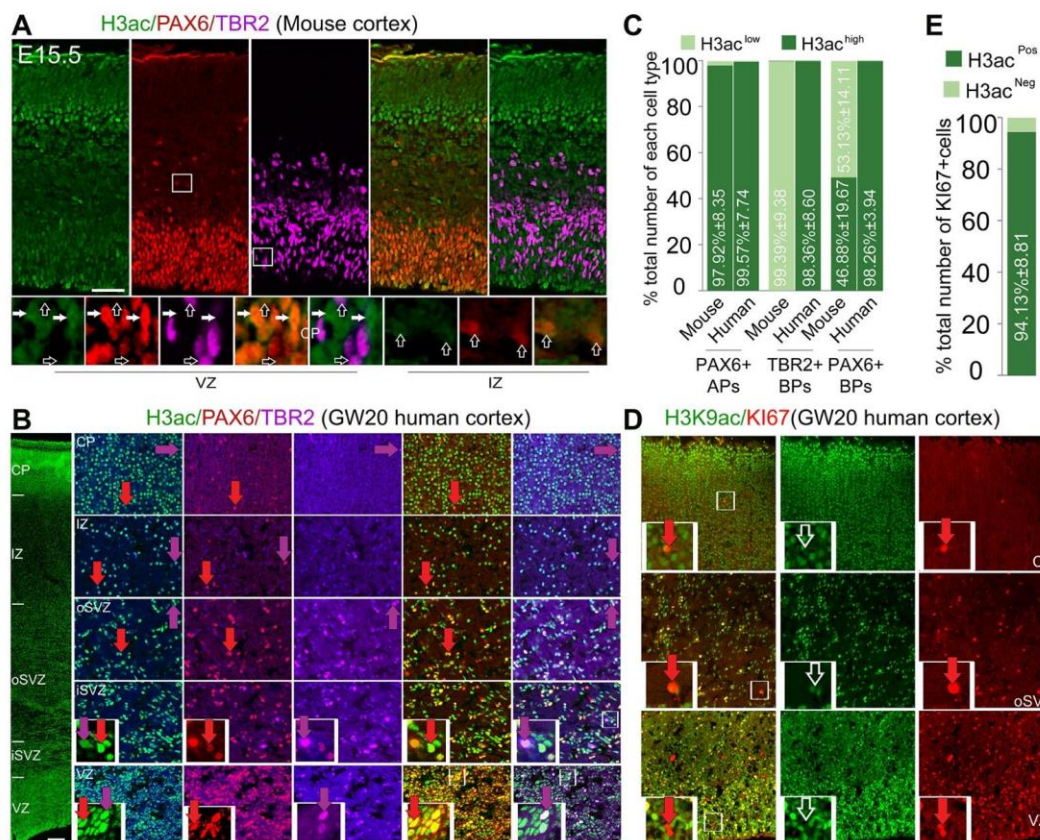
Fig. S2. Histone H3 is acetylated differently in basal progenitors in murine and human developing cortex. (A) Images of triple IHC with cortical section from mouse embryo at E15.5 stained with antibodies against: H3ac, PAX6, and TBR2. The lower panels showing selected frames at VZ and IZ at higher magnification revealed that PAX6^{high+}, TBR2^{low+} APs are H3ac^{high+} cells (filled arrows), whereas many PAX6^{high+} BPs in IZ and all TBR2^{high+} BPs are H3ac^{low+} cells (empty arrows). (B) Images showing triple immunolabeling with antibodies against H3ac, PAX6 and TBR2 and human fetal cortex at GW20 demonstrate that both the PAX6+ APs and PAX6+ BPs (arrows in red) and the TBR2+ BPs (arrows in magenta) are highly immunoreactive with H3ac. (C) Statistical analyses of IHC assay (shown in A and B) comparing the level of H3ac (H3ac^{high} and H3ac^{low}) in progenitor subtypes in developing mouse and human cortex confirm the high level of H3ac mark in both APs and BPs in humans. (D, E) IHC (D) and statistical analyses (E) of GW20 human cortical sections with H3K9ac and KI67 antibodies revealed that not all H3K9ac+ cells co-express KI67. Abbreviations: BPs, Basal progenitors; vRGs, ventricular radial glial progenitors; IPs, intermediate progenitors; oRGs, outer sub-ventricular radial glial progenitors; VZ, ventricular zone; SVZ, subventricular zone; iSVZ, inner subventricular zone; oSVZ, outer subventricular zone; IZ, intermediate zone; CP, cortical plate. Scale bars = 50 μm.
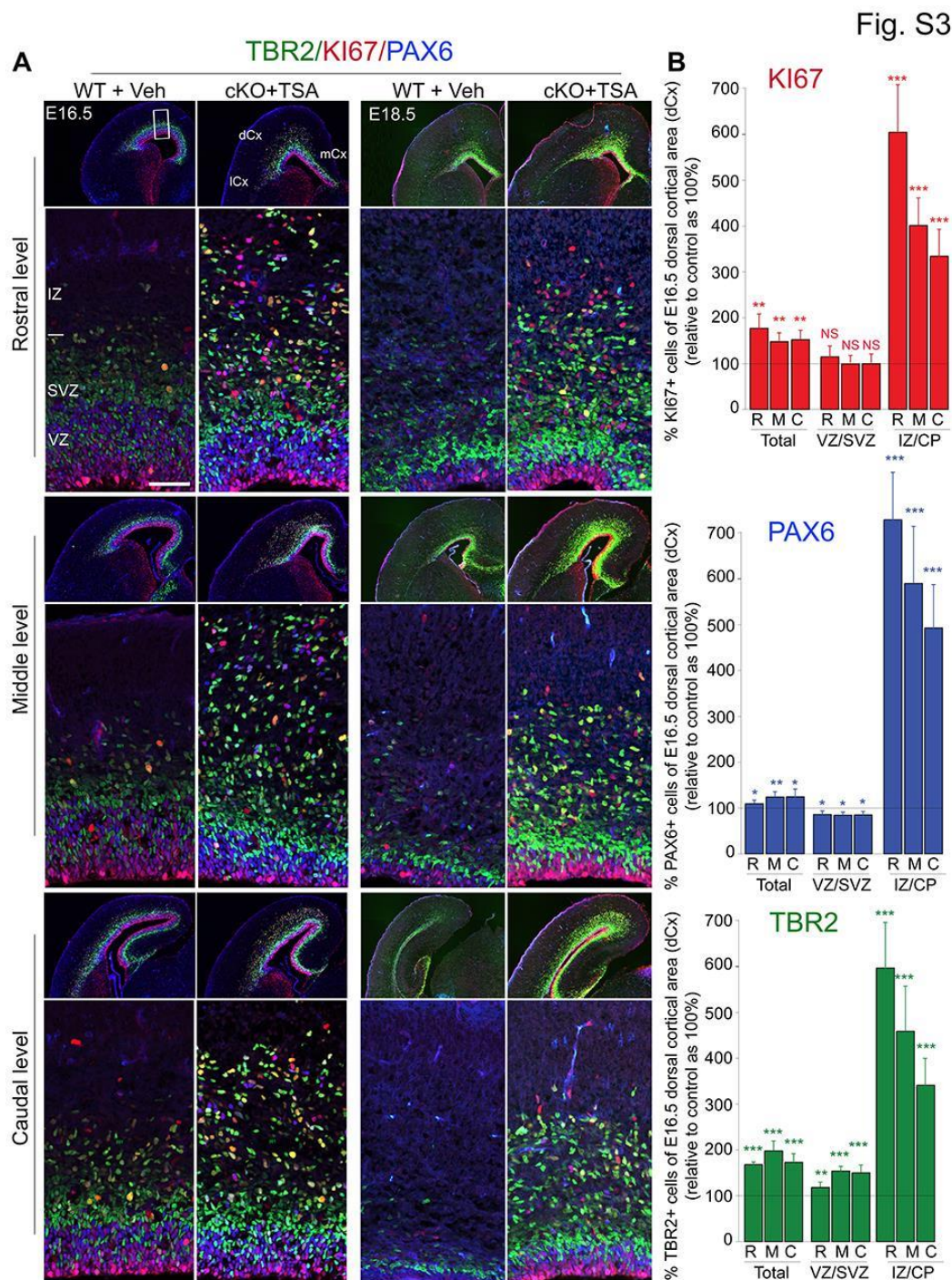
Fig. S3. Area-restricted increases of TBR2+, PAX6+, KI67+ BPs in HDACi− treated cortex

(A) Merged confocal micrographs of fluorescent immunostaining for PAX6 (APs and BPs), TBR2 (BPs), and KI67 (actively-proliferating progenitors in cell cycle) with rostral, middle and caudal levels from unfolded areas of Veh-treated WT and TSA-treated BAF155cKO cortex at stages E16.5 and E18.5. In each level, *upper rows* present coronal mouse brain sections while *lower rows* present confocal images of an area in dorsal

pallium. (B) Statistical analyses of immunostainings shown in A for dorsal cortex (dCx) at E16.5. Values are presented as mean ± SEM (*P <0.05, ***P <0.01, ***P <0.005). Abbreviations: BPs, Basal progenitors; vRGs, ventricular radial glial progenitors; IPs, intermediate progenitors; VZ, ventricular zone; SVZ, subventricular zone; IZ, intermediate zone; CP, cortical plate. Scale bars = 50 μm.
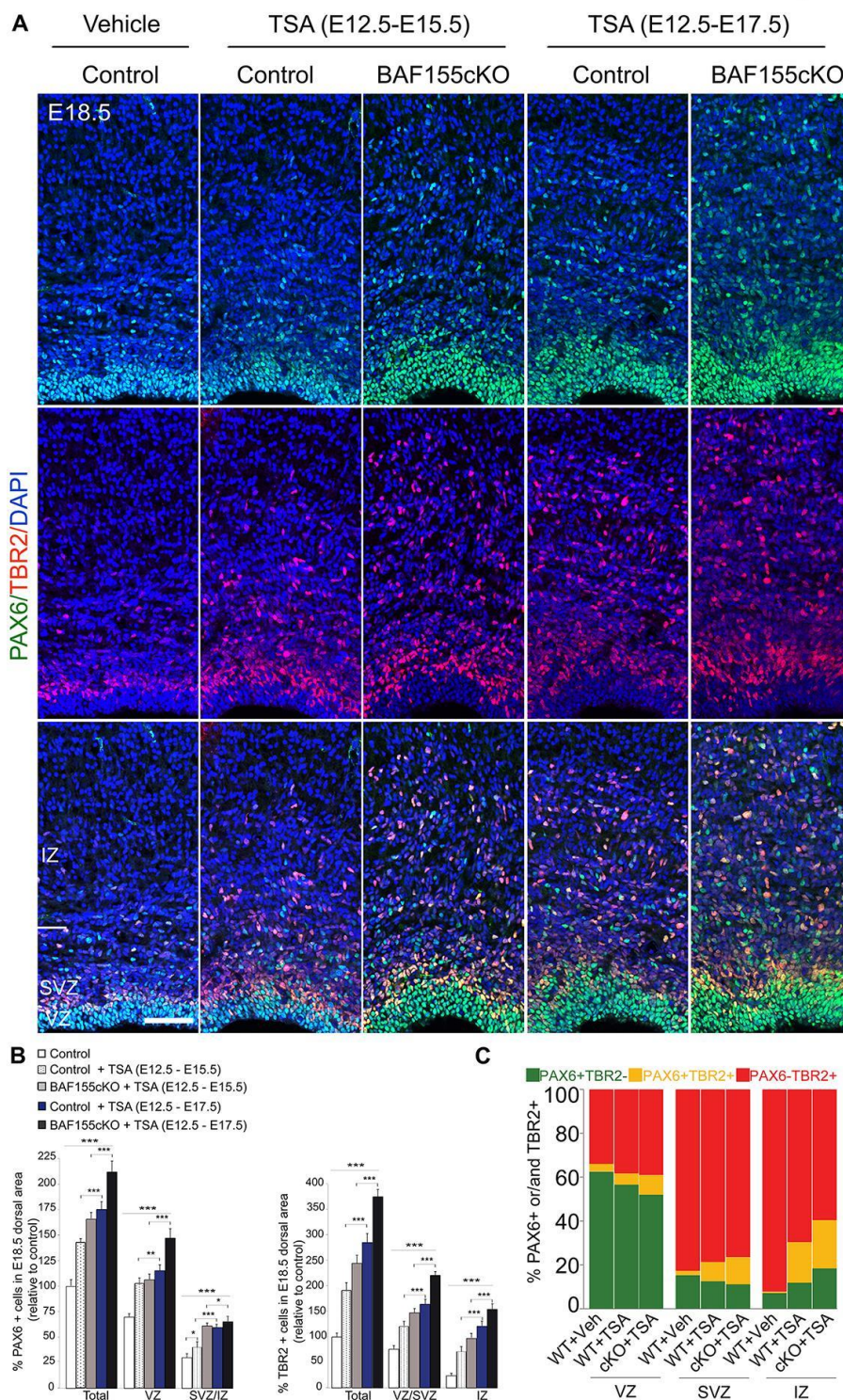
Fig. S4. HDAC inhibition increase the genesis of BPs in dose- dependent fashion.

(A) Immunofluorescence micrographs showing the quantity and distribution of PAX6 and TBR2 at E18.5 in control and BAF155cKO cortices, with or without two regimens of HDACi treatment. (B) Statistical analysis for dorsal cortical area revealed that more PAX6+ BPs

and TBR2+ BPs were found in E18.5 cortex treated with TSA for E13.5-E17.5 than those in those injected with TSA for E13.5-E15.5. Notably, the TSA treatment also led to increased number of PAX6+vRGs at E18.5. (C) Proportions of cortical progenitors expressing PAX6 and/or TBR2 in VZ, SVZ, and IZ from dorsal area of WT+ Veh, WT+ TSA and BAF155cKO+ TSA cortex at E16.5. Values are presented as mean ± SEM (*P <0.05, ***P <0.01, ***P <0.005). Abbreviations: BPs, Basal progenitors; vRGs, ventricular radial glial progenitors; IPs, intermediate progenitors; VZ, ventricular zone; SVZ, subventricular zone; IZ, intermediate zone; CP, cortical plate. Scale bars = 50 μm.
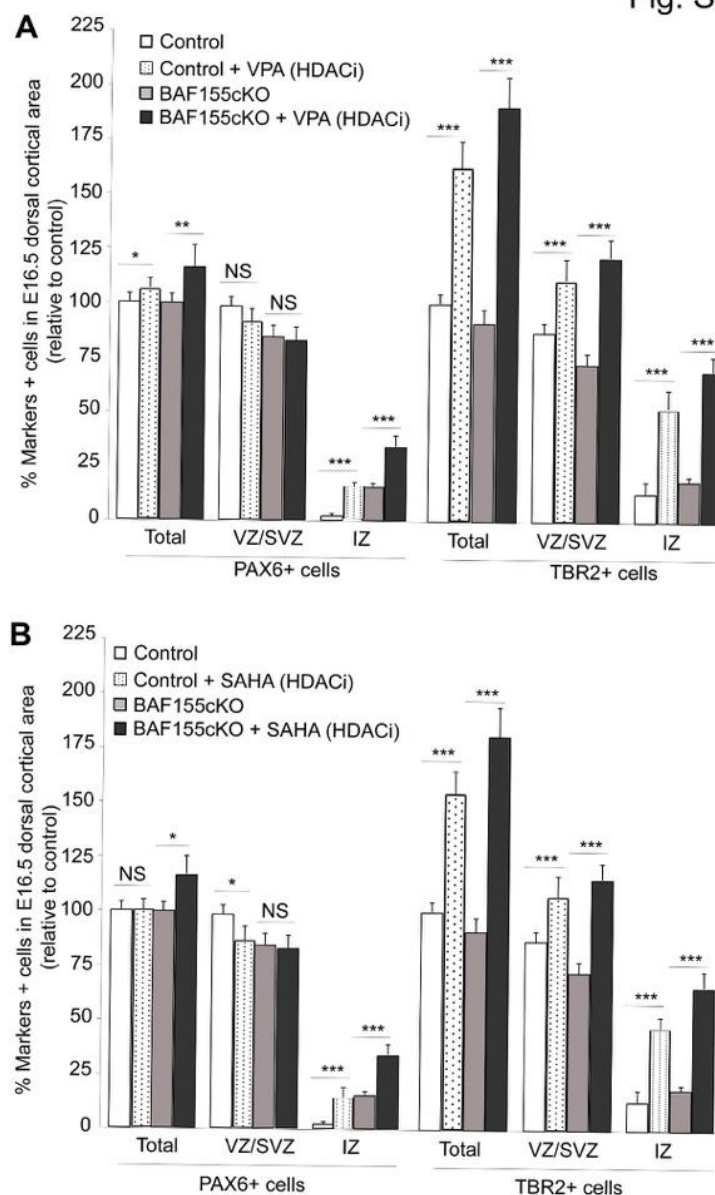
Fig. S5

Fig. S5. Increased level of H3ac promotes genesis and proliferation of BPs in developing mouse cortex.

(F, G) Increasing H3 acetylation by other HDAC inhibitors such as VPA (E) and SAHA (F) also increased the number of TBR2+, PAX6+ BPs in both WT and BAF155cKO developing cortex. Values are presented as mean ± SEM (*P <0.05, ***P <0.01, ***P <0.005). Abbreviations: BPs, Basal progenitors; vRGs, ventricular radial glial progenitors; IPs, intermediate progenitors; VZ, ventricular zone; SVZ, subventricular zone; IZ, intermediate zone; CP, cortical plate. Scale bars = 50 μm.
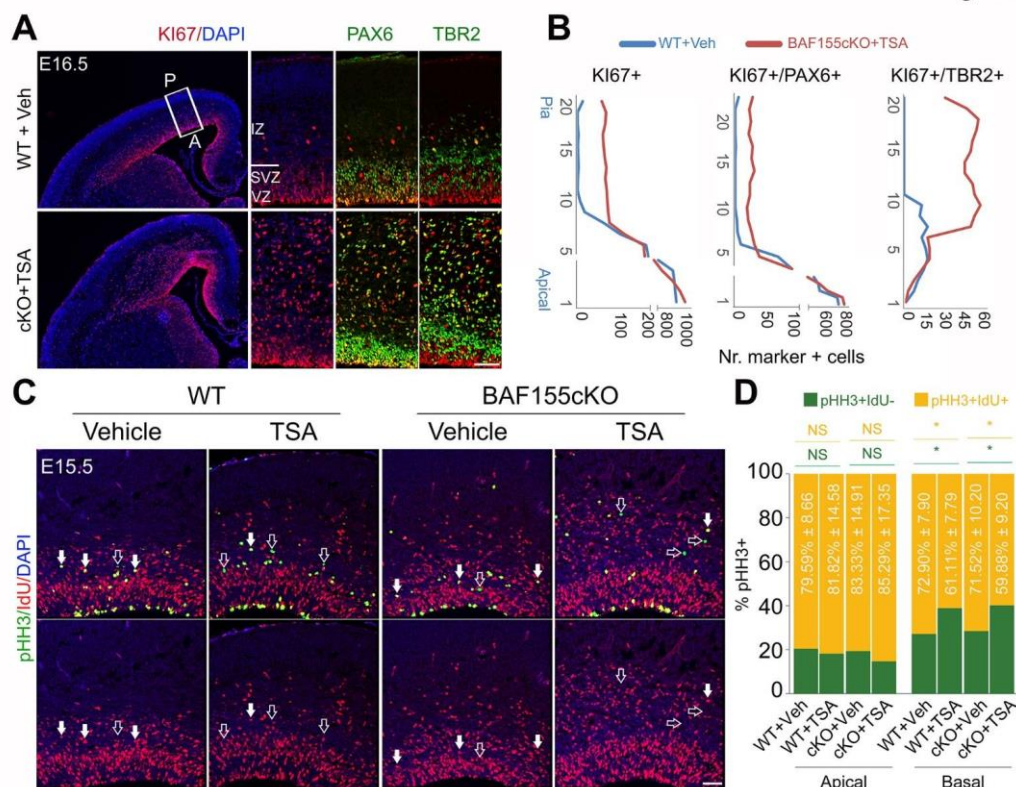
Fig. S6. H3 acetylation specifically promotes proliferation of basal progenitors

(A) IHC images for KI67/PAX6 or KI67/TBR2 staining of Veh-treated WT and TSA-treated BAF155cKO cortex at E16.5. Right panels show higher magnifications of the areas indicated by the white box. (B) Quantified distribution of proliferating KI67+/PAX6+ APs/BPs and KI67+/TBR2+ BPs from the apical surface to the pia in vehicle-treated WT cortex and the corresponding region of the TSA-treated BAF155cKO cortex shown in A. (C, D) 4 hours- IdU pulse-labeling is to trace for the progression of cortical progenitors within the S-M phases. (C) Double IHC analysis with antibodies against IdU and pHH3 (markers for cells in the late G2-M phases) to label APs (apical surface- located IdU+/pHH3+) and BPs (basally located IdU+/pHH3+, filled arrows), which already passed through S- phase and entered into late G+M- phases. (D) Proportion of BPs entered into the late G2-M phases in TSA- treated cortices is smaller compared to that in Veh-treated cortices. Values are presented as mean ± SEM (*P <0.05, ***P <0.01, ***P <0.005). Abbreviations: VZ, ventricular zone; SVZ, subventricular zone; IZ, intermediate zone; CP, cortical plate. Scale bars = 50 μm.
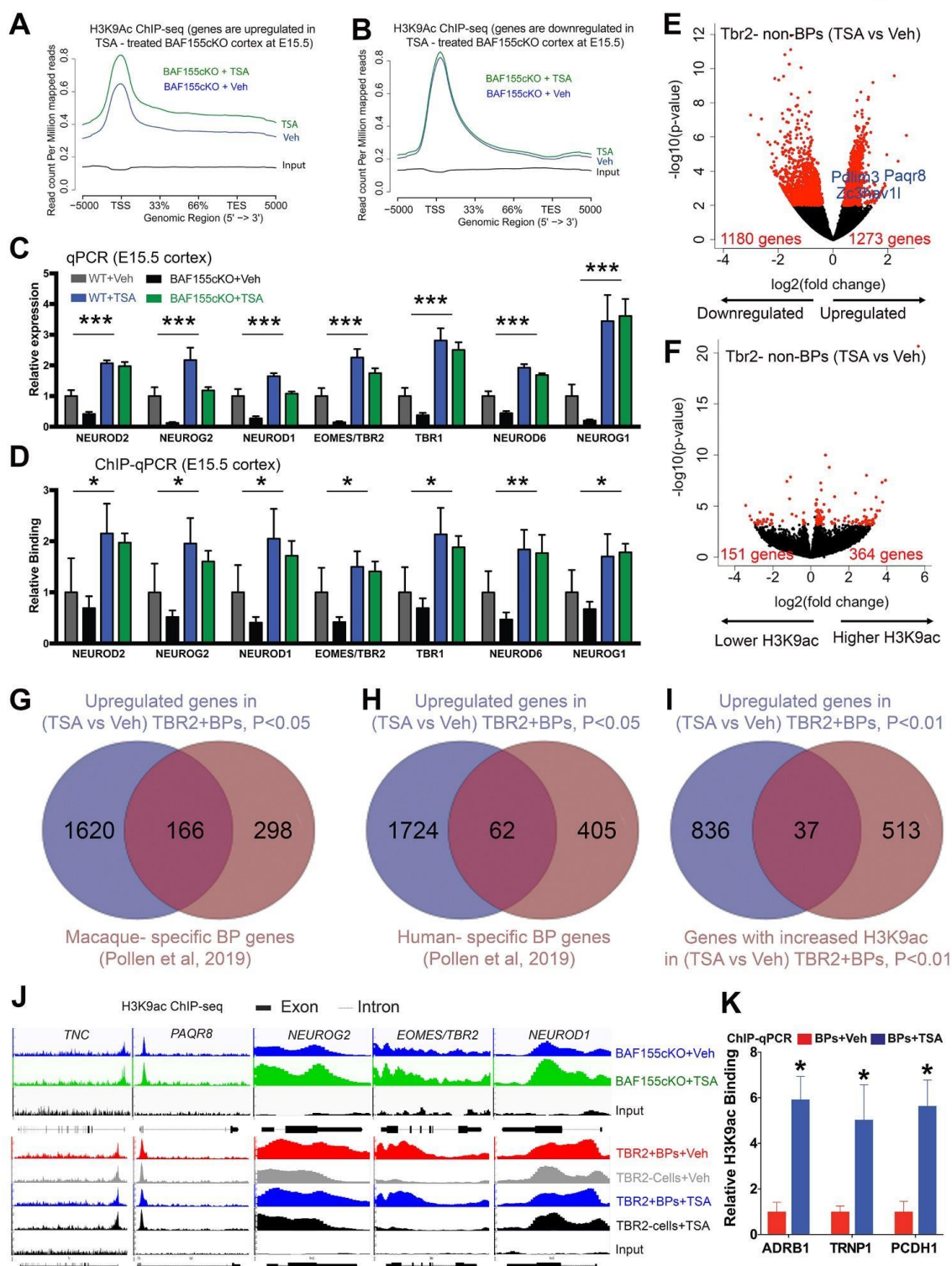
**Fig. S7. HDAC inhibition causes H3K9ac-linked upregulation of gene expression in BAF155cKO developing cortex, TBR2+ BPs and TBR2- cells.**

(A, B, G, H) Volcano plots showing statistically significant changes (Paired Student's *t*-test < 0.01, FC > 1.2) visualized by our RNA-Seq (A, G) and H3K9ac ChIP-Seq (B, H) analyses of E16.5 BAF155cKO cortex (A, B) and TBR2- cells (G, H) in TSA vs. Vehicle

experiments (see also Fig. 2D/E for TBR2+ BPs). (C, D) H3K9ac is increased at loci of upregulated genes (C), but not at those of downregulated genes (D) in TSA treated BAF155cKO cortex. (E, F) Graphs of qPCR to examine gene expression (E) and ChIP-qPCR to quantify H3K9ac level (F). TSA treatment leads to increase in H3K9ac levels at promoters of neurogenesis-related genes (F) and their expression (E) both in WT and in BAF155cKO. ChIP-qPCR and qPCR results for a selection of genes are shown (WT_Veh: n = 3, WT_TSA: n = 4, BAF155 cKO_Veh: n = 4, BAF155 cKO_TSA: n = 4; general TSA effect: mixed linear model p-value < 0.001; general effect of BAF155 knockdown: Two-way ANOVA WT_Veh vs BAF155 cKO_Veh p < 0.01). (G, H) Overlap between the upregulated genes in TSA-treated TBR2+ BP genes and BP/IP genes, which were recently identified specifically for macaque (G), and human (H) (Pollen, Bhaduri et al. 2019). Notably, the TSA treatment provoked expression of a large set of macaque BP genes and in less extend, human BP genes in TBR2+ BPs in the developing mouse cortex. (I) Overlap between up-regulated genes in RNA-seq and genes with increased H3K9ac in ChIP-seq studies of TBR2+ BPs upon the TSA treatment.

(J) Genome browser views of the distribution of H3K9ac along representative BP-enriched genes in BAF155cKO cortex, TBR2+ BPs, TBR2- cells, which were treated with either TSA and or vehicle as control. It should be noted that the bulk level of H3K9ac at their loci was unaltered in TBR2+ BPs and TBR2- cells in response to TSA treatment. In BAF155cKO cortex, the increased level of H3K9ac at these loci is due to expanded pool of BPs upon the treatment of TSA. (K) The increased level of H3K9ac at gene loci of ADRB1, TRNP1, PCDH1 in ChIP-seq upon TSA treatment were validated in ChIP-qPCR.
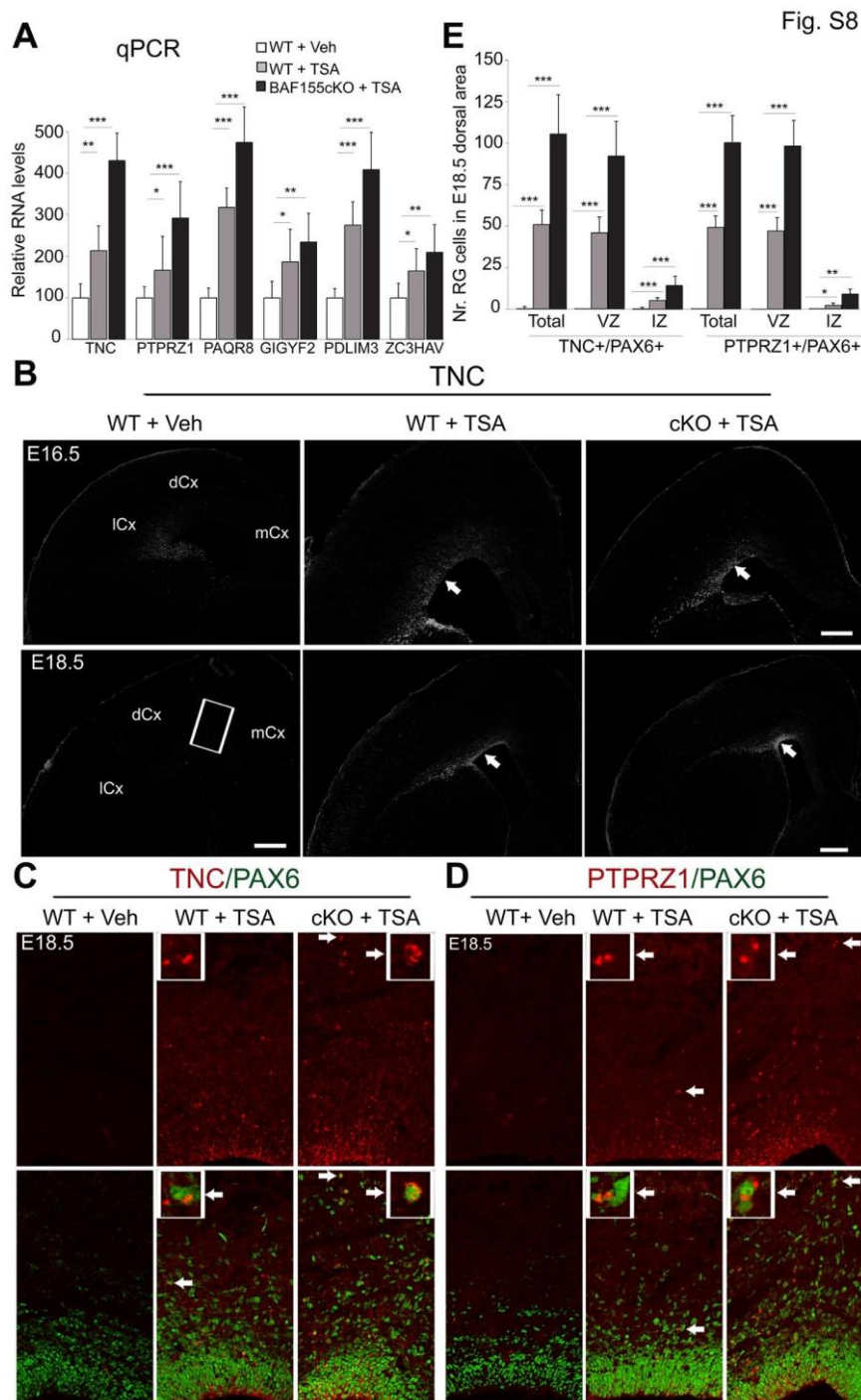
Fig. S8. HDAC inhibition causes upregulation of bRG-enriched genes in developing mouse cortex.

(A) qPCR analysis was used to confirm the upregulated expression of human- enriched bRG markers in TSA-treated cortex as compared to vehicle-treated control. (B-E) Among bRG-enriched genes, expression of TNC and PTPRZ1 have been well characterized in the developing cortex of both human and mouse **(Pollen, Nowakowski et al. 2015)**. TNC and PTPRZ1 antibodies labeled subsets of APs and BPs/bRGs in the entire human cortex,

whereas in mouse their expression was detected only in a sub-population of PAX6+ APs in the lateral cortex (lCx, B) (Pollen, Nowakowski et al. 2015). The increased expression of TNC (B/C) and PTPRZ1 (D) in TSA-treated WT and BAF155cKO cortex at E16.5 and E18.5 was revealed by immunolabeling (B-D) and quantification (E). Notably, the expression of these human bRG markers PTPRZ1 and TNC was found in PAX6+ APs in VZ of dorsal cortex (dCX, white filled arrows in B) and PAX6+ bRGs (white filled arrows in C, D) in TSA-treated cortex. Uppers images show the expression in arrow-pointed cells in lower pictures at higher magnification. Values are presented as mean ± SEM (*P <0.05, ***P <0.01, ***P <0.005). Abbreviations: l/d/mCx, lateral/dorsal/medial cortex. Scale bars = 100 μm.
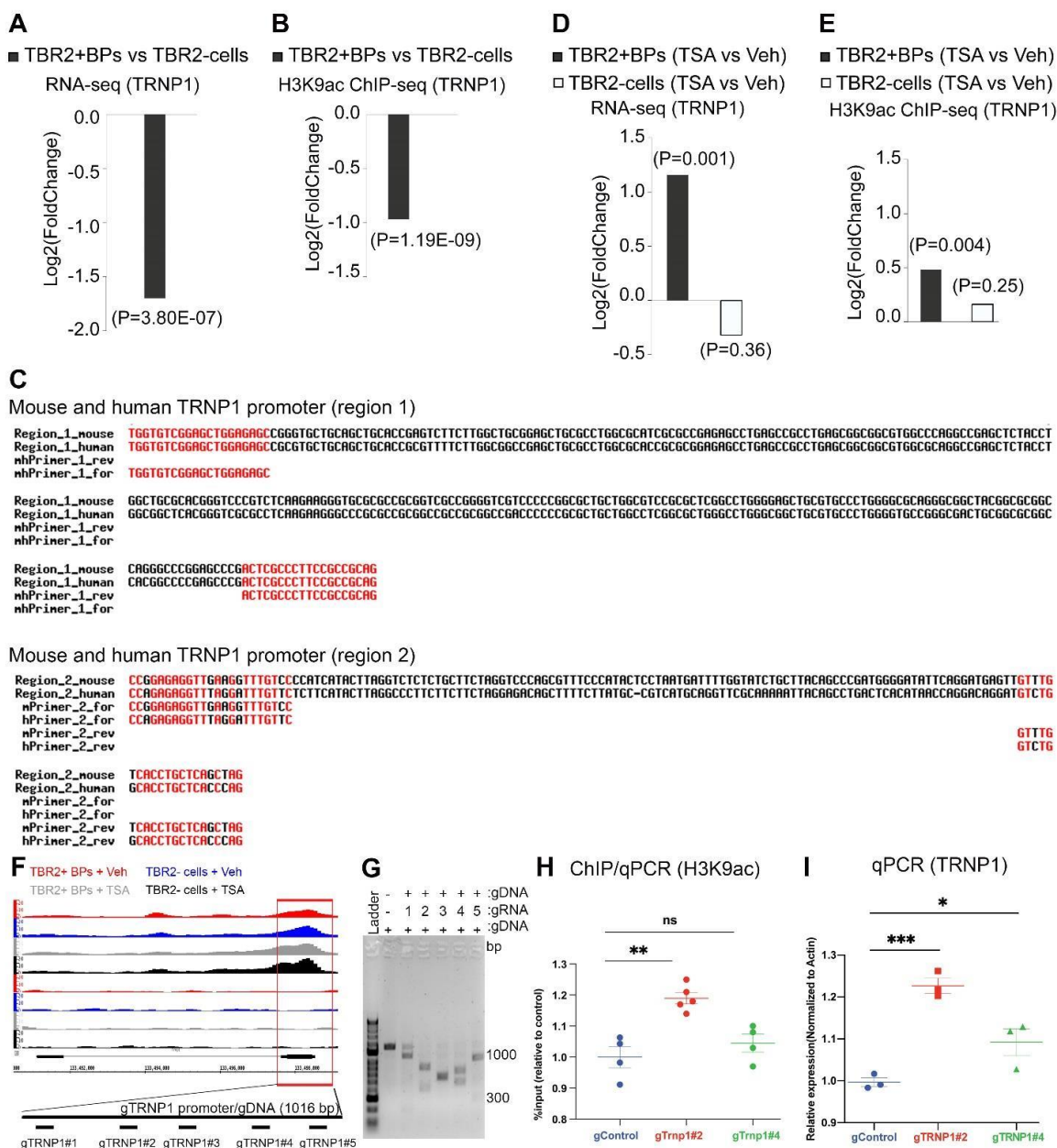
Fig. S9. In vitro testing of TRNP1 gRNAs.

(A/B) TBR2- cells have higher basal levels of H3K9ac at TRNP1 promoter (A) and higher basal expression (B) of this gene compared to TBR2+ BPs. (C) DNA sequence alignment for two ortholog regions of mouse (m) and human (h) TRNP1 promoter and primers, which were used in CHIP/qPCR experiment (see also Fig. 6C). Note that mhPrimer_1 set was used to amplify the region 1 of both mouse and human TRNP1 promoter. The mPrimer_2 and hPrimer_2 sets, which have similar sequence, were used to amplify the region 2. (D/E) TSA treatment significantly increases H3K9ac levels at TRNP1 promoter

(D) and upregulates its expression (E) specifically in TBR2+ BPs but not in TBR2- cells. (F/G) Analysis of the Cas9 cutting efficiency guided by various gRNAs targeting the TRNP1 promoter. (F) Depiction of different gRNAs targeting various regions in TRNP1 promoter that were used for testing. (G) Agarose gel showing the cutting efficiency of each tested gRNA-Cas9 complex on a 1016-bp-long PCR product of TRNP1 promoter region. (H/I) ChIP/qPCR (H) and qPCR (I) analyses indicate that transfection of Neuro2A cells with gTRNP1#2 and #4 increases H3K9ac level at TRNP1 promoter (H) and upregulates its expression (I). Values are presented as means ± SEMs (*P <0.05, ***P <0.01, ***P <0.005).

Fig. S10
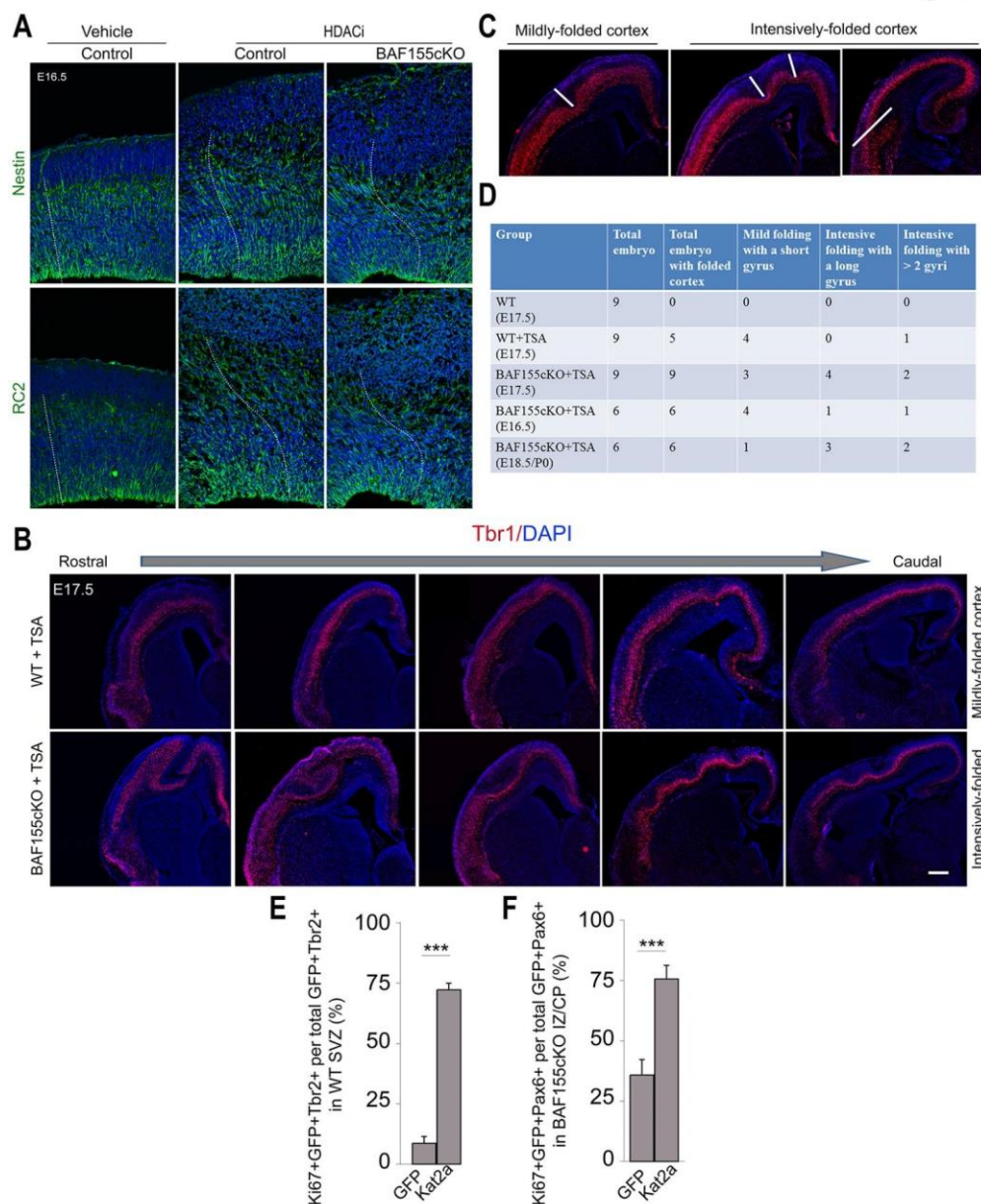
Fig. S10. Augmented level of H3K9ac by HDAC inhibition and overexpression of KAT2A increase basal progenitor proliferation and can induce cortical folding. (A) Staining of RG fibers (NESTIN and RC2) in Veh-treated control, TSA-treated-control and TSA-treated BAF155cKO cortex revealed an increase and divergence of radial fibers at basal sides upon TSA treatment (radial processes indicated by white-dashed lines on confocal images). (B) Cortical tissue was processed for TBR1 immunostaining. Examples with different forms of cortical folding were shown from rostral cortex to caudal cortex at E17.5. The mildly-folded and intensively-folded cortices were usually found in TSA - treated WT embryos and TSA - treated BAF155cKO embryos, respectively. (C) The folding in length was measured from the pial surface to the end of the innervated point outlined

by white lines. Cortical folding phenotypes were scored as a mild folding being shorter than 200µm in length (leftmost image in B) and intensive folding with a gyrus deeper than 200µm from the surface or with more than two gyri (right images). (D) Overview of cortical phenotypes observed after TSA or Veh-treatment. In total, nine Veh-treated control, nine TSA-treated control and nine TSA-treated *BAF155cKO* brains at E17.5, six brains at E16.5 and six brains at E18.5/P0 of TSA-treated BAF155cKO mutants were used to examine folding phenotypes. Six coronal sections per brain covering the rostral, middle-level and caudal cortex were processed for TBR1 IHC. Cortex was scored as "mild folding" or "intensive folding", if at least two sections met the above criteria. (E-F) E13.5 cortices of WT or BAF155cKO embryos were electroporated with eGFP or KAT2A-ires-eGFP, followed by quantification of TBR2+/KI67+ proliferating BPs/IPs (E) and PAX6+/KI67+ proliferating BPs/bRGs (F) at E15.5. Statistical analyses revealed that the overexpression of KAT2A led to increase the proliferation of BPs. Scale bars = 100 μm.

REFERENCES

Albert, M. and W. B. Huttner (2018). "Epigenetic and Transcriptional Pre-patterning-An Emerging Theme in Cortical Neurogenesis." Front Neurosci **12**: 359.

Albert, M., N. Kalebic, M. Florio, N. Lakshmanaperumal, C. Haffner, H. Brandl, I. Henry and W. B. Huttner (2017). "Epigenome profiling and editing of neocortical progenitor cells during development." EMBO J **36**(17): 2642-2658.

Arai, Y., J. N. Pulvers, C. Haffner, B. Schilling, I. Nusslein, F. Calegari and W. B. Huttner (2011). "Neural stem and progenitor cells shorten S-phase on commitment to neuron production." Nature communications **2**: 154.

Bachmann, C., H. Nguyen, J. Rosenbusch, L. Pham, T. Rabe, M. Patwa, G. Sokpor, R. H. Seong, R. Ashery-Padan, A. Mansouri, A. Stoykova, J. F. Staiger and T. Tuoc (2016). "mSWI/SNF (BAF) Complexes Are Indispensable for the Neurogenesis and Development of Embryonic Olfactory Epithelium." PLoS genetics **12**(9): e1006274.

Bae, B. I., I. Tietjen, K. D. Atabay, G. D. Evrony, M. B. Johnson, E. Asare, P. P. Wang, A. Y. Murayama, K. Im, S. N. Lisgo, L. Overman, N. Sestan, B. S. Chang, A. J. Barkovich, P. E. Grant, M. Topcu, J. Politsky, H. Okano, X. Piao and C. A. Walsh (2014). "Evolutionarily dynamic alternative splicing of GPR56 regulates regional cerebral cortical patterning." Science **343**(6172): 764-768.

Bishop, K. M., S. Garel, Y. Nakagawa, J. L. Rubenstein and D. D. O'Leary (2003). "Emx1 and Emx2 cooperate to regulate cortical size, lamination, neuronal differentiation, development of cortical efferents, and thalamocortical pathfinding." The Journal of comparative neurology **457**(4): 345-360.

Bishop, K. M., J. L. Rubenstein and D. D. O'Leary (2002). "Distinct actions of Emx1, Emx2, and Pax6 in regulating the specification of areas in the developing neocortex." The Journal of neuroscience : the official journal of the Society for Neuroscience **22**(17): 7627-7638.

Borrell, V. and M. Gotz (2014). "Role of radial glial cells in cerebral cortex folding." Current Opinion in Neurobiology **27**: 39-46.

Chen, B., L. A. Gilbert, B. A. Cimini, J. Schnitzbauer, W. Zhang, G.-W. Li, J. Park, E. H. Blackburn, J. S. Weissman, L. S. Qi and B. Huang (2013). "Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system." Cell **155**(7): 1479-1491.

Chen, J., E. E. Bardes, B. J. Aronow and A. G. Jegga (2009). "ToppGene Suite for gene list enrichment analysis and candidate gene prioritization." Nucleic Acids Res **37**(Web Server issue): W305-311.

Chenn, A. and C. A. Walsh (2002). "Regulation of cerebral cortical size by control of cell cycle exit in neural precursors." Science **297**(5580): 365-369.

Choi, J., M. Ko, S. Jeon, Y. Jeon, K. Park, C. Lee, H. Lee and R. H. Seong (2012). "The SWI/SNF-like BAF complex is essential for early B cell development." Journal of immunology **188**(8): 3791-3803.

Conti, L., S. M. Pollard, T. Gorba, E. Reitano, M. Toselli, G. Biella, Y. Sun, S. Sanzone, Q. L. Ying, E. Cattaneo and A. Smith (2005). "Niche-independent symmetrical self-renewal of a mammalian tissue stem cell." PLoS Biol **3**(9): e283.

de Juan Romero, C., C. Bruder, U. Tomasello, J. M. Sanz-Anquela and V. Borrell (2015). "Discrete domains of gene expression in germinal layers distinguish the development of gyrencephaly." The EMBO journal **34**(14): 1859-1874.

Dehay, C., H. Kennedy and K. S. Kosik (2015). "The outer subventricular zone and primate-specific cortical complexification." Neuron **85**(4): 683-694.

Del Toro, D., T. Ruff, E. Cederfjall, A. Villalba, G. Seyit-Bremer, V. Borrell and R. Klein (2017). "Regulation of Cerebral Cortex Folding by Controlling Neuronal Migration via FLRT Adhesion Molecules." Cell **169**(4): 621-635 e616.

Djebali, S., C. A. Davis, A. Merkel, A. Dobin, T. Lassmann, A. Mortazavi, A. Tanzer, J. Lagarde, W. Lin, F. Schlesinger, C. Xue, G. K. Marinov, J. Khatun, B. A. Williams, C. Zaleski, J. Rozowsky, M. Roder, F. Kokocinski, R. F. Abdelhamid, T. Alioto, I. Antoshechkin, M. T. Baer, N. S. Bar, P. Batut, K. Bell, I. Bell, S. Chakrabortty, X. Chen, J. Chrast, J. Curado, T. Derrien, J. Drenkow, E. Dumais, J. Dumais, R. Duttagupta, E. Falconnet, M. Fastuca, K. Fejes-Toth, P. Ferreira, S. Foissac, M. J. Fullwood, H. Gao, D. Gonzalez, A. Gordon, H. Gunawardena, C. Howald, S. Jha, R. Johnson, P. Kapranov, B. King, C. Kingswood, O. J. Luo, E. Park, K. Persaud, J. B. Preall, P. Ribeca, B. Risk, D. Robyr, M. Sammeth, L. Schaffer, L. H. See, A. Shahab, J. Skancke, A. M. Suzuki, H. Takahashi, H. Tilgner, D. Trout, N. Walters, H. Wang, J. Wrobel, Y. Yu, X. Ruan, Y. Hayashizaki, J. Harrow, M. Gerstein, T. Hubbard, A. Reymond, S. E. Antonarakis, G. Hannon, M. C. Giddings, Y. Ruan, B. Wold, P. Carninci, R. Guigo and T. R. Gingeras (2012). "Landscape of transcription in human cells." Nature **489**(7414): 101-108.

Englund, C., A. Fink, C. Lau, D. Pham, R. A. Daza, A. Bulfone, T. Kowalczyk and R. F. Hevner (2005). "Pax6, Tbr2, and Tbr1 are expressed sequentially by radial glia, intermediate progenitor cells, and postmitotic neurons in developing neocortex." J Neurosci **25**(1): 247-251.

Feng, J., T. Liu and Y. Zhang (2011). "Using MACS to identify peaks from ChIP-Seq data." Curr Protoc Bioinformatics **Chapter 2**: Unit 2 14.

Fiddes, I. T., G. A. Lodewijk, M. Mooring, C. M. Bosworth, A. D. Ewing, G. L. Mantalas, A. M. Novak, A. van den Bout, A. Bishara, J. L. Rosenkrantz, R. Lorig-Roach, A. R. Field, M. Haeussler, L. Russo, A. Bhaduri, T. J. Nowakowski, A. A. Pollen, M. L. Dougherty, X. Nuttle, M. C. Addor, S. Zwolinski, S. Katzman, A. Kriegstein, E. E. Eichler, S. R. Salama, F. M. J. Jacobs and D. Haussler (2018). "Human-Specific NOTCH2NL Genes Affect Notch Signaling and Cortical Neurogenesis." Cell **173**(6): 1356-1369 e1322.

Fietz, S. A., I. Kelava, J. Vogt, M. Wilsch-Brauninger, D. Stenzel, J. L. Fish, D. Corbeil, A. Riehn, W. Distler, R. Nitsch and W. B. Huttner (2010). "OSVZ progenitors of human and ferret neocortex are epithelial-like and expand by integrin signaling." Nature neuroscience **13**(6): 690-699.

Fietz, S. A., R. Lachmann, H. Brandl, M. Kircher, N. Samusik, R. Schroder, N. Lakshmanaperumal, I. Henry, J. Vogt, A. Riehn, W. Distler, R. Nitsch, W. Enard, S. Paabo

and W. B. Huttner (2012). "Transcriptomes of germinal zones of human and mouse fetal neocortex suggest a role of extracellular matrix in progenitor self-renewal." Proceedings of the National Academy of Sciences of the United States of America **109**(29): 11836-11841.

Florio, M., M. Albert, E. Taverna, T. Namba, H. Brandl, E. Lewitus, C. Haffner, A. Sykes, F. K. Wong, J. Peters, E. Guhr, S. Klemroth, K. Prufer, J. Kelso, R. Naumann, I. Nusslein, A. Dahl, R. Lachmann, S. Paabo and W. B. Huttner (2015). "Human-specific gene ARHGAP11B promotes basal progenitor amplification and neocortex expansion." Science **347**(6229): 1465-1470.

Garcia-Moreno, F., N. A. Vasistha, N. Trevia, J. A. Bourne and Z. Molnar (2012). "Compartmentalization of cerebral cortical germinal zones in a lissencephalic primate and gyrencephalic rodent." Cerebral cortex **22**(2): 482-492.

Gorski, J. A., T. Talley, M. Qiu, L. Puelles, J. L. Rubenstein and K. R. Jones (2002). "Cortical excitatory neurons and glia, but not GABAergic neurons, are produced in the Emx1-expressing lineage." J Neurosci **22**(15): 6309-6314.

Halder, R., M. Hennion, R. O. Vidal, O. Shomroni, R. U. Rahman, A. Rajput, T. P. Centeno, F. van Bebber, V. Capece, J. C. Garcia Vizcaino, A. L. Schuetz, S. Burkhardt, E. Benito, M. Navarro Sala, S. B. Javan, C. Haass, B. Schmid, A. Fischer and S. Bonn (2016). "DNA methylation changes in plasticity genes accompany the formation and maintenance of memory." Nat Neurosci **19**(1): 102-110.

Hansen, D. V., J. H. Lui, P. R. Parker and A. R. Kriegstein (2010). "Neurogenic radial glia in the outer subventricular zone of human neocortex." Nature **464**(7288): 554-561.

Heinz, S., C. Benner, N. Spann, E. Bertolino, Y. C. Lin, P. Laslo, J. X. Cheng, C. Murre, H. Singh and C. K. Glass (2010). "Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities." Mol Cell **38**(4): 576-589.

Johnson, M. B., P. P. Wang, K. D. Atabay, E. A. Murphy, R. N. Doan, J. L. Hecht and C. A. Walsh (2015). "Single-cell analysis reveals transcriptional heterogeneity of neural progenitors in human cortex." Nature Neuroscience **18**(5): 637-646.

Ju, X. C., Q. Q. Hou, A. S. Sheng, K. Y. Wu, Y. Zhou, Y. Jin, T. Wen, Z. Yang, X. Wang and Z. G. Luo (2016). "The hominoid-specific gene TBC1D3 promotes generation of basal neural progenitors and induces cortical folding in mice." eLife **5**.

Kalebic, N., E. Taverna, S. Tavano, F. K. Wong, D. Suchold, S. Winkler, W. B. Huttner and M. Sarov (2016). "CRISPR/Cas9-induced disruption of gene expression in mouse embryonic brain and single neural stem cells in vivo." EMBO reports **17**(3): 338-348.

Kelava, I., I. Reillo, A. Y. Murayama, A. T. Kalinka, D. Stenzel, P. Tomancak, F. Matsuzaki, C. Lebrand, E. Sasaki, J. C. Schwamborn, H. Okano, W. B. Huttner and V. Borrell (2012). "Abundant occurrence of basal radial glia in the subventricular zone of embryonic neocortex of a lissencephalic primate, the common marmoset Callithrix jacchus." Cerebral cortex **22**(2): 469-481.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin and S. Genome Project Data Processing (2009). "The Sequence Alignment/Map format and SAMtools." Bioinformatics **25**(16): 2078-2079.

Lienhard, M., C. Grimm, M. Morkel, R. Herwig and L. Chavez (2014). "MEDIPS: genome-wide differential coverage analysis of sequencing data derived from DNA enrichment experiments." Bioinformatics **30**(2): 284-286.

Liu, J., W. Liu, L. Yang, Q. Wu, H. Zhang, A. Fang, L. Li, X. Xu, L. Sun, J. Zhang, F. Tang and X. Wang (2017). "The Primate-Specific Gene TMEM14B Marks Outer Radial Glia Cells and Promotes Cortical Expansion and Folding." Cell Stem Cell **21**(5): 635-649 e638.

Love, M. I., W. Huber and S. Anders (2014). "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2." Genome biology **15**(12): 1.

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

MacLean, B., D. M. Tomazela, N. Shulman, M. Chambers, G. L. Finney, B. Frewen, R. Kern, D. L. Tabb, D. C. Liebler and M. J. MacCoss (2010). "Skyline: an open source document editor for creating and analyzing targeted proteomics experiments." Bioinformatics **26**(7): 966-968.

Martínez-Balbás, M. A., U. M. Bauer, S. J. Nielsen, A. Brehm and T. Kouzarides (2000). "Regulation of E2F1 activity by acetylation." The EMBO journal **19**(4): 662-671.

Martinez-Martinez, M. A., C. De Juan Romero, V. Fernandez, A. Cardenas, M. Gotz and V. Borrell (2016). "A restricted period for formation of outer subventricular zone defined by Cdh1 and Trnp1 levels." Nature communications **7**: 11812.

Narayanan, R., L. Pham, C. Kerimoglu, T. Watanabe, R. Castro Hernandez, G. Sokpor, P. A. Ulmke, K. A. Kiszka, A. B. Tonchev, J. Rosenbusch, R. H. Seong, U. Teichmann, J. Frahm, A. Fischer, S. Bonn, A. Stoykova, J. F. Staiger and T. Tuoc (2018). "Chromatin Remodeling BAF155 Subunit Regulates the Genesis of Basal Progenitors in Developing Cortex." iScience **4**: 109-126.

Narayanan, R., M. Pirouz, C. Kerimoglu, L. Pham, R. J. Wagener, K. A. Kiszka, J. Rosenbusch, R. H. Seong, M. Kessel, A. Fischer, A. Stoykova, J. F. Staiger and T. Tuoc (2015). "Loss of BAF (mSWI/SNF) Complexes Causes Global Transcriptional and Chromatin State Changes in Forebrain Development." Cell reports **13**(9): 1842-1854.

Nicol, J. W., G. A. Helt, S. G. Blanchard, Jr., A. Raja and A. E. Loraine (2009). "The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets." Bioinformatics **25**(20): 2730-2731.

Nonaka-Kinoshita, M., I. Reillo, B. Artegiani, M. A. Martinez-Martinez, M. Nelson, V. Borrell and F. Calegari (2013). "Regulation of cerebral cortex size and folding by expansion of basal progenitors." The EMBO journal **32**(13): 1817-1828.

Pilz, G. A., A. Shitamukai, I. Reillo, E. Pacary, J. Schwausch, R. Stahl, J. Ninkovic, H. J. Snippert, H. Clevers, L. Godinho, F. Guillemot, V. Borrell, F. Matsuzaki and M. Gotz (2013).

"Amplification of progenitors in the mammalian telencephalon includes a new radial glial cell type." <u>Nature communications</u> **4**: 2125.

Pollen, A. A., A. Bhaduri, M. G. Andrews, T. J. Nowakowski, O. S. Meyerson, M. A. Mostajo-Radji, E. Di Lullo, B. Alvarado, M. Bedolli, M. L. Dougherty, I. T. Fiddes, Z. N. Kronenberg, J. Shuga, A. A. Leyrat, J. A. West, M. Bershteyn, C. B. Lowe, B. J. Pavlovic, S. R. Salama, D. Haussler, E. E. Eichler and A. R. Kriegstein (2019). "Establishing Cerebral Organoids as Models of Human-Specific Brain Evolution." <u>Cell</u> **176**(4): 743-756 e717.

Pollen, A. A., T. J. Nowakowski, J. Chen, H. Retallack, C. Sandoval-Espinosa, C. R. Nicholas, J. Shuga, S. J. Liu, M. C. Oldham, A. Diaz, D. A. Lim, A. A. Leyrat, J. A. West and A. R. Kriegstein (2015). "Molecular identity of human outer radial glia during cortical development." <u>Cell</u> **163**(1): 55-67.

Reillo, I., C. de Juan Romero, M. A. Garcia-Cabezas and V. Borrell (2011). "A role for intermediate radial glia in the tangential expansion of the mammalian cerebral cortex." <u>Cerebral cortex</u> **21**(7): 1674-1694.

Ross-Innes, C. S., R. Stark, A. E. Teschendorff, K. A. Holmes, H. R. Ali, M. J. Dunning, G. D. Brown, O. Gojis, I. O. Ellis, A. R. Green, S. Ali, S. F. Chin, C. Palmieri, C. Caldas and J. S. Carroll (2012). "Differential oestrogen receptor binding is associated with clinical outcome in breast cancer." <u>Nature</u> **481**(7381): 389-393.

Sahara, S. and D. D. O'Leary (2009). "Fgf10 regulates transition period of cortical stem cell differentiation to radial glia controlling generation of neurons and basal progenitors." <u>Neuron</u> **63**(1): 48-62.

Shen, L., N. Shao, X. Liu and E. Nestler (2014). "ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases." <u>BMC Genomics</u> **15**: 284.

Shitamukai, A., D. Konno and F. Matsuzaki (2011). "Oblique radial glial divisions in the developing mouse neocortex induce self-renewing progenitors outside the germinal zone that resemble primate outer subventricular zone progenitors." <u>The Journal of neuroscience : the official journal of the Society for Neuroscience</u> **31**(10): 3683-3695.

Stahl, R., T. Walcher, C. De Juan Romero, G. A. Pilz, S. Cappello, M. Irmler, J. M. Sanz-Aquela, J. Beckers, R. Blum, V. Borrell and M. Gotz (2013). "Trnp1 regulates expansion and folding of the mammalian cerebral cortex by control of radial glial fate." <u>Cell</u> **153**(3): 535-549.

Stewart, S. A., D. M. Dykxhoorn, D. Palliser, H. Mizuno, E. Y. Yu, D. S. An, D. M. Sabatini, I. S. Chen, W. C. Hahn, P. A. Sharp, R. A. Weinberg and C. D. Novina (2003). "Lentivirus-delivered stable gene silencing by RNAi in primary cells." <u>RNA</u> **9**(4): 493-501.

Suzuki, I. K., D. Gacquer, R. Van Heurck, D. Kumar, M. Wojno, A. Bilheu, A. Herpoel, N. Lambert, J. Cheron, F. Polleux, V. Detours and P. Vanderhaeghen (2018). "Human-Specific NOTCH2NL Genes Expand Cortical Neurogenesis through Delta/Notch Regulation." <u>Cell</u> **173**(6): 1370-+.

Taverna, E., M. Gotz and W. B. Huttner (2014). "The cell biology of neurogenesis: toward an understanding of the development and evolution of the neocortex." Annual review of cell and developmental biology **30**: 465-502.

Thomsen, E. R., J. K. Mich, Z. Z. Yao, R. D. Hodge, A. M. Doyle, S. M. Jang, S. I. Shehata, A. M. Nelson, N. V. Shapovalova, B. P. Levi and S. Ramanathan (2016). "Fixed single-cell transcriptomic characterization of human radial glial diversity." Nat Methods **13**(1): 87-+.

Toda, T., Y. Shinmyo, T. A. Dinh Duong, K. Masuda and H. Kawasaki (2016). "An essential role of SVZ progenitors in cortical folding in gyrencephalic mammals." Sci Rep **6**: 29578.

Tuoc, T. C., S. Boretius, S. N. Sansom, M. E. Pitulescu, J. Frahm, F. J. Livesey and A. Stoykova (2013). "Chromatin Regulation by BAF170 Controls Cerebral Cortical Size and Thickness." Developmental Cell **25**(3): 256-269.

Tuoc, T. C., K. Radyushkin, A. B. Tonchev, M. C. Pinon, R. Ashery-Padan, Z. Molnar, M. S. Davidoff and A. Stoykova (2009). "Selective cortical layering abnormalities and behavioral deficits in cortex-specific Pax6 knock-out mice." The Journal of neuroscience : the official journal of the Society for Neuroscience **29**(26): 8335-8349.

Tuoc, T. C., K. Radyushkin, A. B. Tonchev, M. C. Pinon, R. Ashery-Padan, Z. Molnar, M. S. Davidoff and A. Stoykova (2009). "Selective cortical layering abnormalities and behavioral deficits in cortex-specific Pax6 knock-out mice." Journal of neuroscience **29**(26): 8335-8349.

Tuoc, T. C. and A. Stoykova (2008). "Trim11 modulates the function of neurogenic transcription factor Pax6 through ubiquitin-proteosome system." Genes Dev **22**(14): 1972-1986.

Tuoc, T. C. and A. Stoykova (2008). "Trim11 modulates the function of neurogenic transcription factor Pax6 through ubiquitin-proteosome system." Genes & development **22**(14): 1972-1986.

Vaid, S., J. G. Camp, L. Hersemann, C. Eugster Oegema, A. K. Heninger, S. Winkler, H. Brandl, M. Sarov, B. Treutlein, W. B. Huttner and T. Namba (2018). "A novel population of Hopx-dependent basal radial glial cells in the developing mouse neocortex." Development **145**(20).

Wang, L., S. Hou and Y. G. Han (2016). "Hedgehog signaling promotes basal progenitor expansion and the growth and folding of the neocortex." Nature Neuroscience **19**(7): 888-896.

Wang, X., J. W. Tsai, B. LaMonica and A. R. Kriegstein (2011). "A new subtype of progenitor cell in the mouse embryonic neocortex." Nat Neurosci **14**(5): 555-561.

Xie, Y., R. Castro-Hernandez, G. Sokpor, L. Pham, R. Narayanan, J. Rosenbusch, J. F. Staiger and T. Tuoc (2019). "RBM15 Modulates the Function of Chromatin Remodeling Factor BAF155 Through RNA Methylation in Developing Cortex." Mol Neurobiol.

Zheng, Y., P. M. Thomas and N. L. Kelleher (2013). "Measurement of acetylation turnover at distinct lysines in human histones identifies long-lived acetylation sites." <u>Nat Commun</u> **4**: 2203.

ACKNOWLEDGMENTS

AUTHOR CONTRIBUTIONS

CK performed ChIP, RNA-Seq, ChIP-Seq analyses and the corresponding bioinformatics analyses; LP and TT performed most histological analyses of IUE and HDACi experiments; AT, LP, RM, EK, MA and TT collected human fetuses and performed analyses of expression patterns of mouse and human cortices; MSS contributed to ChIP-Seq library preparation, establishment of TBR2+ cell sorting, and RNA-Seq. LK contributed to RNAseq library preparation. YX, GS, HN, JR, PAU and EA contributed in histological analyses, epigenome editing, and statistical quantification. AM contributed to qPCR and ChIP-qPCR. VC performed QC analysis and alignment of ChIP-Seq and RNA-Seq. MA and WH contributed in epigenome editing. ME and MG contributed in TRNP1 study. AS and JS provided transgenic lines and research tools; JFS, AS, WH, MG and AF offered suggestions for the study; TT conceived the study and wrote the draft. CK, GS, JFS, MA, ME, AT and AS revised the manuscript.

COMPETING INTERESTS: The authors declare no competing financial interests.

DATA AVAILABILITY

All RNA-Seq and ChIP-Seq data have been deposited in GEO and will be released to public upon acceptance of the manuscript.

Manuscript 2

in vivo cKO of Setd1b gene in neurons and genome wide analysis

4. Manuscript 2

SETD1B controls cognitive function via cell type specific regulation of neuronal identity genes.

Alexandra Michurina[1*], **M Sadman Sakib**[1*], Cemil Kerimoglu[1*$], Dennis Manfred Krüger[1],

Lalit Kaurani[1], Rezaul Islam[1], Tonatiuh Pena Centeno[1], Julia Cha[1], Xingbo Xu[2], Elisabeth

M. Zeisberg[2,3,4] Andrea Kranz[5], Francis Adrian Stewart[5,6], Andre Fischer[1,4,6,7, $]

My contributions:
Shared first author for major contribution to the experimental part.

- Established Neuronal (NeuN+) and non-neuronal (NeuN-) nuclear RNAseq from frozen mouse hippocampal tissue.
- Established and performed ChIP-seq for 4 different histone marks (H3K4me1, H3K4me3, H3K27ac, H3K9ac) from low input neuronal chromatin (~60,000 cells) from WT vs Setd1B cKO mice hippocampal tissues.
- Generated bulk neuronal nuclear data from WT vs Setd1B neuronal nuclei
- Utilized isolated nuclei to perform single nuclei RNA sequencing using 10X genomics platform.
- Contributed to data analysis.

Manuscript publication status:

Ongoing

Submitted to *"EMBO Journal"*: 7 August 2020

Reviewers' comments received: 15 September 2020

Responses to Reviewers being prepared

Pre-print available on Biorxiv (Since 7 August 2020):

https://www.biorxiv.org/content/10.1101/2020.08.07.240853v1.full

SETD1B controls cognitive function via cell type specific regulation of neuronal identity genes.

Alexandra Michurina[1]*, **M Sadman Sakib**[1]*, Cemil Kerimoglu[1]*$, Dennis Manfred Krüger[1], Lalit Kaurani[1], Rezaul Islam[1], Tonatiuh Pena Centeno[1], Julia Cha[1], Xingbo Xu[2], Elisabeth M. Zeisberg[2,3,4] Andrea Kranz[5], Francis Adrian Stewart[5,6], Andre Fischer[1,4,6,7, $]

[1]Department for Systems Medicine and Epigenetics, German Center for Neurodegenerative Diseases (DZNE), Von Siebold Str. 3a, 37075, Göttingen, Germany

[2] Department of Cardiology and Pneumology, University Medical Center of Göttingen, Georg-August University, Germany, Robert Koch Str. 40, 37075 Göttingen, Germany

[3] German Centre for Cardiovascular Research (DZHK), Partner Site Göttingen, Germany

[4]Cluster of Excellence "Multiscale Bioimaging: from Molecular Machines to Networks of Excitable Cells" (MBExC), University of Göttingen, Germany

[5]Biotechnology Center, Center for Molecular and Cellular Bioengineering, Dresden University of Technology, Dresden 01307, Germany

[6] Max-Planck-Institute for Cell Biology and Genetics, Dresden, 01307, Germany

[7]Department of Psychiatry and Psychotherapy, University Medical Center Göttingen, Von Siebold Str 5, 37075 Göttingen, Germany

§Corresponding authors: cemil.kerimoglu@dzne.de; andre.fischer@dzne.de

* equal contribution

ABSTRACT

Histone-3-lysine-4-methylation (H3K4me) is mediated by six different lysine methyltransferases (KMTs). Amongst these enzymes SET domain containing 1b (SETD1B) has been linked to intellectual disability but its role in the adult brain has not been studied yet. Here we show that mice lacking *Setd1b* from excitatory neurons of the adult forebrain exhibit severe memory impairment. By combining neuron-specific ChIP-seq, RNA-seq and single cell RNA-seq approaches we show that *Setd1b* controls the expression of neuronal-identity genes with a broad H3K4me3 peak linked to learning and memory processes. Our data furthermore suggest that basal neuronal gene-expression is ensured by other H3K4 KMTs such as *Kmt2a* and *Kmt2b* while the additional presence of *Setd1b* at the single cell level provides transcriptional consistency to the expression of genes important for learning & memory.

One Sentence Summary: Forebrain excitatory neuron specific Setd1B knockout leads to memory impairment due to its regulatory role of controlling important neuron related genes through histone methylation.

INTRODUCTION

Cognitive diseases are a heterogeneous group of disorders that depend on complex interactions of genetic and environmental factors that activate epigenetic processes(Fischer 2014). In addition, mutations in genes that control epigenetic gene-regulation are over-represented in cognitive diseases (Kleefstra, Schenck et al. 2014). Therefore, targeting the epigenome has emerged as a promising therapeutic avenue to treat neurodegenerative and neuropsychiatric diseases (Nestler, Peña et al. 2015) (Fischer 2014). To understand the regulation of epigenetic gene-expression in the adult brain is thus of utmost importance. Histone 3 lysine 4 methylation (H3K4) is enriched around transcription start site (TSS) regions of actively transcribed genes when trimethylated (H3K4me3) (Guenther, Jenner et al. 2006). In the human brain reduced H3K4me3 has been observed in cognitive diseases such as autism spectrum disorder (Shulha, Cheung et al. 2012)or Alzheimer's disease (Gjoneska, Pfenning et al. 2015) (Kerimoglu, Sakib et al. 2017). In mammals, H3K4 methylation is mediated by six different lysine-methlytransferases (KMT's), namely KMT2A (Mll1), KMT2B (Mll2), KMT2C (Mll3), KMT2D (Mll4), SETD1A, and SETD1B that catalyze mono-, di- and trimethylation (Shilatifard 2012). The role of these enzymes in the adult brain is only beginning to emerge. Recent reports showed that *Kmt2a* and *Kmt2b* are required for hippocampus-dependent memory formation (Gupta, Kim et al. 2010) (Kerimoglu, Agis-Balboa et al. 2013) (Kerimoglu, Sakib et al. 2017), while *Setd1a* has been linked to schizophrenia (Singh T, Chheda H et al. 2016, Takata, Ionita-Laza et al. 2016, Mukai, Cannavò et al. 2019). *Setd1b* has been studied during development (Brici, Zhang et al. 2017, Schmidt, Zhang et al. 2018). Virtually nothing is known about the function of *Setd1b* in the adult brain, although mutations in *Setdb1* have been linked to intellectual disability (Labonne, Lee et al. 2016, Hiraide, Nakashima et al. 2018). To elucidate the role of *Setd1b* in the brain we generated mice that lack *Setd1b* from excitatory neurons of the adult forebrain. Our data reveal that *Sed1b* is essential for memory formation. Moreover, we provide evidence that *Setd1b* controls the expression of neuronal-identity genes that are characterized by a broad H3K4 trimethylation peak at the TSS, high expression levels and are intimately linked to learning and memory processes. Comparison of our data to those from other H3K4 KMTs suggest that this role is specific to *Setd1b* which provides to neurons transcriptional consistency to the expression of learning and memory genes.

RESULTS

Loss of Setd1b in adult forebrain neurons impairs hippocampus-dependent memory formation.
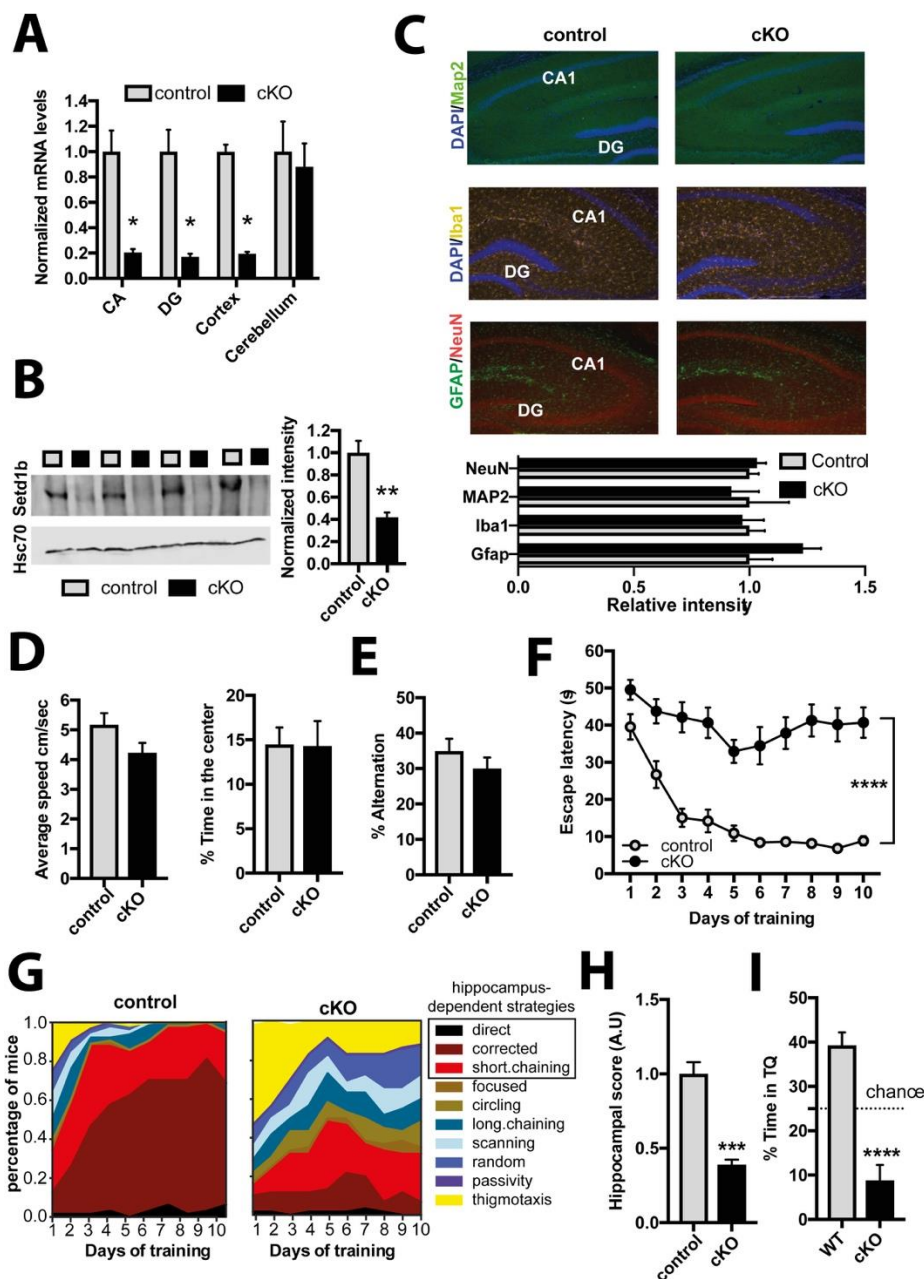
Figure 1: ***Setd1b is required for hippocampus-dependent memory.*** **A.** qPCR analysis shows loss of Setd1b in forebrain regions while levels in the cerebellum are not affected (CA: Control, n = 6; cKO, n = 6. DG: Control, n = 6; cKO, n = 6. Cortex: Control, n = 7; cKO, n = 7. Cerebellum: Control, n = 4; cKO, n = 4). * p-value < 0.05 (Student t-test). **B.** Immunoblot analysis shows loss of SETD1B in the hippocampus of Setd1b cKO mice (Control, n = 4; cKO, n = 4). ** p-value < 0.01 (Student t-test). **C.** Immunohistochemical staining (upper level) for marker proteins of neuronal integrity and quantification (lower panel) shows no difference in control and Setd1b cKO mice (NeuN: Control, n = 6; cKO, n = 6; Student t-test

p-value = 0.57. MAP2: Control, n = 4; cKO, n = 4; Student t-test p-value = 0.72. Iba1: Control, n = 5; cKO, n = 5; Student t-test p-value = 0.8. Gfap: Control, n = 5; cKO, n = 5; Student t-test p-value = 0.09.). Scale bar: 100 μm. **D.** Average speed (left panel) and time spent in the center (right) panel during exposure to the open field test was similar in control and Setd1b cKO mice (Average speed: Control, n = 15; cKO, n = 15; Student t-test p-value = 0.075. Time spent in center: Control, n = 15; cKO, n = 15; Student t-test p-value = 0.96). **E.** Short term memory was not affected in control and Setd1b cKO mice as indicated by similar percent of alternations in the Y-maze test (Control, n = 15; cKO, n = 15; Student t-test p-value = 0.3). **F.** Escape latency during water maze training indicated severe learning impairment in Setd1b cKO mice (Control: n = 15, cKO: n = 15. Repeated measures ANOVA, genotype effect: $F_{(1,28)} = 82.34$, **** p-value < 0.0001). **G.** Plots showing the specific search strategies during water maze training. Note the failure of Setd1b cKO mice to adapt hippocampus-dependent search strategies. **H.** The cognitive score calculated on the basis of the hippocampal search strategies is severely impaired in Setd1b cKO mice (Student t-test: *** p-value < 0.001). **I.** Time spent in the target quadrant during the probe test is impaired in Setd1b cKO mice (Control: n = 15, cKO: n = 15. **** Student t-test < 0.0001). Error bars indicate SEM.

To study the role of *Setd1b* in the adult brain, we crossed mice in which exon 5 of the *Setd1b* gene is flanked by loxP sites to mice that express CRE-recombinase under control of the CamKII promoter. This approach ensures deletion of *Setd1b* from excitatory forebrain neurons of the adult brain (cKO mice). Quantitative PCR (qPCR) analysis confirmed decreased expression of Setd1b from the hippocampal Cornu Ammonis (CA) area, the dentate gyrus (DG) and the cortex when compared to corresponding control littermates that carry loxP sites but do not express CRE recombinase (control group). Expression in the cerebellum was not affected confirming the specificity of the approach **(Fig 1A).** Residual expression of *Setd1b* is most likely due to the fact that deletion is restricted to excitatory neurons while other cell types are unaffected. In line with the qPCR data, SETD1B protein levels were reduced in the hippocampal CA region of *Setd1b* cKO mice **(Fig 1B)**. *Setd1b* cKO mice did not show any gross abnormalities in brain anatomy as evidenced by immunohistological analysis of DAPI staining, staining of marker-proteins for neuronal integrity Neuronal N (NEUN), microtubule-associated protein 2 (MAP2) as well as ionized calcium-binding adapter molecule 1 (IBA1) as a marker for microglia and glial fibrillary acidic protein (GFAP) as a marker for astrocytes **(Fig. 1C)**. Next, we subjected *Setd1b* cKO and control mice to behavior testing. Notably, it was previously shown that heterozygous mice expressing CRE under control of the CamKII

promoter do not differ from wild type littermates (Kuczera, Stilling et al. 2010) (Stilling, Rönicke et al. 2014) and we have confirmed this in the context of the present study also for behavior testing **(Fig. S1)**. There was no difference amongst groups in the open field test, suggesting that explorative behavior is normal in *Setd1b* cKO mice **(Fig 1D)**. Short term memory was assayed via the T-maze and was also similar amongst groups **(Fig 1E)**. Next, we subjected mice to the Morris Water Maze test to study hippocampus-dependent spatial memory. While control mice were able to learn the task as indicated by a reduced escape latency throughout the 10 days of training, *Setd1b* cKO mice were severely impaired **(Fig 1F)**. We also performed a more sensitive analysis using a modified version of the MUST-C algorithm to measure the different spatial strategies that represent either hippocampus-dependent or independent abilities (Illouz, Madar et al. 2016). Our results indicate that *Setd1b* cKO mice fail to adapt hippocampus-dependent search strategies such as "direct", "corrected" and "short-chaining" **(Fig 1G)**. Consistently, the cumulative learning score calculated on the basis of these search strategies was severely impaired in *Setd1b* cKO mice **(Fig 1H)**. To assess memory retrieval, a probe test was performed. *Set1b* cKO mice were severely impaired during the probe test performed at the end of the training **(Fig 1I)**. These data show that deletion of *Setd1b* from excitatory neurons of the adult forebrain leads to severe impairment of hippocampus-dependent learning and memory abilities.

Setd1d controls neuronal H3K4 methylation

To elucidate the molecular mechanisms by which *Setd1b* contributes to memory formation we decided to test its impact on epigenetic gene-expression in hippocampal neurons. To this end we isolated the hippocampal CA region from *Setd1b* cKO and control mice and prepared nuclei using modified fixation protocols that allowed us to perform neuron-specific chromatin-immunoprecipitation (ChIP) to study histone-modifications and RNA-sequencing to assay gene-expression from the same samples **(Fig 2A, Fig S2)**. Since SETD1B is a histone 3 lysine 4 (H3K4) methyltransferase we decided to analyze tri-methylation (H3K4me3) of histone 3 lysine 4 that is enriched at the transcription start site (TSS) of active genes and is associated with euchromatin and active gene-expression. H3K4 methylation is believed to be a stepwise process and recent data suggest that the different methylation states (from mono- to tri-methylation) at the TSS of a gene form a gradient reflecting its specific transcriptional state (Soares, He et al. 2017, Choudhury, Singh et al. 2019). Thus, we also analyzed mono-methylation of histone 3 at lysine 4 (H3K4me1). In addition, we analyzed histone 3 lysine 9 acetylation (H3K9ac), an eu-chromatin mark that was shown to partially depend on H3K4 methylation (Kerimoglu, Agis-Balboa et al. 2013, Kerimoglu, Sakib et al. 2017). Finally, we also performed Chip-seq for histone 3 lysine 27 acetylation (H3K27ac), another euchromatin mark that is linked to active gene-expression and marks promoter elements around the TSS but also enhancer regions and has not be directly linked to H3K4me3 in brain tissue. We observed that loss of *Setd1b* leads to a substantial decrease in neuronal H3K4me3 across the genome while the majority of significant changes are localized to regions in close proximity to the transcriptional start site (TSS) **(Fig 2B, C)**.

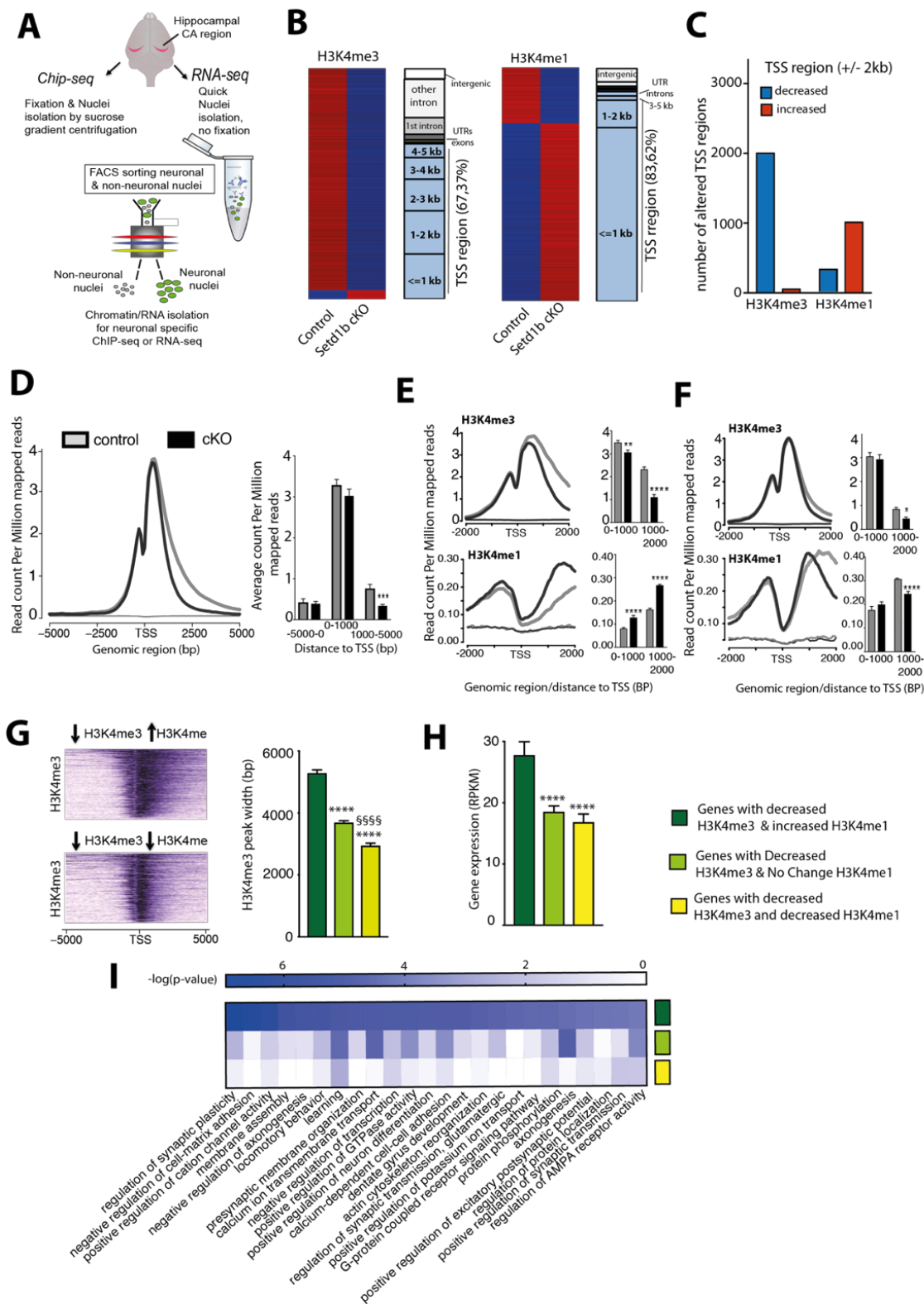Figure 2: *Setd1b controls histone-methylation and H3K4me3 peak width.* **A.** Experimental scheme showing our approach to perform cell-type specific ChIP-seq and RNA-seq. For Chip-Seq we employed n = 4 for control and n = 4 form Setd1b cKO. **B.** Left panel: Heat map showing genes with significantly differing H3K4me3 sites at the TSS (+/-2kb) in Setd1b cKO mice and the overall genomic locations of altered H3K4me3 levels. Right

panel shows the same analysis for H3K4me1 (FDR < 0.05 & |fold change| > 1.5). **C.** Bar plot showing the number of genes with decreased and increased H3K4me3 and H3K4me1 at the TSS region in Setd1b cKO mice (FDR < 0.05 & |fold change| > 1.5). **D.** NGS plot showing H3K4me3 across all genes with significantly reduced H3K4me3 in Setd1b cKO mice. Left panel shows a bar chart indicating that reduced H3K4me3 in Setd1b cKO mice is mainly occurring downstream of the TSS (*** Student t-test p-value < 0.001). **E.** NGS plots showing the distribution of H3K4me3 and H3K4me1 at the close vicinity of TSS of genes that show significantly reduced H3K4me3 and increased H3K4me1 in Setd1b cKO mice. Bar graphs on the left show corresponding quantification (Student t-test: ** p-value < 0.01, **** p-value < 0.0001). **F.** NGS plot showing the distribution of H3K4me3 and H3K4me1 at the TSS of genes that show both reduced H3K4me3 and H3K4me1 in Setd1b cKO mice. Bar graphs on the left show corresponding quantification (Student t-test: * p-value < 0.05, **** p-value < 0.0001). **G.** Heatmap (left panel) showing basal state H3K4me3 peak width for genes characterized by decreased H3K4me3 in combination with either increased or decreased H3K4me1 in Setd1b cKO mice. Right panel: Quantification of the peak width in genes with decreased H3K4me3 in combination with either increased, decreased or not altered H3K4me1 in Setd1b cKO mice (One-way ANOVA: p-value < 0.0001. Post-hoc multiple comparisons, Tukey's test: increased H3K4me1 vs no change H3K4me1, **** p-value < 0.0001; increased H3K4me1 vs decreased H3K4me1, **** p-value < 0.0001; no change H3K4me1 vs decreased H3K4me1, §§§§ p-value < 0.0001). **H.** Bar graph showing the basal wild type expression level for the 3 categories of genes that display altered H3K4me3 in Setd1b cKO mice. Please note that basal expression level is highest for genes with decreased H3K4me3 in combination with increased H3K4me1 that are characterized by broad H3K4me3 peaks (One-way ANOVA: p-value < 0.0001. Post-hoc multiple comparisons, Tukey's test: increased H3K4me1 vs no change H3K4me1, **** p-value < 0.0001; increased H3K4me1 vs decreased H3K4me1, **** p-value < 0.0001; no change H3K4me1 vs decreased H3K4me1, p-value = 0.6967). **I.** Heat map showing functional pathways for the 3 categories of genes affected by reduced H3K4me3 in Setd1b cKO mice. Error bars indicate SEM.

Similar changes were observed for neuronal H3K9ac and H3K27ac, although less regions were affected when compared to H3K4me3 (**Fig. S3**). We also observed significantly altered H3K4me1 in neurons of *Setd1b* cKO mice **(Fig 2B).** These changes were also almost exclusively detected in vicinity to the TSS **(Fig 2B)** but in contrast to the other investigated histone-modifications, many of the significantly altered genomic regions exhibited increased H3K4me1 levels in *Setd1b* cKO mice **(Fig. 2B, C)**. To further analyze these data, we first asked if the observed changes in histone-modifications occur within the same genomic regions. As expected the number of genes showing reduced H3K4me3 exceeded by far the number of genes showing reduced levels of H3K9ac, H3K27ac or

altered H3K4me1 (Fig S3). Nevertheless, almost all regions exhibiting decreased H3K9ac where also marked by decreased H3K4me3, while the regions showing decreased H3K27ac were mainly localized to different genes (**Fig. S3**). These data support previous findings, showing that H3K4me3 is functionally linked to H3K9ac (Kerimoglu, Agis-Balboa et al. 2013, Kerimoglu, Sakib et al. 2017)and suggest that the observed changes in H3K27ac are mainly due to secondary effects. Interestingly, decreased H3K4me3 in *Setd1b* cKO manifested exclusively downstream of the TSS, indicating that loss of *Sedt1b* may affect peak width **(Fig 2D)**. We decided to further explore this observation and noticed that there was an obvious difference amongst the genes that exhibit decreased H3K4me3 and increased H3K4me1 **(Fig. 2E)** when compared to genes that show exclusively decreased H3K4 methylation around the TSS **(Fig 2F)**. Namely, the change in H3K4me3 was most significant in genes with decreased H3K4me3 and increased H3K4me1 and was characterized by a substantially reduced H3K4me3 peak width **(Fig 2E)**, when compared to genes with decreased H3K4me3 and H3K4me1 **(Fig 2F)**. Findings from other cell types suggest a gradient of H3K4 methylation states in which the proximity of the mark to the TSS is correlated to the level of gene-expression. Thus, genes with broader H3K4me3 peaks at the TSS exhibit the highest and most consistent expression levels and represent genes of particular importance for cellular identity (Benayoun, Pollina et al. 2015, Soares, He et al. 2017). Indeed, our data revealed that the genes which are characterized by decreased H3K4me3 and increased H3K4me1 in *Setd1b* cKO mice, already exhibit significantly broader H3K4me3 peaks under basal conditions, when compared to genes characterized by decreased H3K4me3 but either decreased or unchanged H3K4me1 levels **(Fig 2 G)**. Interestingly, these genes were also expressed at significantly higher levels under baseline conditions **(Fig 2 H)**. Taken together, our findings suggest that *Setd1b* may be of particular importance for the expression of genes linked to the specific function of hippocampal neurons. In line with this, functional pathway analysis revealed that the genes with decreased H3K4me3 and increased H3K4me1 and thus having the broadest H3K4me3 peak under basal conditions, represent pathways intimately linked to the function of excitatory hippocampal neurons **(Fig 2I)**. Most importantly, this was not the case for the genes of the other two categories **(Fig 2I).** In summary, our data show that loss of *Setd1b* from hippocampal neurons leads to distinct changes in neuronal histone-methylation and

point to a specific role of *Setd1b* in the expression of genes essential for neuronal identity and cognitive function.

Setd1b controls the levels of highly expressed neuronal genes characterized by a broad H3K4me3 peak at the TSS

To test the impact of *Setd1b* on gene-expression directly, we analyzed the RNA-sequencing data obtained from neuronal nuclei of the same hippocampi used to generate ChIP-seq data (See Fig 2A, Fig S2). In line with the established role of H3K4me3 in active gene-expression, we mainly detected down-regulated genes when comparing control to *Setd1b* cKO mice **(Fig 3A)**.

(Continued to next page)
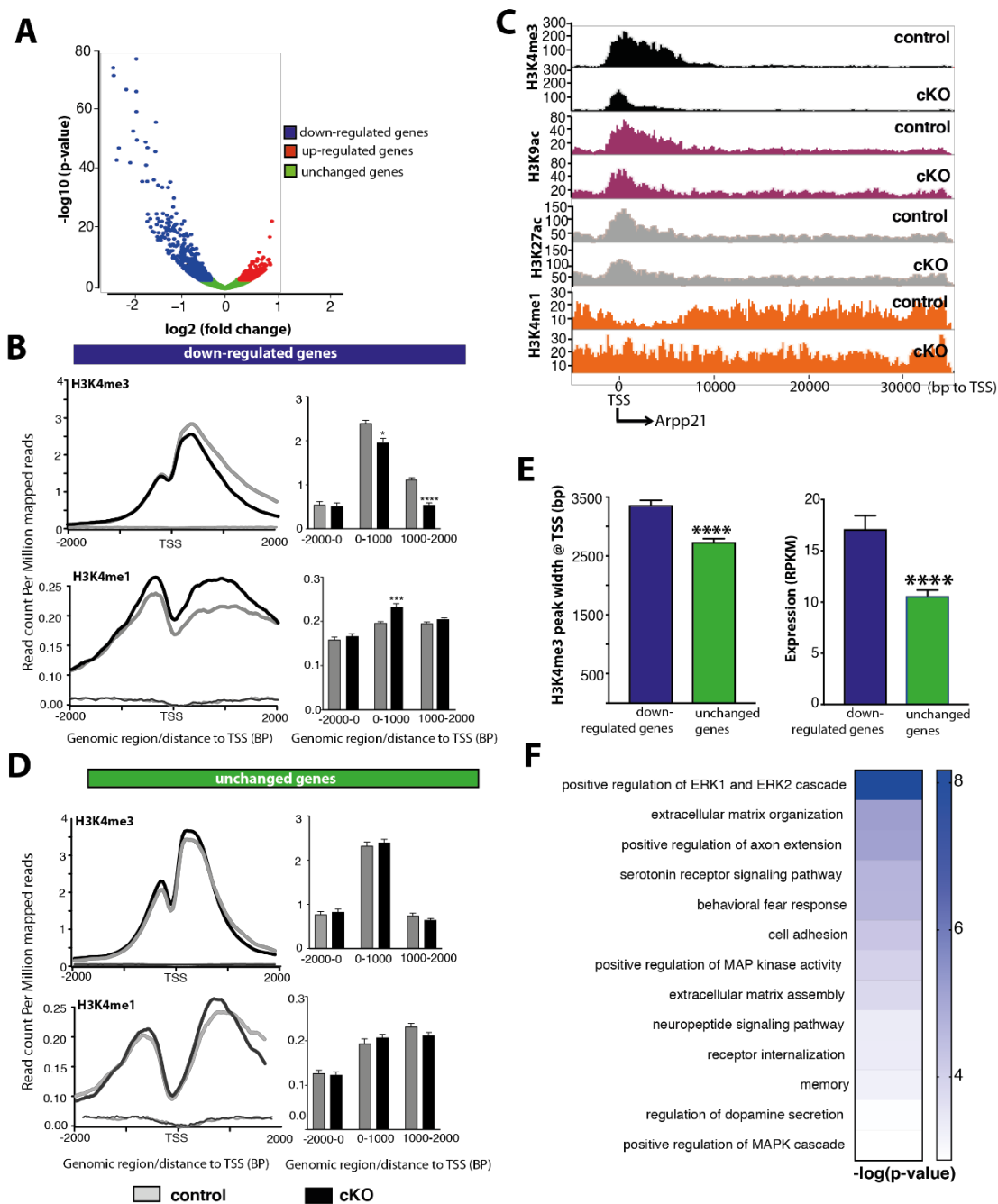
Figure 3: ***Hippocampal Setd1b controls highly expressed learning and memory genes characterized by a broad H3K4me3 peak***. A. Volcano plot showing genes differentially expressed in hippocampal neurons of Setd1b cKO mice. n = 3/group. B. NGS plots showing H3K4me3 and H3K4me1 at the TSS of genes down-regulated in Setd1b cKO mice. Bar plots (right panel) show quantification. C. Arpp21 (CAMP Regulated Phosphoprotein 21) was selected as a representative gene down-regulated in hippocampal neurons of Setd1b cKO mice to illustrate changes of the analyzed histone-modifications. Please note that the H3K4me3 peak-width is substantially shrinking in Setd1b cKO mice. At the same time there is an obvious increase of H3K4me1 at the TSS of Arpp21 in Setd1b cKO mice. D. NGS plots showing H3K4me3 and H3K4me1 at the TSS of a random set of genes that were not altered in Setd1b cKO mice. Bar plot (right panel) show quantification. E. Left panel: H3K4me3 peaks are significantly broader in genes that

are down-regulated in Setd1b cKO mice, when compared to a random set of genes that were unaffected. Right panel: Genes down-regulated in Setd1b cKO mice are characterized by higher baseline expression when compared to a random set of genes that were unaffected. F. Heat map showing functional pathways affected by genes down-regulated in Sed1b cKO mice. Error bars indicate SEM. Student t-test: * p-value < 0.05, **** p-value < 0.0001

In fact, the comparatively few up-regulated genes were all lowly expressed at baseline conditions suggesting rather unspecific effects (RPKM down-regulated genes =18.77 +/- 1.45 vs. up-regulated genes RPKM = 3.08 +/- 0.26; *P* < 0.0001). Further analysis revealed that the TSS of genes down-regulated in *Setd1b* cKO mice is characterized by significantly reduced H3K4me3 peak-width and increased H3K4me1 **(Fig 3B, C)**. This observation was specific to the genes down-regulated in *Setd1b* cKO mice, since random sets of genes that were not de-regulated in *Setd1b* cKO mice show normal H3K4me3 and H3K4me1 levels at the TSS **(Fig 3D)**. We also observed that the genes down-regulated as a result of *Setd1b* deletion were characterized by a significantly broader H3K4me3 peak and higher expression under basal conditions **(Fig 3C, E)**. A functional pathway analysis revealed that the genes down-regulated in *Setd1b* cKO mice are intimately linked to synaptic plasticity and learning and memory related processes **(Fig 3F)**. Taken together, these data further suggest that *Setd1b* controls a specific set of genes that are characterized by a broad H3K4me3 peak at the TSS, are highly expressed in hippocampal neurons under basal conditions and play a specific role in neuronal plasticity and identity.

The regulation of highly expressed neuronal identity genes with broad H3K4me3 peaks is a specific feature of Setd1b.

To provide further evidence for the specific role of *Setd1b* in the regulation of neuronal plasticity and neuronal identity genes we decided to compare *Setd1b* to other mammalian H3K4 *KTM*s. We have previously generated comparable H3K4me3 and H3K4me1 ChiP-seq data from neuronal nuclei obtained from the hippocampal CA region of mutant mice that lack either *Kmt2a* or *Kmt2b* from excitatory forebrain neurons (Kerimoglu, Sakib et al. 2017). To ensure reliable comparison we reanalyzed in parallel the H3K4me3 and H3K4me1 Chip-seq datasets obtained from hippocampal neuronal nuclei of *Kmt2a*, *Kmt2b* and *Set1b* cKO mice. In line with the previous findings, all 3 *Kmt* mutant mice exhibit a substantial amount of TSS regions with deceased H3K4me3 **(Fig 4A)**. Interestingly, we also detected TSS regions with decreased H3K4me1 in all mutant mice, but only in *Setd1b* cKO mice a substantial number of TSS regions exhibited increased H3K4me1 **(Fig 4B)**.
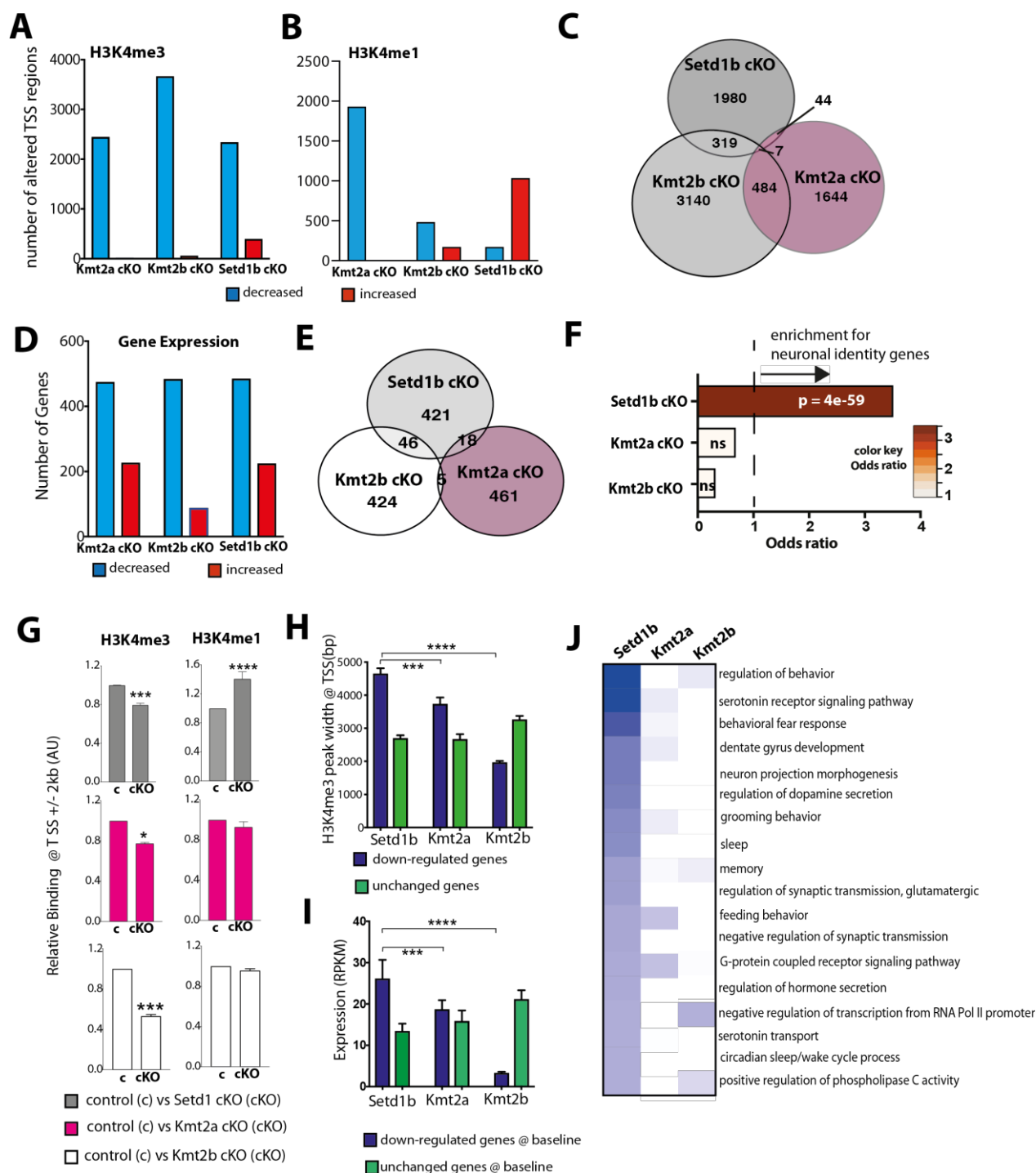
Figure 4: ***Comparative analysis of the hippocampal transcriptome in Setd1b, Kmt2a and Kmt2b cKO mice. A.*** Bar chart showing the number of genes that exhibit significantly altered H3K4me3 at the TSS (Kmt2a: control, n = 5; cKO, n = 3. Kmt2b: control, n = 6; cKO, n = 5. Setd1b: control, n = 4; cKO, n = 4). **B.** Bar chart showing the number of genes that exhibit significantly altered H3K4me1 at the TSS. **C.** Venn diagram comparing the genes with significantly decreased H3K4me3 at the TSS amongst in the 3 respective cKO mice. **D.** Bar chart showing the number of differentially expressed genes from bulk RNA-seq in each of the 3 KMT cKO mice. Kmt2a: control, n = 5; cKO, n = 6. Kmt2b: control, n = 8; cKO, n = 11. Setd1b: control, n = 6; cKO, n = 6) **E.** Venn diagram comparing the significantly

down-regulated genes H3K4me3 amongst the 3 respective cKO mice. **F.** Donw-regualted genes with decreased H3K4me3 in each of the 3 KMT cKO mice were tested for the overlap to the 836 neuronal identity genes we had defined for the hippocampal CA region (See Fig S2). Only for Setd1b cKO mice a highly significant odds ratio (Fisher's exact test) was observed, while there was no significant association amongst neuronal identity genes and the genes affected in Kmt2a and Kmt2b cKO mice. **G.** Left panel: Bar graphs showing H3K4me3 binding around the TSS of downregulated genes exhibiting significantly decreased H3K4me3 in either of the 3 KMT cKO mice (Two-way ANOVA: * p-value < 0.05, *** p-value < 0.001). Right panel depicts H3K4me1 for the same TSS regions (Two-way ANOVA **** p-value < 0.0001). Note that only in Setd1b cKO mice decreased H3K4me3 is accompanied by significantly increased H3K4me1. **H.** Genes exhibiting decreased H3K4me3 and reduced expression in Kmt2a, Kmt2b or Setd1b cKO mice were analyzed for H3K4me3 peak-width at the TSS under basal conditions. Genes affected in Setd1b cKO mice displayed significantly broader H3K4me3 peak-width when compared to genes down-regulated in Kmt2a or Kmt2b cKO mice. H3K4me3 peak-width at unchanged genes are shown for comparison. **I.** Bar graphs showing average basal expression of genes down-regulated with decreased H3K4me3 levels at the TSS in Kmt2a, Kmt2b or Setd1b cKO mice. Genes affected in Setd1cKO mice are expressed at significantly higher levels und basal conditions when compared to genes affected in Kmt2a or Kmt2b cKO mice. **J.** Heat map showing functional pathways of genes affected in Kmt2a, Kmt2b or Setd1b cKO mice. Note that genes affected by loss of Setd1b specifically represent pathways linked to neuronal function. Error bars indicate SEM.

In addition, there was little overlap amongst the TSS regions with decreased H3K4me3 in *Kmt2a*, *Kmt2b* and *Setd1b* mutant mice, providing further evidence that *Setd1b* controls a unique gene-expression program in neuronal cells **(Fig 4C)**. To test this hypothesis directly, we decided to compare the corresponding gene-expression changes in the hippocampal CA1 region of the 3 KMT mutant mice. While the ChIP-seq data available for *Kmt2a* and *Kmt2b* mutant mice had been generated from neuronal nuclei, the corresponding gene-expression analysis represents RNA-seq data obtained from bulk tissue of the hippocampal CA1 region (Kerimoglu, Sakib et al. 2017). To allow optimal comparison of these RNA-seq data to the gene-expression changes in *Setd1b* mutant mice, we also performed bulk RNA-seq from the hippocampal CA1 region of *Setd1b* cKO mice and control littermates. We observed 485 genes that were significantly down-regulated when comparing control to *Setd1b* cKO mice **(Fig 4D)**. These genes largely overlapped with the down-regulated genes detected via neuronal specific RNA-seq in *Setd1b* cKO mice **(Fig S4)**. While the total number of genes differentially expressed in the hippocampal CA1 region of *Kmt2a*, *Kmt2b* and *Setd1b* cKO mice was comparable, there

was little overlap amongst them **(Fig 4E)**. Recently, RNA-sequencing data was reported for mice that were heterozygous for *Setd1a*. Although these mutants were heterozygous constitutive knock out mice and furthermore cortical tissue was analyzed instead of the hippocampus (Mukai, Cannavò et al. 2019), it is interesting to note that there was virtually no overlap regarding the genes down-regulated in *Setd1a* knock out mice, when compared to the data obtained from our *Setd1b* cKO mice **(Fig. S5)**. Further support for a specific role of *Setd1b* in neuronal genes expression was revealed by the finding that the genes down-regulated in *Setd1b* cKO mice exhibited a significant enrichment for neuronal identity genes, while this was not the case for genes down-regulated in *Kmt2a* or *Kmt2b* cKO mice **(Fig 4F)**. In line with these data we observed that genes down-regulated in *Kmt2a* or *Kmt2B* cKO mice display decreased H3K4me3 at the TSS, while the levels of H3K4me1 were unaffected **(Fig 4G)**. In striking contrast, only the genes down-regulated in *Setd1b* cKO mice were characterized by reduced H3K4me3 and also increased H3K4me1 **(Fig 4G)**. Consequently, the genes that exhibit decreased H3K4me3 levels and were down-regulated in *Setd1b* cKO mice displayed significantly broader H3K4me3 peak at the TSS **(Fig 4H)** and were expressed a higher level under basal conditions when compared the genes controlled by KMT2A or KMT2B **(Fig 4I)**. Functional pathway analysis showed that genes affected in the 3 KMT mutant mice represent different functional pathways. Interestingly, when compared to the gene-expression data obtained from *Kmt2a* or *Kmt2b* cKO mice, genes down-regulated in *Setd1b* cKO mice represent pathways intimately linked to learning and memory and the function of hippocampal neurons **(Fig 4J)**. Further analysis revealed that the genes affected in *Kmt2a* cKO mice are enriched for more general cellular processes, for processes related to gene-expression control and also neuronal plasticity related functions **(Fig S6A)** while the genes decreased in *Kmt2b* cKO mice represent almost exclusively pathways important for basal cellular function but not specifically important for neurons **(Fig S6B)**. In sum, these data support the view that *Setd1b* is of particular importance for the expression of genes essential for the identity of hippocampal neurons and synaptic plasticity. This view is further supported by the direct comparison of hippocampus-dependent memory function in *Setd1b*, *Kmt2a* and *Kmt2b* cKO mice. While loss of any of the 3 KMT's leads to impaired spatial reference memory in the Morris water maze task (Kerimoglu, Agis-Balboa et al. 2013, Kerimoglu, Sakib et al. 2017) (see Fig 1), memory impairment is more pronounced in *Setd1b* cKO **(Fig S7)**.

Single cell expression pattern likely contributes to the distinct role of Setd1b on neuronal gene-expression.

In our effort to further elucidate the specific role of *Setd1b* on neuronal gene-expression and memory function we noticed that the levels of the H3K4 methyltransferases differ substantially in hippocampal neurons of the adult mouse brain. Surprisingly, our RNA-seq data from neuronal nuclei revealed *Setd1b* as the least expressed H3K4 methyltransferases when compared to *Kmt2a* or *Kmt2b* **(Fig 5A)**. These data might indicate that *Sedt1b* is generally expressed at very low levels or that alternatively only few cells may express *Setd1b*, a question that cannot be addressed on the basis our neuron-specific bulk RNA-seq. Thus, we decided to perform single nuclei sequencing.
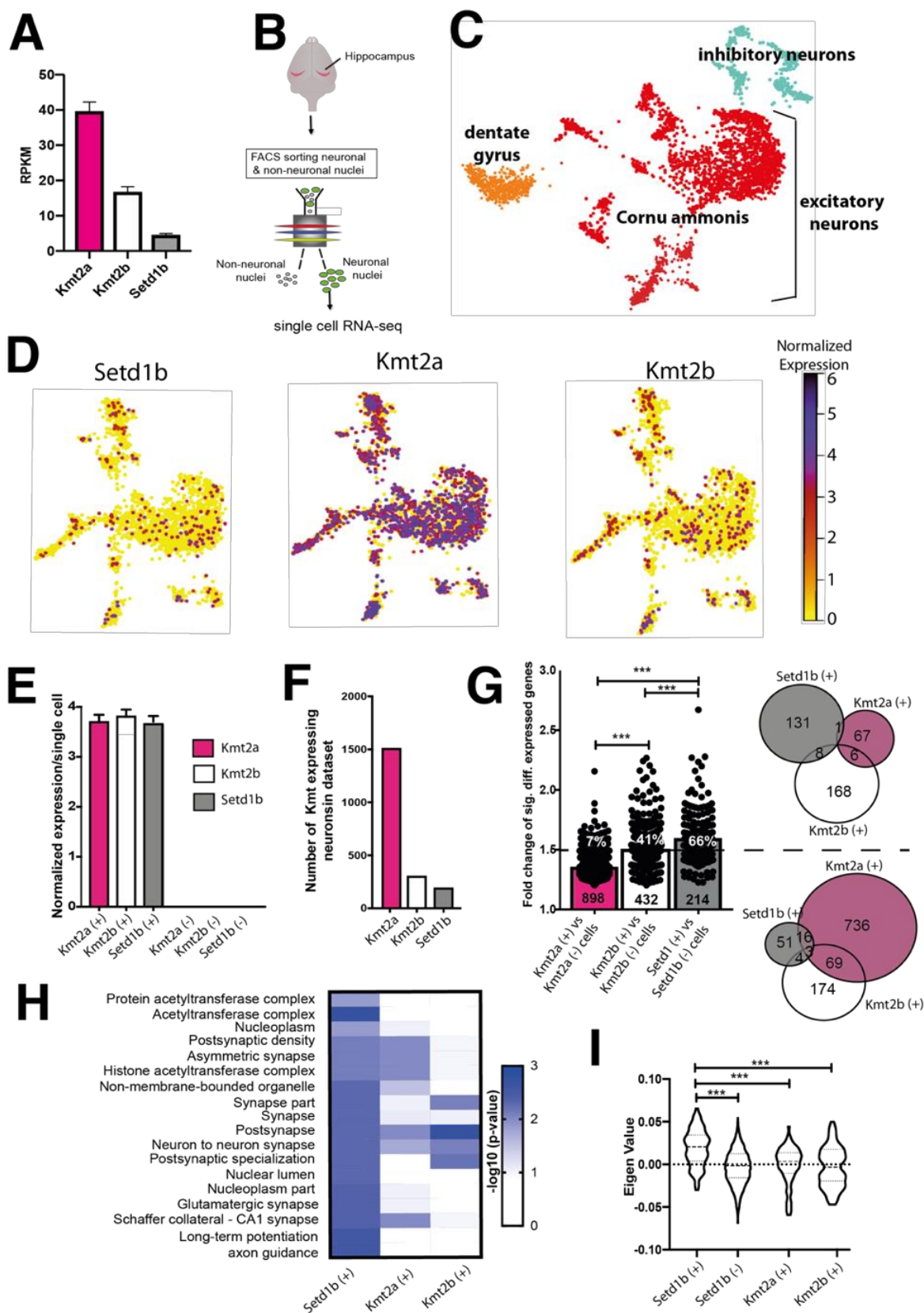
Figure 5: *Kmt2, Kmt2b and Setd1b expression at the single cell level reveals a specific role for Setd1b.* **A.** Bar graph showing the expression of Kmt2, Kmt2b and Setd1b expression in neuronal nuclei form the hippocampal CA region n = 3798. **B.** Experimental

scheme for the single nuclei RNAseq experiment. **C.** UMAP plot showing the data from 3798 neuronal nuclei. **D.** UMAP plot showing the clustering of 2619 nuclei from hippocampal excitatory CA neurons indicating the normalized expression of Setd1b (left panel), Kmt2a (middle panel) and Kmt2b (right panel). **E.** Bar graph showing the normalized expression of Kmt2, Kmt2b and Setd1b in the respective positive cells. Note that the absolute expression amongst the 3 KTMs in the respective positive cells is not different. **F.** Number of nuclei positive for of the Kmt2a, Kmt2b or Setd1b in our dataset. **G.** We performed a differential expression analysis for Kmt2a, Kmt2b and Setd1b positive nuclei vs. the nuclei that did not express the corresponding KMT. Left panel: The bar graph shows the fold change of genes significantly increased in either Kmt2a, Kmt2b or Setd1b positive nuclei. The number in the bars refer to the number of differentially expressed genes. Please note that the majority of the genes significantly enriched in Kmt2a (+) cells exhibit a rather low fold change, while this is the opposite for Setd1b (+) cells. White number within the individual plotted genes indicate the percentage of genes that are significantly increased in Kmt2a, Kmt2b or Setd1b positive nuclei by a fold change greater than 1.5. Right panel: Venn diagram comparing the genes significantly increased in Kmt2a, Kmt2b or Setd1b positive nuclei with a fold change above 1.5 (upper diagram) or below a fold change of 1.5 (lower diagram). **H.** Top GO and Kegg pathways representing the genes increased in Setd1b positive nuclei. For comparison the enrichment of the same GO-terms/pathways is shown for genes enriched in Kmt2a or Kmt2b positive nuclei. **I.** Violin plot showing the eigen-value of the gene significantly altered when comparing Setd1b (+) to Setd1b (-) cells in Setd1b (+), Setd1b (-), Kmt2a (+) and Kmt2b (+) cells. Please note that this Setd1b specific gene-set is also significantly higher expressed when compared to Kmt2a or Kmt2b positive cells (One-way ANOVA $P < 0.0001$; $F = 61.62$; asterisks indicate unpaired t-Test; ***$P < 0.001$)

We isolated the hippocampus from 3-month old wild type mice and sorted NeuN + nuclei using our established protocol **(Fig 5B)**. These nuclei were then subjected to sequencing. As expected, we detected excitatory neurons of the cornu ammonis (CA) and dentate gyrus region as well as inhibitory neurons **(Fig 5C)**. Since our analysis so far was focused on the hippocampal CA region of mice that lack *Setd1b* from excitatory neurons, we selected the CA excitatory neurons and plotted the expression of *Kmt2a*, *Kmt2b* and *Setd1b*. In line with the data obtained from bulk sequencing of hippocampal neuronal nuclei, we observed that *Kmt2a* expression was most prominent when compared to *Kmt2b* or *Setd1b* **(Fig 5D)**. However, this difference was not due to the absolute expression value per cell, which was comparable for all 3 KMT's **(Fig 5E)**. Rather, we observed that *Kmt2a* was expressed in the majority of the analyzed CA excitatory neurons, while *Kmt2b* and *Setd1b* were expressed in a comparatively small subset of cells

(Fig 5F). Next, we compared the gene-expression in neurons that either express *Kmt2a*, *Kmt2b* or *Setd1b* to cells that do not express the corresponding gene **(Fig 5E)**. A comparison of *Kmt2a* (+) vs *Kmt2a* (-) cells revealed that 897 genes were significantly increased in *Kmt2a* (+) cells, while 432 genes were expressed at higher levels in *Kmt2b* (+) and 214 genes in *Setd1b* (+) cells **(Fig 5G)**. Although the presence of *Kmt2a* affected more genes when compared to *Kmt2b* and *Setd1b,* the impact on gene-expression was comparatively small, which is indicated by the corresponding fold change of significantly differentially expressed genes. In fact, although the presence of *Setd1b* affected the least genes, the observed fold change of these genes was significantly greater when compared to *Kmt2* and *Kmt2b* **(Fig 5G)**. Only 7% of the genes significantly increased in *Kmt2a* (+) vs *Kmt2a* (-) cells showed a fold change greater than 1.5 **(Fig 5G)**. In case of *Kmt2b* (+) cells 41% of the regulated genes showed a fold change greater than 1.5 and for *Setd1b* this was true for 66 % of the regulated genes **(Fig 5G)**. These data suggest that the presence of *Kmt2a* contributes to the expression of many genes while its impact seems to be limited, when compared for example to the impact that the presence of *Setd1b* has on gene-expression. Hence, *Setd1b* affects comparatively less genes but with greater impact. When we subjected the genes specifically enriched in *Kmt2a* (+), *Kmt2b* (+) or *Setd1b* (+) cells to gene ontology and pathways analysis we observed that genes increased in the presence of *Setd1b* at the single cell level represent pathways linked to hippocampal function and interestingly also histone-acetylation **(Fig 5H)**, while these pathways were much less affected in *Kmt2a* (+) and *Kmt2b* (+) cells **(Fig. 5H)**. This is line with the finding that also at the single cell level, there is little overlap between the genes specifically enriched in either *Kmt2a* (+), *Kmt2b* (+) or *Setd1b* (+) cells **(Fig 5G.)**. We also calculated the eigen-value of the genes increased in *Setd1b* (+) cells and analyzed its expression *Setd1b* (+) and (-) cells as well as in cells positive for *Kmt2a* or *Kmt2b* **(Fig 5I)**. In line with our data, the eigen-value of *Setd1b*-specific genes was significantly increased in *Set1b* (+) cells compared to *Setd1b* (-) cells but also in cells positive for *Kmt2a* or *Kmt2b* **(Fig5I)**. In summary, these data further support a specific role of Setd1b in neuronal function and suggest that hippocampal neurons expressing *Setd1b* in addition to other H3K4 KMTs may have a "plasticity benefit".

DISCUSSION

We show that loss of *Setd1b* from excitatory forebrain neurons impairs learning and memory in mice. These data are in line with previous findings showing that hippocampal H3K4me3 increases in response to memory training in rodents (Gupta, Kim et al. 2010), while its levels are reduced in the hippocampus of a mouse for AD-like neurodegeneration (Gjoneska, Pfenning et al. 2015)and in postmortem human brain samples of patients suffering from cognitive diseases (Shulha, Cheung et al. 2012). Our data furthermore support previous genetic studies linking mutations in *Setd1b* to intellectual disability (Hiraide, Nakashima et al. 2018) (Labonne, Lee et al. 2016). Since in our work *Setd1b* is not deleted during brain development but only in the postnatal brain, the presented findings suggest that reduced *Setd1b* expression can lead to cognitive dysfunction independent of developmental alterations. Impaired hippocampus-dependent memory has been also observed in mice that lack *Kmt2a* (Gupta, Kim et al. 2010, Kerimoglu, Sakib et al. 2017) or *Kmt2b* (Kerimoglu, Agis-Balboa et al. 2013)from excitatory neurons of the adult forebrain. Interestingly, mice heterozygous for *Setd1a*, the close homologue to *Setd1b* that is genetically linked to schizophrenia, show no impairment in the water maze task but rather exhibit impaired working memory and schizophrenia-like phenotypes (Mukai, Cannavò et al. 2019). These data suggest that the different H3K4 KMT's, at least *Kmt2a*, *Kmt2b*, *Setd1a* and *Setd1b*, serve distinct functions in the adult brain. The molecular characterization of *Setd1b* cKO further confirms this view. In line with the role of *Setd1b* in regulating H3K4me4, we observed a substantial decrease of neuronal H3K4me3 and the vast majority of these changes were observed at the TSS region of genes. Our data furthermore revealed that many genes with decreased H3K4me3 also exhibited reduced H3K9ac, which is in line with previous data showing that H3K4me3 appears to be a pre-requisite for H3K9ac, most likely since H3K4 KMT's interact with histone-acetyltransferases (Wang, Zang et al. 2009, Kerimoglu, Agis-Balboa et al. 2013). For example, both *Setd1b* and the histone-acetlytransferase *Kat2a* where shown to interact with WDR5 (Lin, Min et al. 2016, Ma, Zhang et al. 2018), which is interesting since loss of *Kat2a* from excitatory forebrain neurons also leads to severely impairment of spatial reference memory (Stilling, Rönicke et al. 2014). Somewhat unexpected was the observation that H3K4me1 levels were increased at a substantial number of TSS regions that exhibited decreased H3K4me3. A similar observation has however been made in yeast that expresses only one H3K4 KMTs, namely *Set1* (Soares, He et al. 2017). The

authors show that highly expressed genes have the broadest H3K4me3 peak at the TSS, while moderate to low expressed genes are characterized by a narrow H3K4me3 peak and comparatively higher H3K4me1 levels. This H3K4me-dependent pattern of gene-expression was directly correlated to degree of SET1 activity at the TSS. In agreement with these data we observed that the genes with decreased H3K4me3 and increased H3K4me1 in *Setd1b* cKO mice are indeed characterized by a broad H3K4me3 peak at the TSS and high expression levels at the basal state. It is in this context interesting to note that *Setd1b* is the closest mammalian orthologue of the yeast SET1 protein, from which the other H3K4 methyltransferases have evolved (Shilatifard 2012). Importantly, when we analyzed gene-expression in *Setd1b* cKO mice we observed that not all genes that exhibit reduced H3K4me3 display reduced mRNA expression. Rather, we found that specifically genes with decreased H3K4me3 and increased H3K4me1, hence only the genes with the broadest H3K4me3 distribution at the TSS were significantly down-regulated in *Setd1b* cKO mice. These genes represent key pathways linked to memory formation and the identity of hippocampal neurons, which is also in line with a recent study reporting that memory training specifically activates hippocampal genes with broad H3K4me3 peaks at the TSS (Collins, Sweatt et al. 2019). In sum, these data suggest that *Setd1b* is important for the expression of neuronal identity genes in the hippocampus, that are linked to learning and memory processes, a function that might be specifically associated with *Setd1b*. Thus, when we directly compared genes down-regulated in *Setd1b*, *Kmt2a* or *Kmt2b* cKO mice, the genes down-regulated in *Setd1b* cKO mice were characterized by significantly broader H3K4me3 peaks at the TSS, significantly higher baseline expression and they were enriched for neuronal identity genes and pathways specifically important for learning-related processes. A significant gradient *Setd1b* > *Kmt2a* > *Kmt2b* was observed for all of these comparisons. In line with these data, also the memory performance in the water maze training was more severely affected in *Setd1b* cKO mice followed by mice lacking *Kmt2a* or *Kmt2b*. These data may suggest that H3K4 KTM's other than *Setd1b* are essential to ensure the sufficient expression of genes important for basal cellular processes in neurons, while *Setd1* enables to preeminent expression of neuronal identity genes. This view is in line with a recent study in mouse embryonic stem cells in which *Setd1b* was associated with the expression of highly expressed genes that exhibit a broad H3K4me3 peak, while *Kmt2b* was linked to the expression of genes with narrow H3K4me3 peaks (Sze, Ozakr et al. 2020). Interestingly, this study suggested a

functional redundancy of *Setd1b* and *Setd1a*. It is likely that this is true for gene-expression programs related to more general cellular processes. However, the situation might be different for genes specifically important in post-mitotic neurons. Moreover, mutations in either *Setd1a* or *Setd1b* lead to distinct neuropsychiatric diseases and unlike Setd1b cKO mice, *Setd1a* heterozygous mutant mice do not exhibit impairment of long-term memory consolidation (Mukai, Cannavò et al. 2019). At present we cannot conclusively answer the question how *Setd1b* affects the expression of specifically neuronal identity genes. Previous data suggest that H3K4 KMT's associate with different co-activators (Lee, Lee et al. 2006), (Dreijerink, Mulder et al. 2006, Shilatifard 2012) (Hughes, Rozenblatt-Rosenm O et al. 2004, Yokoyama, Wang et al. 2004) (Scacheri, Davis et al. 2006) which could explain the regulation of specific gene-expression programs. Our data provides a complimentary explanation that should be considered in addition. Using single-nucleus-sequencing we observed that *Setd1b* is expressed in a comparatively small number of hippocampal neurons, when compared for example to *Kmt2a*. Yet, the impact on memory function is most significant when hippocampal neurons lack *Setd1b*, as for example compared to the more abundantly expressed *Kmt2a* or *Kmt2b*. The comparison of *Setd1b* expressing hippocampal neurons to cells that do not express *Setd1b* confirmed a specific role for *Setd1b* in the expression of genes intimately linked to neuronal function and memory processes. This was different for *Kmt2a* and *Kmt2b* expressing neurons suggesting that the presence of *Setd1b* at the single cell level enables particularly efficient expression of genes important for the function of hippocampal neurons and memory consolidation. It has to be re-iterated that these genes are also detectable in neurons that lack *Setd1b* and express *Kmt2a* or *Kmt2b,* but to a lesser extent (See Fig 5I). This allows for some interesting hypotheses. For example, a number of studies demonstrated that upon learning a specific set of neurons – in most cases neurons that initiate a *cFos*-dependent gene-expression program – become part of a neuronal circuitry important for memory encoding. Particularly important are those neurons that are later reactivated during memory retrieval (Reijmers, Perkins et al. 2007, Tonegawa, Morrissey et al. 2018) (Josselyn SA 2019). All of these studies found that only a small fraction of the originally activated cells became reactivated during memory retrieval. It is thus tempting to speculate that the activity of genes such as *Setd1b* might help to shape the neuronal ensemble that will indeed be reactivated during memory retrieval, a hypothesis that would need to be tested in further studies. Taking into account

that decreased neuronal H3K4me3 levels have been observed in cognitive and neurodegenerative diseases therapeutic strategies that reinstate specifically the expression of neuronal plasticity genes controlled by *Setd1b* might be particularly helpful. We suggest that the various epigenetic drugs currently tested in pre-clinical and clinical settings for cognitive diseases should especially be analyzed for their potential to reinstate the H3K4me3 peak width at neuronal identity genes.

In conclusion, we show that *Setd1b* is essential for memory consolidation and ensures the proper expression of neuronal identity genes. Since *Setd1b* is expressed only in a subset of hippocampal neurons it may provide a plasticity benefit to those cells thereby regulating memory formation at the molecular level. In turn, *Setd1b*-related gene-expression programs could be a suitable therapeutic target to treat cognitive diseases and help patients suffering from intellectual disability.

MATERIALS AND METHODS

Animals
All animals used in this study were C57BL/6J mice and of 3-6 months of age. The experimental groups were age and sex matched. Mice were kept in standard home cages with food and water provided *ad libitum*. All experiments were performed according to the animal protection law of the state of Lower Saxony.

Behavior experiments
The behavioral experiments were performed as described previously(Kerimoglu, Sakib et al. 2017). For in depth feature analysis from water maze data, a modified version of MUST-C algorithm was used (Illouz, Madar et al. 2016).

Tissue isolation and processing
Hippocampal CA tissues were dissected from WT and *Setd1b* cKO mice, flash frozen in liquid nitrogen and stored at -80°C until further processing.

Cell-type specific nuclear RNA isolation and sequencing
Frozen CA tissues from left and right hemisphere of two mice were pooled together and processed on ice to maintain high RNA integrity. Tissue was homogenized using a plastic

pestle in a 1.5mL Eppendorf tube containing 500 uL EZ prep lysis buffer (Sigma, NUC101-1KT) with 30 strokes. The homogenate was transferred into 2 mL microfuge tubes, lysis buffer was added up to 2 mL and incubated on ice for 7 minutes. After centrifuging for 5 minutes at 500g supernatant was removed and the nuclear pellet was resuspended into 2 mL lysis buffer and incubated again on ice (7 minutes). After centrifuging for 5 minutes at 500g, the supernatant was removed and the nuclei pellet was resuspended into 500ul nuclei storage buffer (NSB: 1x PBS; Invitrogen, 0.5% RNase free BSA;Serva, 1:200 RNaseIN plus inhibitor; Promega, 1x EDTA-free protease inhibitor; Roche) and filtered through 40 µm filter (BD falcon) with additional 100 µL NSB to collect residual nuclei from the filter. Nuclei were stained with anti-NeuN-Alexa488 conjugated antibody (1:1000) for 45 minutes and washed once with NSB. Stained nuclei were then FACS-sorted with FACSaria III using 85 µm nozzle. Nuclei were gated by their size, excluding doublets and neuronal nuclei were separated from non-neuronal nuclei by their NeuN-Alexa488 fluorescence signal. Sorted nuclei were collected into a 15 mL falcon tube precoated with NSB, spun down and RNA was isolated using Trizol LS. After addition of chloroform according to the Trizol LS protocol, aqueous phase was collected and RNA was isolated by using Zymo RNA clean & concentrator-5 kit with DNAse treatment. Resulting RNA concentration were measured in Qubit and RNA-seq was performed using 100ng of neuronal RNA with illumina TruSeq RNA Library Prep Kit. Since glial nuclei are smaller and contains very little amount of RNA, neuronal nuclear RNA was scaled down and 1ng from both neuronal and glial nuclear RNA was used to make RNA-seq libraries using Takara SMART-Seq v4 Ultra Low Input RNA Kit. Libraries were sequenced using single-end 75 bp in Nextseq 550 or single-end 50 bp in HiSeq 2000, respectively.

Cell-type specific chromatin isolation and ChIP sequencing
Frozen tissues were homogenized, formaldehyde (1%) fixed for 10 minutes and quenched with 125mM glycine for 5 minutes. Debris was removed by sucrose gradient centrifugation. The resulting nuclear pellet was stained with anti-NeuN-Alexa488 conjugated antibody (1:1000) for 25 minutes and washed 3 times with PBS. Stained nuclei were then FACS sorted with FACSaria III using 85 µm nozzle. Nuclei were gated similarly as described previously(Halder, Hennion et al. 2016). Sorted nuclei were collected into a 15mL falcon tube and transferred into 1.5mL tubes. The nuclear pellet was flash frozen in liquid nitrogen and saved at -80°C for further processing. For

chromatin shearing, the pellet was resuspended into 100uL RIPA buffer (containing 1% SDS) and sonicated for 25 cycles in Diagenode bioruptor plus with high power and 30 cycles on/ 30 cycles off. Chromatin shearing was checked by taking a small aliquot and decrosslinking the DNA by 30 minutes RNAse and 2 hours of proteinase K treatment. DNA was isolated using SureClean Plus protocol. Sheared chromatin size was determined using Bioanalyzer 2100(DNA high sensitivity kit) and the concentration was measured using Qubit 2.0 fluorometer (DNA high sensitivity kit). 0.3µg chromatin was used along with 1 µg of antibody to do ChIP for H3K4me3 (Abcam ab8580), H3K4me1 (Abcam ab8895), H3K27ac (Abcam ab4729) and H3K9ac (Millipore 07-352). ChIP was performed as previously described(Halder, Hennion et al. 2016). The resulting ChIP DNA was subjected to library preparation using NEBNext Ultra II DNA library preparation kit and sequenced for single end 50bp at illumina HiSeq 2000.

ChIP-Seq Analysis

Base calling and fastq conversion were performed using Illumina pipeline. Quality control was performed using fastqc (www.bioinformatics.babraham.ac.uk/projects/fastqc). Reads were mapped to mm10 mouse reference genome with STAR aligner v2.3.0.w. PCR duplicates were removed by *rmdup -s* function of samtools. BAM files with unique reads belonging to the same group were merged into a single BAM file with the *merge* function of samtools. Profile plots were created from these merged BAM files with NGSPlot. Peak calling was performed using MACS2 against the input corresponding to the particular group (i.e., control or cKO) using q < 0.1. Consensus peaksets were generated for each histone modification individually using the Diffbind package of Bioconductor with the command *dba.count* and the parameter *minOverlap=1*. Then, these consensus peaksets were intersected with each other using the *intersect* function of bedtools with default parameters.. The differential binding analysis for each histone mark between control and *Setd1b* cKO was then performed using Diffbind with this common peakset as input. For the comparison of H3K4me3 and H3K4me1 changes in *Kmt2a* cKO, *Kmt2b* cKO and *Setd1b* cKO common peaksets for each individual histone mark from three separate ChIP-Seq experiments were extracted. In this case, first, consensus peaksets for a histone mark from each individual ChIP-Seq experiment (i.e., "Control vs Kmt2a cKO", "Control vs Kmt2b cKO" and "Control vs Setd1b cKO") were determined using Diffbind. For the purpose of comparing the effects of the three KMT knockdowns on H3K4me3 or H3K4me1 the differential binding analyses for each individual ChIP-Seq experiment were

performed always utilizing these common consensus peaksets. Diffbind package was used for differential binding analysis with in-built DESEQ2 option for differential analysis. The annotation of the genomic regions was performed with HOMER.

RNA-Seq Analysis

Base calling, fastq conversion, quality control, mapping of reads to mouse reference genome (mm10) were performed as described before (Kerimoglu, Sakib et al. 2017)Seq. Reads were counted using FeaturesCount (http://bioinf.wehi.edu.au/featureCounts/). Differential expression was analyzed with DESeq2 package of Bioconductor RPKM values were calculated using edgeR package of Bioconductor.

Single-nucleus RNA-Seq

Unfixed NeuN+ neuronal nuclei were isolated as mentioned above (section: Cell-type specific nuclear RNA isolation and sequencing). Sorted neuronal nuclei were counted in a Neubauer chamber with 10% trypan blue (in PBS) and nuclei concentration were adjusted to 1000 nuclei/µL. The nuclei were further diluted to capture and barcode 4000 nuclei according to Chromium single cell 3′ reagent kit v3 (10X genomics). Single nuclei barcoding, GEM formation, reverse transcription, cDNA synthesis and library preparation were performed according to 10X genomics guidelines. Finally, the library was sequenced in Illumina NextSeq 550 according to manufacturer's protocol. Gene counts were obtained by aligning reads to the mm10 genome (GRCm38.p4)(NCBI:GCA_000001635.6) using CellRanger software (v.3.0.2) (10XGenomics). The CellRanger count pipeline was used to generate a gene-count matrix by mapping reads to the pre-mRNA as reference to account for unspliced nuclear transcripts. The dataset contained 3841 cells with a mean of 31.053 total read counts over protein-coding genes.

The SCANPY package was used for pre-filtering, normalization and clustering (Wolf, Angerer et al. 2018) Initially, cells that reflected low-quality cells (either too many or too few reads, cells isolated almost exclusively, cells expressing less than 10% of house-keeping genes (Eisenberg and Levanon 2013) were excluded remaining in 3801 cells. Next, counts were scaled by the total library size multiplied by 10.000, and transformed to log space. A total of 3066 highly variable genes were identified based on dispersion

and mean, the technical influence of the total number of counts was regressed out, and the values were rescaled. Principal component analysis (PCA) was performed on the variable genes, and UMAP was run on the top 50 principal components (PCs) (Becht, McInnes et al. 2018), The top 50 PCs were used to build a k-nearest-neighbours cell−cell graph with k= 200 neighbours. Subsequently, spectral decomposition over the graph was performed with 50 components, and the Louvain graph-clustering algorithm was applied to identify cell clusters. We confirmed that the number of PCs captures almost all the variance of the data. For each cluster, we assigned a cell-type label using manual evaluation of gene expression for sets of known marker genes. Two cell-type clusters identified as neurons from dentate gyrus and inhibitory neurons were excluded. Remaining excitatory neuronal cells from CA region were re-clustered using the same settings as described above. For each cluster, differentially expressed genes were detected using the Wilcoxon rank-sum test as implemented in the function rank_genes_groups in SCANPY.

Data availability

All RNA and ChIP-seq datasets will be made available via GEO database

SUPPLEMENTAL DATA

Supplemental Fig. S1.
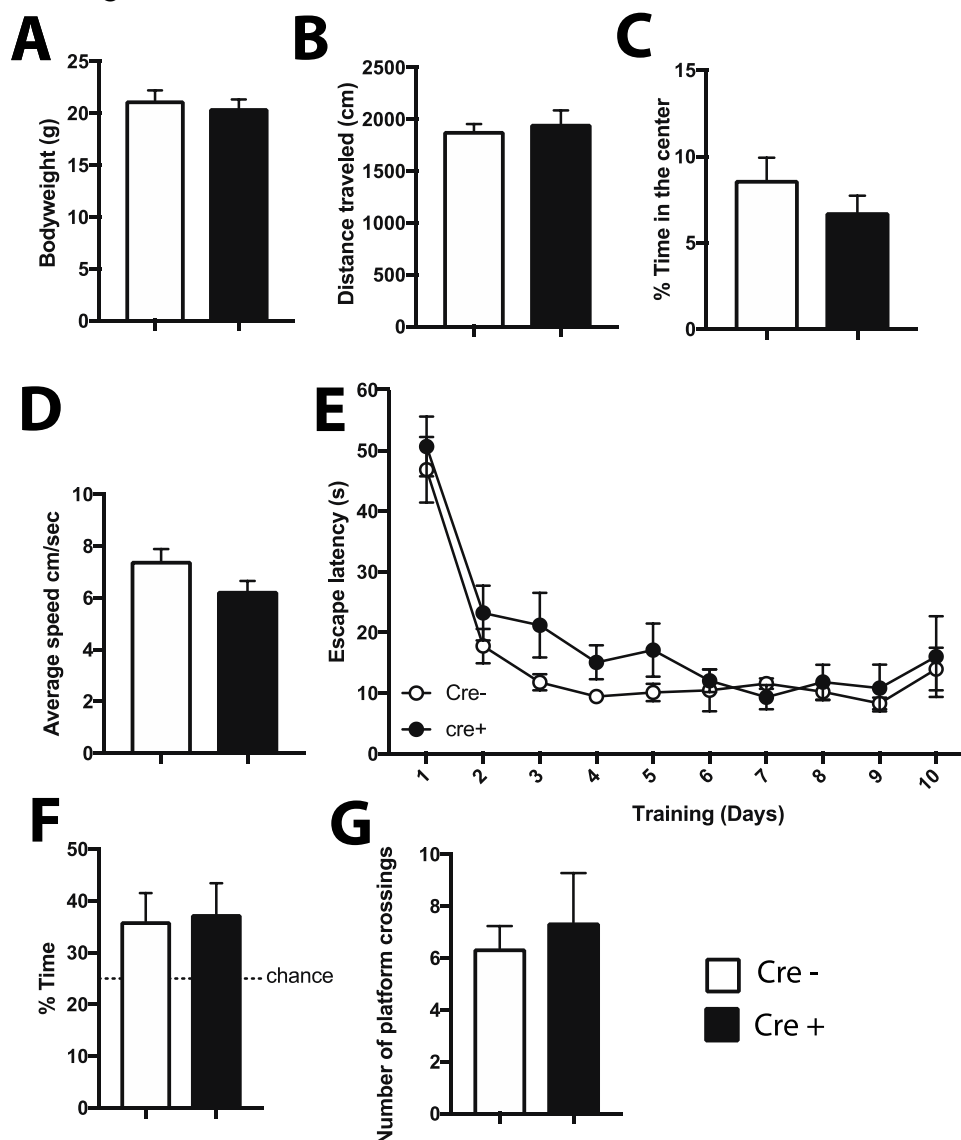
Supplemental Fig S1. Behavioral analysis of mice expressing CamKII-driven Cre recombinase. A. Transgenic mice expressing CRE under control of the CamKII promoter were subjected to behavior testing (n=8, Cre +) comparing them to wild type mice from the same breeding colony that did not express CRE (n=8; Cre -). No difference was observed in body weight. **B.** The distance traveled in the open filed test and **(C)** the time spent in the center of the arena was similar amongst groups. **D.** No difference in the swimming speed was observed amongst groups when subjected to the water maze test. **E.** Escape latency during water maze training was similar in CRE - and CRE + mice. **F.** During the probe test performed after 10 trainings days, CRE - and CRE + mice showed similar performance then time spent in the target quadrant and **(G)** the number of platform crossings were analyzed. Error bars indicate SEM.

Supplemental Fig S2.
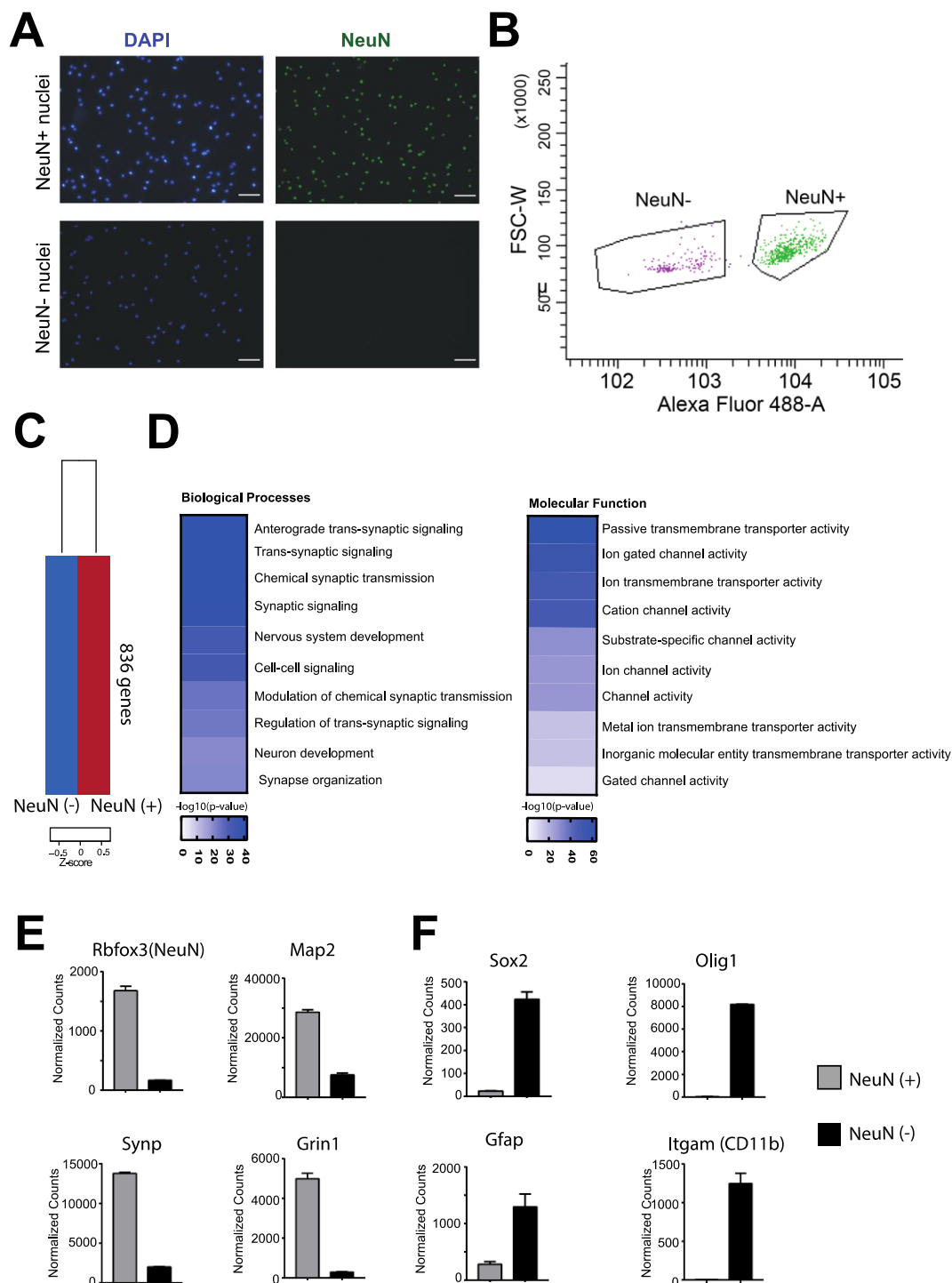
Supplemental Fig S2: Sorting neuronal and non-neuronal nuclei for RNA-analysis. Nuclei from the hippocampal CA region were subjected to FACS as depicted in Fig 2A. **A**. Representative images showing nuclei that were sorted using the neuronal marker NeuN**.** Note that no NeuN positive nuclei are detected in the NeuN (-) fraction confirming the puirity of the approach.   Scale bar: 50μm **C.** Gating strategy for NeuN (+) and NeuN (-) nuclei sorting. **C**. RNA-sequencing (n=2/group) was performed from NeuN (+) and NeuN (-) nuclei and a differential expression analysis was performed. Heat map shows 836 genes specifically enriched in NeuN (+) nuclei when compared to NeuN (-) nuclei. The criteria to select those genes were: adjusted p value <0.01, basemean >150, fold change

> 5. **D.** GO-term analysis showing that the top 10 enriched biological processes and molecular functions for the 836 genes enriched in NeuN (+) nuclei all represent specific neuronal processes. **E.** Normalized expression values obtained from the RNA-seq experiment showing the expression of selected genes known to be enriched in neurons. **F.** Normalized expression values of genes that are known to be enriched in non-neuronal cells including glia cells. Error bars indicate SEM.

Supplemental Fig. S3.
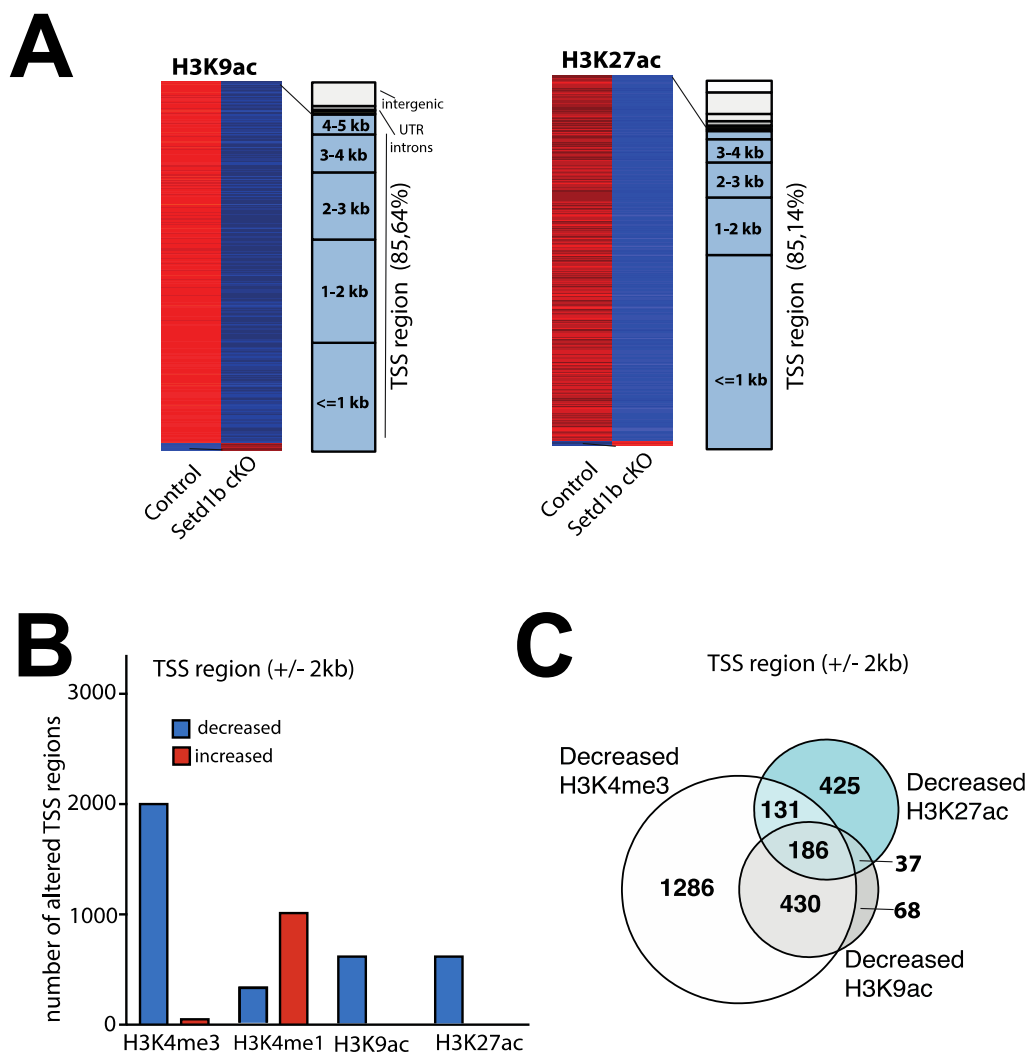
Supplemental Fig. S3. Decreased H3K9ac and H3K27ac in _Setd1b_ cKO mice. A. Left panel: Heat map showing genes with differential H3K9ac sites at the TSS in neuronal nuclei from _Setd1b_ cKO mice and their genomic location. Right panel shows the same analysis for H3K27ac. **B.** Bar chart showing the number of genes with decreased and increased H3K9ac and H3K27ac marks at the TSS region. Data for H3K4me3 and H3K4me1 are shown for comparison. As expected, the most affected histone-mark is H3K4me3. **C.** Venn diagram showing that most of the sites exhibiting decreased H3K9ac at the TSS also exhibit reduced H3K4me3, while this was not the case for H3K27ac. TSS, transcription start site.

Supplemental Fig. S4.
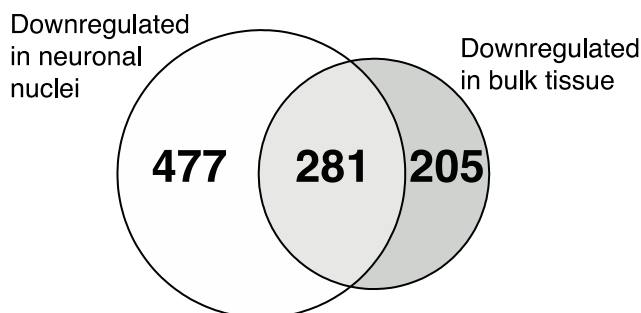
Supplemental Fig S4. Comparison of gene-expression changes in *Setd1b* cKO mice detected from cell-type specific and bulk tissue RNA-seq. Venn diagram showing the overlap of genes down-regulated in *Setd1b* cKO mice detected via RNA-seq from neuronal nuclei or bulk hippocampal CA tissue. Please note that more genes are detected when neuronal nuclei are analyzed suggesting that some of the difference are masked by cell type heterogeneity when bulk tissue is analyzed.

Supplemental Fig. S5
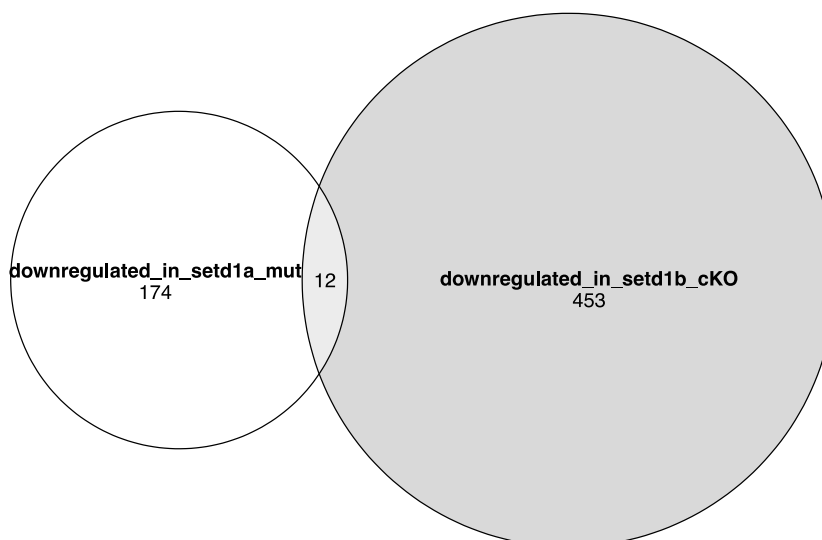
Supplemental Fig. S5. Comparison of the genes down-regulated in *Setd1A* and *Setd1b* mutant mice. Venn diagram showing genes down-regulated in *Setd1A* vs *Setd1b* cKO mice. Please note that genes affected in the different mutant mice are very different. Of course, care has to be taken since the data from *Setd1A* mutant mice stems from a recent publication by Mukai et al., 2019 (PMID:31606247). In this study cortical tissue from heterozygous mice constitutively lacking *Setd1A* were analyzed, while our data stems from the hippocampus of conditional knock out mice.

Supplemental Fig S6.

A

Downregulated Genes with Decreased H3K4me3 in cKO
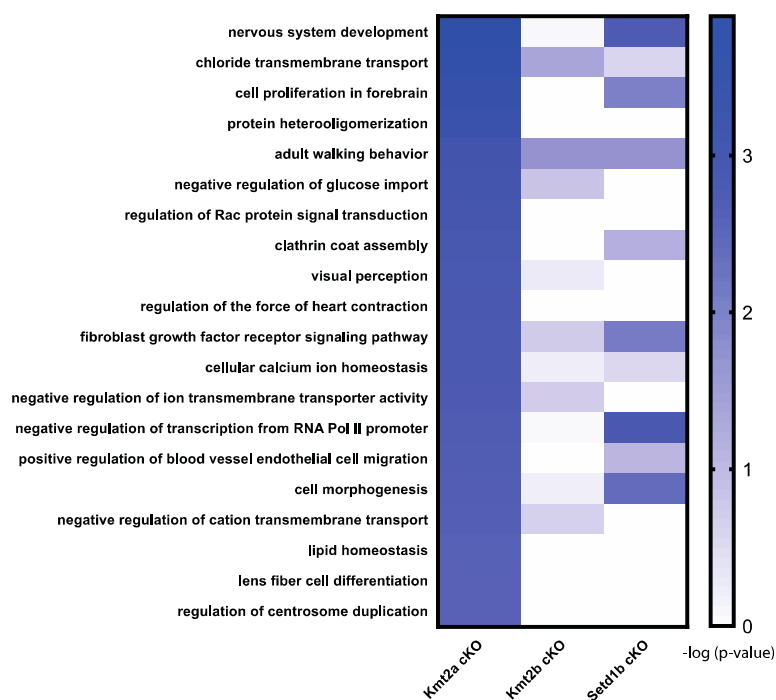
B

Downregulated Genes with Decreased H3K4me3 in cKO
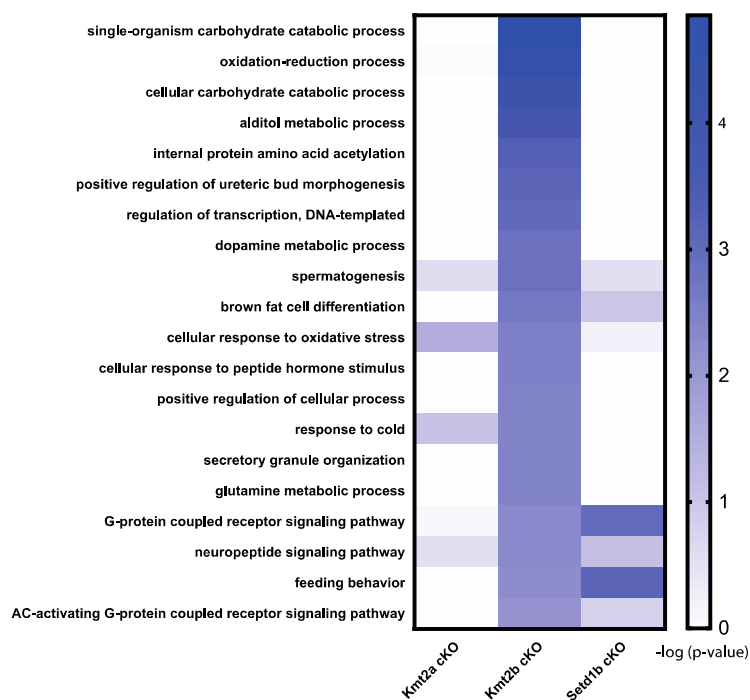
Supplemental Fig.S6. Functional pathways affected in *Kmt2a* and *Kmt2b* cKO mice. A. Heat map showing functional pathways analysis for genes down-regulated in *Kmt2a* cKO mice. Enrichment of the same pathways is also shown for *Kmt2b* and *Setd1b* cKO mice.

B. Heat map showing functional pathways analysis for genes down-regulated in *Kmt2b* cKO mice. Enrichment of the same pathways is also show for *Kmt2a* and *Setd1b* cKO mice. Please note that the pathways affected in *Kmt2a* or *Kmt2b* cKO mice differ substantially from those affected in *Setd1b* cKO mice. All data is based on comparable RNA-seq data generated from bulk hippocampal CA1 region.

Supplemental Fig S7.
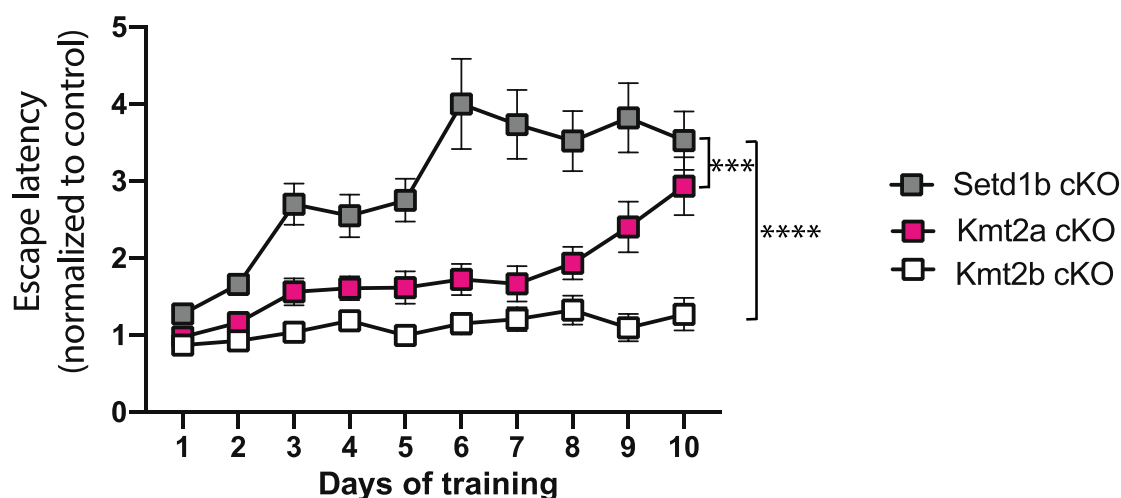
Supplemental Fig. S7. Spatial reference learning in *Kmt2a*, *Kmt2b* and *Setd1b* cKO mice. To compare spatial reference learning in the 3 different mutant mice, we normalized the data to the corresponding control group. This is important, since the experiments were performed at different time points. The data on *Kmt2a* and *Kmt2b* cKO mice was generated in our laboratory and is already published (Kerimoglu, Agis-Balboa et al. 2013) (Kerimoglu, Sakib et al. 2017). The data on *Setd1b* cKO mice were generated as part of this project using the same protocol. In this plot an increase in the normalized escape latency shows the difference to the corresponding control group. Hence a higher normalized escape latency indicates a greater difference to the corresponding control and this more severe learning impairment. This difference in normalized escape latency is significantly greater in *Setd1b* cKO mice when compared to *Kmt2a* or *Kmt2b* cKO mice. It is interesting to note that the degree of memory impairment seems to parallel the gene-expression data, hence *Setd1* appears to be most important for the expression of neuronal identity genes linked to learning and memory *Setd1b* cKO mice are also most affected in spatial memory formation. Loss of *Kmt2a* affects some pathways specific to neuronal function but also more general cellular processes while loss of *Kmt2b* seem to affect genes that are not specific to neuronal function. In line with this, loss of Kmt2b has the least effect on spatial memory and in fact memory defects in *Kmt2b* cKO mice only become obvious during prolonged training and the probe test[23]. *Setd1b* cKO (n = 14) vs *Kmt2a* cKO (n = 13): Repeated measures ANOVA, genotype effect: $F_{(1,25)} = 16.83$, *** *p*-value < 0.001. *Setd1b* cKO (n = 14) vs *Kmt2b* cKO (n = 22): Repeated measures ANOVA, genotype effect: $F_{(1,34)} = 70.66$, **** *p*-value < 0.0001.

REFERENCES

Becht, E., L. McInnes, J. Healy, C. A. Dutertre, I. Kwok, L. G. Ng, F. Ginhoux and E. W. Newell (2018). "Dimensionality reduction for visualizing single-cell data using UMAP. ." Nature biotechnology, **10.1038/nbt.4314. Advance online publication.** .

Benayoun, B. A., E. A. Pollina, D. Ucar, S. Mahmoudi, K. Karra, E. D. Wong, K. Devarajan, A. C. Daugherty, A. B. Kundaje, E. Mancini, B. C. Hitz, R. Gupta, T. A. Rando, J. C. Baker, M. P. Snyder, J. M. Cherry and A. Brunet (2015). " H3K4me3 Breadth Is Linked to Cell Identity and Transcriptional Consistency." Cell **165**(5): 1281–1286.

Brici, D., Q. Zhang, S. Reinhardt, A. Dahl, H. Hartmann, K. Schmidt, N. Goveas, J. Huang, L. Gahurova, G. Kelsey, K. Anastassiadis, A. F. Stewart and A. Kranz (2017). "The histone 3 lysine 4 methyltransferase Setd1b is a maternal effect gene required for the oogenic gene expression program. ." Development **144**.

Choudhury, R., S. Singh, S. Arumugam, A. Roguev and A. F. Stewart (2019). "The Set1 complex is dimeric and acts with Jhd2 demethylation to convey symmetrical H3K4 trimethylation." Genes & development **33**(9-10): 550-564.

Collins, B. E., J. D. Sweatt and C. B. Greer (2019). "Broad domains of histone 3 lysine 4 trimethylation are associated with transcriptional activation in CA1 neurons of the hippocampus during memory formation." Neurobiol Learn Mem. **5**(161): 149-157.

Dreijerink, K. M., K. W. Mulder, G. S. Winkler, J. W. Hoppener, C. J. Lips and H. T. Timmers (2006). "Menin links estrogen receptor activation to histone H3K4 trimethylation. ." Cancer Research **66**: 4929-4935.

Eisenberg, E. and E. Y. Levanon (2013). " Human housekeeping genes, revisited. ." Trends in genetics **29**(10): 569–574.

Fischer, A. (2014). "Epigenetic memory: the Lamarckian brain." EMBO J **33**(9): 945-967.

Gjoneska, E., A. Pfenning, H. Mathys, G. Quon, A. Kundaje, L. Tsai and M. Kellis (2015). "Conserved epigenomic signals in mice and humans reveal immune basis of Alzheimer's disease." Nature **518**(7539): 365-369.

Gjoneska, E., A. R. Pfenning, H. Mathys, G. Quon, A. Kundaje, L. H. Tsai and M. Kellis (2015). "Conserved epigenomic signals in mice and humans reveal immune basis of Alzheimer's disease." Nature **518**(7539): 365-369.

Guenther, M. G., R. G. Jenner, B. Chevalier, T. Nakamura, C. M. Croce, E. Canaani and R. A. Young (2006). "Global and Hox-specific roles for the MLL1 methyltransferase." Proc Natl Acad Sci U S A. **102**(24): 8603-8608.

Gupta, S., S. Y. Kim, S. Artis, D. L. Molfese, A. Schumacher, J. D. Sweatt, R. E. Paylor and F. D. Lubin (2010). "Histone methylation regulates memory formation." J Neurosci **30**(10): 3589-3599.

Halder, R., M. Hennion, R. O. Vidal, O. Shomroni, R. U. Rahman, A. Rajput, T. P. Centeno, F. van Bebber, V. Capece, J. C. Vizcaino, A. L. Schuetz, S. Burkhardt, E. Benito, M. N. Sala, S.

B. Javan, C. Haass, B. Schmid, A. Fischer and S. Bonn (2016). "DNA methylation changes in plasticity genes accompany the formation and maintenance of memory." <u>Nat Neurosci</u> **19**(1): 102-110.

Hiraide, T., M. Nakashima, K. Yamoto, T. Fukuda, M. Kato, H. Ikeda, Y. Sugie, K. Aoto, T. Kaname, K. Nakabayashi, T. Ogata, N. Matsumoto and H. Saitsu (2018). "De novo variants in SETD1B are associated with intellectual disability, epilepsy and autism." <u>Hum Genet.</u> **137**(1): 95-104.

Hughes, C., M,, Rozenblatt-Rosenm O, A. Milnem T, D. Copelandm T, Levinemm S.S, J. C. Lee, D. N. Hayes, K. S. Shanmugam, A. Bhattacharjee, C. A. Biondi, G. F. Kay, N. K. Hayward, J. L. Hess and M. Meyerson (2004). "Menin associates with a trithorax family histone methyltransferase complex and with the hoxc8 locus." **Mol Cell**(13): 4.

Illouz, T., R. Madar, Y. Louzon, K. J. Griffioen and E. Okun (2016). "Unraveling cognitive traits using the Morris water maze unbiased strategy classification (MUST-C) algorithm. ." <u>Brain Behav Immun.</u> **52**(2): 132-144.

Josselyn SA, K. S., Frankland PW. Finding the engram. Nat Rev Neurosci. 2015;16(9):521-534. doi:10.1038/nrn4000 (2019). "Finding the engram." <u>Nat Rev Neurosci.</u> **16**(9): 521-534.

Kerimoglu, C., R. C. Agis-Balboa, A. Kranz, R. Stilling, S. Bahari-Javan, E. Benito-Garagorri, R. Halder, S. Burkhardt, A. F. Stewart and A. Fischer (2013). "Histone-methyltransferase mll2 (kmt2b) is required for memory formation in mice." <u>J Neurosci</u> **33**(8): 3452-3464.

Kerimoglu, C., M. Sakib, G. Jain, E. Benito, S. Burkhardt, V. Capece, L. Kaurani, R. Halder, R. Agis-Balboa, R. Stilling, H. Urbanke, A. Kranz, A. Stewart and A. Fischer (2017). "KMT2A and KMT2B mediate memory function by affecting distinct genomic regions." <u>Cell Rep</u> **20**(3): 538-548.

Kerimoglu, C., M. S. Sakib, G. Jain, E. Benito, S. Burkhardt, V. Capece, L. Kaurani, R. Halder, R. C. Agís-Balboa, R. Stilling, H. Urbanke, A. Kranz, A. F. Stewart and A. Fischer (2017). "KMT2A and KMT2B Mediate Memory Function by Affecting Distinct Genomic Regions." <u>Cell reports</u> **20**(3): 538-548.

Kleefstra, T., A. Schenck, J. M. Kramer and H. van Bokhoven (2014). "The genetics of cognitive epigenetics." <u>Neuropharmacology</u> **80**(1): 83-94.

Kuczera, T., R. M. Stilling, H. E. Hsia, S. Bahari-Javan, S. Irniger, K. Nasmyth, F. Sananbenesi and A. Fischer (2010). "The anaphase promoting complex is required for memory function in mice." <u>Learn Mem</u> **18**(1): 49-57.

Labonne, J. D., K. H. Lee, S. Iwase, I. K. Kong, M. P. Diamond, L. C. Layman, C. H. Kim and H. G. Kim (2016). "An atypical 12q24.31 microdeletion implicates six genes including a histone demethylase KDM2B and a histone methyltransferase SETD1B in syndromic intellectual disability." <u>Hum Genet.</u> **135**(7): 757-771.

Lee, S., D. K. Lee, Y. Dou, J. Lee, B. Lee, E. Kwak, Y. Y. Kong, S. K. Lee, R. G. Roeder and J. W. Lee (2006). "Coactivator as a target gene specificity determinant for histone H3 lysine 4 methyltransferases. ." <u>Proc Natl Acad Sci U S A.</u> **103**: 15392-15397.

Lin, H., Z. Min and Q. Tao (2016). "The MLL/Setd1b methyltransferase is required for the Spemann's organizer gene activation in Xenopus. ." <u>Mech Dev. </u>**142** (1-9).

Ma, M., Y. Zhang, M. Weng, Y. Hu, Y. Xuan, Y. Hu and K. Lv (2018). "lncRNA GCAWKR Promotes Gastric Cancer Development by Scaffolding the Chromatin Modification Factors WDR5 and KAT2A." <u>Mol Ther.</u> **26**(11): 2658-2668.

Mukai, J., E. Cannavò, G. W. Crabtree, Z. Sun, A. Diamantopoulou, P. Thakur, C. Y. Chang, Y. Cai, S. Lomvardas, A. Takata, B. Xu and J. A. Gogos (2019). ". Recapitulation and Reversal of Schizophrenia-Related Phenotypes in Setd1a-Deficient Mice. Neuron." <u>Neuron</u> **104**(3): 471–487.

Nestler, E. J., C. J. Peña, M. Kundakovic, A. Mitchell and S. Akbarian (2015). "Epigenetic Basis of Mental Illness." <u>Neuroscientist</u> **8**(Epub ahead of print]).

Reijmers, L. G., B. L. Perkins, N. Matsuo and M. Mayford (2007). "Localization of a stable neural correlate of associative memory." <u>Science</u> **317**(5842): 1230-1233.

Scacheri, P. C., S. Davis, D. T. Odom, G. E. Crawford, S. Perkins, M. J. Halawi, S. K. Agarwal, S. J. Marx, A. M. Spiegel, P. S. Meltzer and F. S. Collins (2006). "Genome-wide analysis of menin binding provides insights into MEN1 tumorigenesis." <u>PLoS Genet</u> **2**(4): e51.

Schmidt, K., Q. Zhang, A. Tasdogan, A. Petzold, A. Dahl, B. M. Arneth, R. Slany, H. J. Fehling, A. Kranz, A. F. Stewart and K. Anastassiadis (2018). "The histone 3 lysine 4 methyltransferase Setd1b is essential for hematopoietic stem and progenitor cell homeostasis in mice." <u>eLife</u> **7**(e27157.).

Shilatifard, A. (2012). "The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis." <u>Annual reviews of Biochemistry</u> **81**: 65-95.

Shulha, H. P., I. Cheung, C. Whittle, J. Wang, D. Virgil, C. L. Lin, Y. Guo, A. Lessard, S. Akbarian and Z. Weng (2012). "Epigenetic signatures of autism: trimethylated H3K4 landscapes in prefrontal neurons." <u>Arch Gen Psychiatry</u> **69**(3): 314-324.

Singh T, K. M., Curtis D, Purcell SM, Crooks L, McRae J, Suvisaari J,, B. D. Chheda H, Breen G, Pietiläinen O, Gerety SS, Ayub M, Blyth M, Cole, C. D. T, Coomber EL, Craddock N, Daly MJ, Danesh J, DiForti M, Foster A,, G. D. Freimer NB, Johnstone M, Joss S, Kirov G, Körkkö J, Kuismin O,, H. C. Holmans P, Iyegbe C, Lönnqvist J, Männikkö M, McCarroll SA, McGuffin , M. A. P, McQuillin A, Moilanen JS, Moore C, Murray RM, Newbury-Ecob R,, P. T. Ouwehand W, Prigmore E, Rees E, Roberts D, Sambrook J, Sklar P, St, V. J. Clair D, Walters JT, Williams H; Swedish Schizophrenia Study; INTERVAL, S. P. Study; DDD Study; UK10 K Consortium, Hurles ME, O'Donovan MC, and O. M. Palotie A, Barrett JC. (2016). "Rare loss-of-function variants in SETD1A are associated with schizophrenia and developmental disorders." <u>Nat Neurosci</u> **19**(4): 571-577.

Soares, L. M., P. C. He, Y. Chun, H. Suh, T. Kim and S. Buratowski (2017). "Determinants of Histone H3K4 Methylation Patterns. Molecular cell." <u>Molecular cell</u> **68**(4): 773–785.

Stilling, R., R. Rönicke, E. Benito-Garagorri, H. Urbanke, V. Capece, S. Burckhard, S. Bahari-Javan, J. Barth, F. Sananbenesi, A. L. Schütz, J. Dyczkowski, A. Martinez-Hernandez, C.

Kerimoglu, S. R. Dent, S. Bonn, K. G. Reymann and A. Fischer (2014). "K-Lysine acetlytransferase 2A regualtes a hippocampal gene-expression network linked to memory formation." <u>EMBO J</u> **33**(17): 1912-1927.

Sze, C. C., P. A. Ozakr, K. Cao, M. Ugarenko, S. Das, L. Wang, S. A. Marshall, E. J. Rendleman, C. A. Ryan, D. Zha, D. Douillet, C. F.X. and A. Shilatifart (2020). "Coordinated regulation of cellular identity−associated H3K4me3 breadth by the COMPASS family." <u>Science Advances</u> **6**(eaaz4764).

Takata, A., I. Ionita-Laza, J. A. Gogos, B. Xu and M. Karayiorgou (2016). "De Novo Synonymous Mutations in Regulatory Elements Contribute to the Genetic Etiology of Autism and Schizophrenia." <u>Neuron</u> **89**(5): 940-947.

Tonegawa, S., M. D. Morrissey and T. Kitamura (2018). "The role of engram cells in the systems consolidation of memory." <u>Nature reviews. Neuroscience,</u> **19**(8): 485−498. .

Wang, Z., C. Zang, K. Cui, D. E. Schones, A. Barski, W. Peng and K. Zhao (2009). "Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes." <u>Cell</u> **138**(5): 1019-1031.

Wolf, F. A., P. Angerer and F. J. Theis (2018). "SCANPY: large-scale single-cell gene expression data analysis. ." <u>Genome biology</u> **19**(1): 15.

Yokoyama, A., Z. Wang, J. Wysocka, M. Sanyal, D. J. Aufiero, I. Kitabayashi, W. Herr and M. L. Cleary (2004). "Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. Molecular and cellular biology." <u>Molecular and cellular biology</u> **24**: 5639-5649.

ACKNOWLEDGMENTS

AUTHOR CONTRIBUTIONS

AM initiated the project as part of her PhD thesis, performed behavior experiments, performed and analyzed immunohistochemistry and analyzed mutant mice; MSS generated cell-type specific ChIP/RNA-seq and single nucleus RNA-seq data and contributed to the analysis of cell type specific RNA-seq, CK analyzed and interpreted ChIP-seq and RNA-seq data and supervised all bioinformatic data analysis, DMF analyzed

single nucleus RNA-seq data, LK, XX & EMZ helped with the generation of single nucleus RNA-seq; AK & FAS provided material and analyzed data, AM, SS, CK and AF designed experiments and generated figures. CK and AF wrote the paper.

Competing interests

The authors declare no competing interests

Manuscript 3
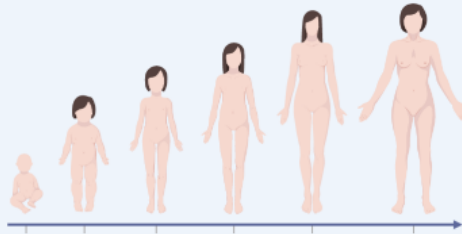
Transcriptome and epigenome changes in early aging of the hippocampus

5. Manuscript 3

Cell-type specific characterization of aging related histone methylation and transcriptome changes and detection of dysfunctional pathways in mouse hippocampal CA1.

M. Sadman Sakib[1], Cemil Kerimoglu[1], Md. Rezaul Islam[1], Lalit Kaurani[1], Alexandra Michurina[1], Susanne Burkhardt[1], Tea Berulava[1], Dennis Manfred Krüger[1], Tonatiuh Pena Centeno[1], Andre Fischer[1, 2, §]

My contributions:
First author

- Designed, planned and executed experiments.
- Utilized Neuronal (NeuN+) and non-neuronal (NeuN-) nuclear RNAseq and ChIPseq from frozen mouse hippocampal tissue to study epigenome and transcriptome of young vs aged mouse hippocampal CA1 region.
- Bioinformatic data analysis of ChIP-seq, RNA-seq and interpretation.

Manuscript publication status:

**Manuscript prepared
for submission**

Cell-type specific characterization of aging related histone methylation and transcriptome changes and detection of dysfunctional pathways in mouse hippocampal CA1.

M. Sadman Sakib[1], Cemil Kerimoglu[1], Md. Rezaul Islam[1], Lalit Kaurani[1], Alexandra Michurina[1], Susanne Burkhardt[1], Tea Berulava[1], Dennis Manfred Krüger[1], Tonatiuh Pena Centeno[1], Andre Fischer[1, 2, §]

[1]Department for Systems Medicine and Epigenetics, German Center for Neurodegenerative Diseases (DZNE), Von Siebold Str. 3a, 37075, Göttingen, Germany

[2]Department of Psychiatry and Psychotherapy, University Medical Center Göttingen, Von Siebold Str 5, 37075 Göttingen, Germany

§ Corresponding authors: andre.fischer@dzne.de

ABSTRACT

Age induced cognitive decline is one of the major etiological factors for neurodegenerative diseases. It is therefore of paramount importance to find out molecular mechanisms towards cognitive dysfunction at middle age, which lead towards further deterioration. Epigenetic mechanism deregulated in aging, such as histone modifications can act as a molecular imprint of aging in a cell, which can potentially be reversed, thereby undoing the effects of aging. As hippocampus is subjected to aging induced cognitive dysfunctions, in order to better understand the gene regulation, we sought to characterize promoter epigenetic marks, i.e. H3K4me3 and H3K27me3, along with gene expression, in a neuronal and non-neuronal cell type specific manner in 16 months old mouse hippocampal CA1 region. Along with the cell type specific characterization of epigenome and transcriptome, which will act as a valuable resource for the whole community, we found aging to affect non-neuronal cells dramatically at both epigenetic and transcriptomic levels, while neurons were less affected. Our study provides crucial evidences and important pathways showing neuronal support function, governed by non-neuronal cells were being downregulated via epigenetic and transcriptomic regulation, which can be used as a starting point of designing therapeutic interventions against dementia.

One Sentence Summary: Aging leads to non-neuronal epigenetic and transcriptomic dysfunctions related to neuronal support function, while neurons lose their inhibitory histone marks, thereby being vulnerable to degeneration in later life and cognitive decline in mice.

INTRODUCTION

Age induced cognitive decline is one of the major public health burden and primary risk factor for neurodegenerative disease across the entire world population(Yankner, Lu et al. 2008). Although severe manifestation of age-related cognitive disorders appears prominently at the end stage of life, it is a gradual process(Murman 2015). At the molecular level, aging brain across species has been shown to bear inflammatory signaling, oxidative stress, deregulated calcium signaling, mitochondrial malfunction, alternative splicing (Prolla 2002, Blalock, Chen et al. 2003, Simen, Bordner et al. 2011, Ianov, Rani et al. 2016, Ianov, De Both et al. 2017, Kerimoglu, Sakib et al. 2017). While numerous studies have showed altered epigenetic marks such as histone modifications due to aging in other cell or tissue types(O'Sullivan, Kubicek et al. 2010, Cruickshanks, McBryan et al. 2013, Shah, Donahue et al. 2013), chromatin-immunoprecipitation and sequencing (ChIP-seq) based histone modification studies on brain aging or age related neurodegenerative diseases have just started to emerge (Nativio, Donahue et al. 2018, Nativio, Lan et al. 2020). However, these studies lack the brain cell type specific analysis of histone modification which is crucial to uncover specific mechanisms in aspect of neurons or non-neuronal cells of the brain.

In previous attempts to screen global histone modification changes by western blot in brain aging, authors have shown that histone methylation marks, including H3K27me3 tend to decreased with age in 22 months of mouse brain(Gong, Qian et al. 2015). Genome wide ChIP-seq for promoter mark (H3K4me3), enhancer mark(H3K27mac) and gene expression revealed inflammatory signals across different tissues, including brain cerebellar regions(Benayoun, Pollina et al. 2019). Previous studies have mostly used whole brain tissue samples containing mixture of cells, even though it is known that gene expression or epigenetic marks vary greatly in different cell types of the brain(Kozlenkov, Roussos et al. 2014, Lake, Ai et al. 2016, Ecker, Geschwind et al. 2017). Among all other epigenetic marks, activatory H3K4me3 and inhibitory H3K27me3 marks are the most prominent marks at the promoter region, that correlate well with gene expression and associate with RNA polymerase II(Akkers, van Heeringen et al. 2009). Until now, genome wide studies on promoter epigenetic marks and gene expression in aging hippocampus in a cell type specific manner is still missing, which will be crucial to uncover

epigenomic/transcriptomic mechanisms in aging hippocampus and to find potential cure against age associated cognitive decline.

In this study, first, we utilized two different nuclei isolation methods for sorting chromatin and RNA from NeuN+ (neuronal) and NeuN- (non-neuronal) cell nuclei from the mouse brain tissues and combined basal epigenetic and transcriptomic data to corroborate activatory H3K4me3 and inhibitory H3K27me3 histone modifications with gene expression data. Applying these cell type specific genome wide techniques to young vs aged mouse hippocampal CA1 tissues led us find massive epigenetic and moderate transcriptomic deregulation at the non-neuronal cells, in contrast to neurons which were mildly affected. We further provide correlation between deregulated epigenome and gene expression of the aged non-neuronal cells to find specific pathways that potentially mediates age induced cognitive decline.

RESULTS

Cell type specific nuclei sorting reveals neuronal and non-neuron specific coding and non-coding RNAs

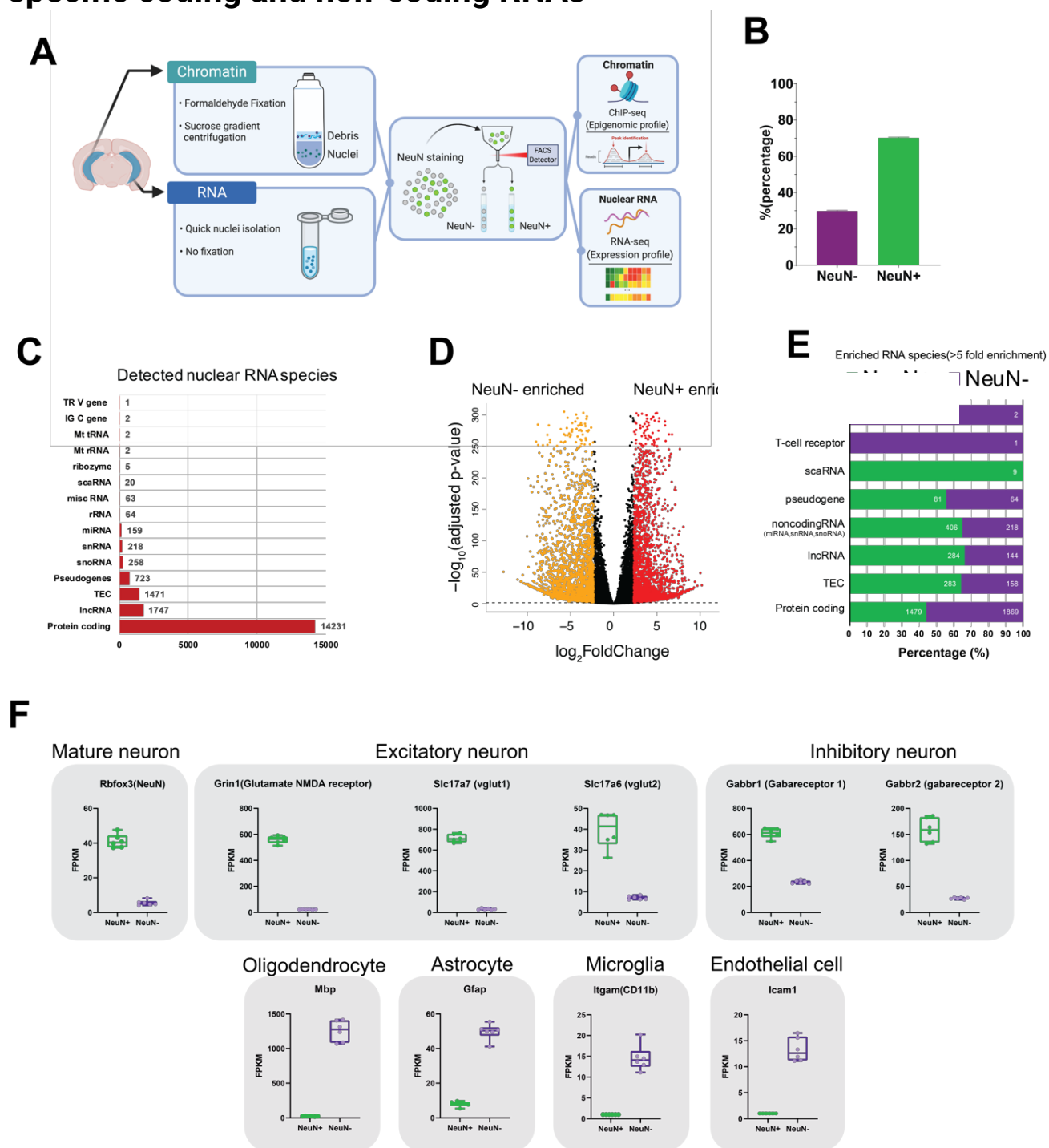**Figure 1:** *Cell type specific transcriptome and epigenome from mouse hippocampal CA1.* (A)Experimental scheme to generate both NeuN+ and NeuN- transcriptome and epigenome data using two different methods from mouse hippocampal CA1 region. (B) Percentages of NeuN+ (n=6) and NeuN- (n=6) nuclei in CA1 region. (C) Among the expressed genes in both types of nuclei, the biotypes of detected transcripts. (D)

Differential expression between NeuN+ and NeuN- nuclei at FDR 0.05 and fold change more or less than 5. (E) Biotypes of highly enriched RNAs in NeuN+ and NeuN- nuclei; TEC: Poly-A transcripts, to be experimentally validated (F) Normalized expression values of selected marker genes in NeuN+ and NeuN- nuclei RNA seq data. All differences were significantly different (Student's t-test). Error bars indicate mean ± SEM.

For unbiased transcriptome and epigenome characterization of neuronal and non-neuronal cells from the young hippocampal CA1, we utilized two different nuclei isolation methods to sort NeuN+ and NeuN- nuclei for Chromatin and RNA (Figure 1**A**). Using NeuN antibody as pan-neuronal marker, we stained neuronal nuclei and separated them from non-neuronal nuclei containing astrocytes, oligodendrocytes, microglia, endothelial cells etc. FACS sorting NeuN+/NeuN- nuclei showed that mouse hippocampal CA1 contains 70% neuronal and 30% non-neuronal cells (Figure 1**B**). Due to differences in sizes and amounts between neuronal and non-neuronal nuclei, RNA concentrations were significantly different as well (***Supplementary figure 1***). Therefore, we normalized the RNA amounts from those different cell types and performed total RNA sequencing to obtain complete transcriptome including coding and non-coding genes from neuronal and non-neuronal nuclei. Although it was nuclear RNA sequencing, majority of the expressed genes between NeuN+ and NeuN- were protein coding (Figure 1**C**). PCA plot and heatmaps showed clear and strong separation between those sample groups (Supplementary figure 2**A, B**). Using stringent cutoffs (FDR 0.05, fold change 5), we found 2391 and 2400 genes specific for NeuN+ and NeuN- (Figure 1**D**). Gene ontology and KEGG pathway analysis showed typical neuronal and non-neuronal pathways enriched in NeuN+ and NeuN- genes, respectively (Supplementary figure 2**C**). RNA biotypes among NeuN+ and NeuN- showed NeuN+ having more of all classes of RNA, except NeuN- having more protein coding genes (Figure 1**E**). Interestingly, 9 scaRNA (Small Cajal body-specific RNAs) are found to be exclusively NeuN+ enriched. scRNA are generally involved in performing modifications (e.g. methylation) for other RNA types (Cao, Rajasingh et al. 2018). Our method of generating neuronal and non-neuronal transcriptome data is highly specific as it showed different excitatory, inhibitory neuronal marker genes to be highly enriched in NeuN+ and other glial marker to be exclusively enriched in NeuN- nuclei as well (Figure 1**F**).

Cell type specific activatory H3K4me3 and inhibitory H3K27me3 histone marks explain neuronal and non-neuronal gene promoter expression
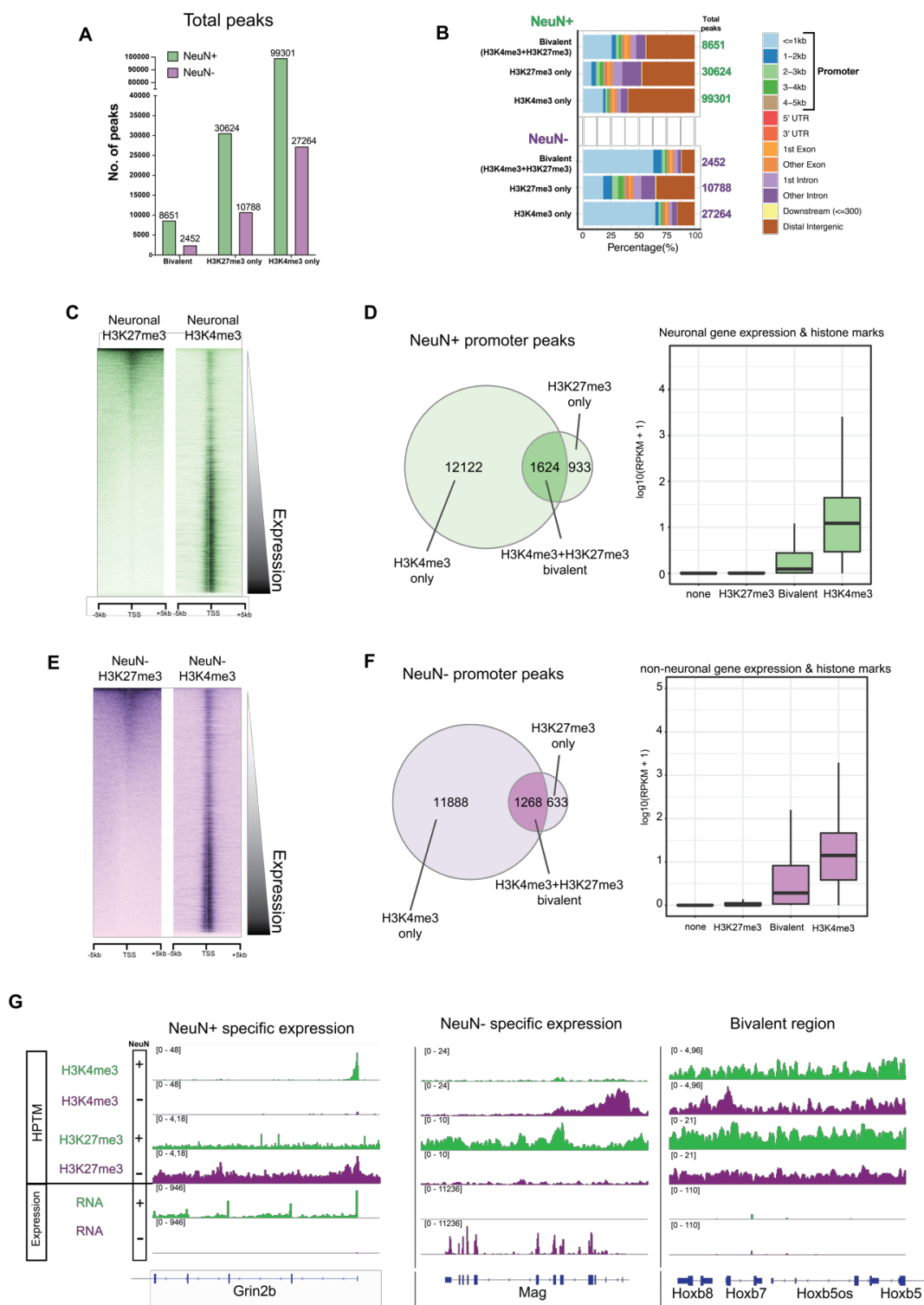

Figure 2: **H3K4me3 and H3K27me3 marks activatory, inhibitory and bivalent promoters of NeuN+/NeuN- nuclei in the brain**. (A) Total number of bivalent, H3K4me3 only and H3K27me3 only peaks detected in 3 months old NeuN+ and NeuN- samples. (B) Peaks

are annotated to find out their genomic locations in different elements such as promoters, exons etc. Despite NeuN+ having higher number of peaks, majority of them falls into distal intergenic regions. (C, E) ChIP-seq heatmap of H3K27me3 and H3K4me3 at transcription start sites, ranked by ascending level of corresponding gene expression levels for (C)NeuN+ and (E) NeuN-. (D, F) Genes either bearing H3K4me3, H3K27me3 or bivalent regions at their promoters (Venn diagram) and log2(RPKM + 1) expression pattern (box plot) are similar in both NeuN+ and NeuN- (D, F, respectively).

Since H3K4me3 and H3K27me3 are archetypical activatory and inhibitory epigenetic marks, respectively, we did genome wide ChIPseq for those two marks in NeuN+ and NeuN- cells to understand gene expression regulation of corresponding cells in the adult hippocampus. After peak calling for H3K4me3 and H3K27me3 in each cell types, peaks were sub-grouped into three classes of histone marks, (i) H3K4me3 only, (ii) H3K27me3 only, and (iii)bivalent region, which contains both H3K4me3 and H4K27me3 marks. H3K4me3 peaks were typically higher in number than H3K27me3 in both cell types and interestingly, NeuN+ had the highest amounts of peaks among all three classes, compared to NeuN- (Figure 2**A**). Despite having higher amounts of peaks in general, majority of NeuN+ peaks were at the distal intergenic regions, compared to NeuN- (Figure 2**B,** upper**).** Since NeuN- had lower peak number compared to NeuN-, majority of it were annotated around promoter (TSS) regions (Figure 2**B,** lower**).** But peaks associated around TSS (+/- 2000 bp) were similar in number (14679 for NeuN+ and 13789 for NeuN-, respectively). When ranked according to gene expression levels, both H3K4me3 and H3K27me3 marks showed oppositely positioned patterns in both NeuN+ and NeuN- cell types around TSS regions (Figure 2**C,E).** Genes bearing all three classes of histone marks around their TSS regions were also found to be similar for both NeuN+ and NeuN- cells (Figure 2**D,F;** Venn diagrams). H3K4me3 mark bearing genes were expressed at the highest RPKM levels, in contrast to H3K27me3 mark bearing genes being not expressed, while bivalent mark bearing genes were being expressed at either very low levels or did not express at all (Figure 2**D,F;** box plots). In terms of peak height and width, NeuN+ peaks were significantly higher than NeuN- in all epigenetic marks, except for bivalent H3K4me3 peaks, which were similar in sizes between NeuN+ and NeuN- (Supplementary figure **3**). Gene ontology analysis (Supplementary figure **4**) for unique genes bearing H3K4me3, H3K27me3 or bivalent marks at their promoter suggests H3K4me3 is related to active NeuN+ or NeuN- related functional pathways, while inhibitory H3K27me3 marks genes to repress pathways that are not usually expressed in neuronal or non-neuronal cells.

Bivalent marks are in similar development related pathways for both NeuN+ or NeuN-, which are also not regularly expressed in healthy situation.

Genome browser view (Figure 2**G**) of those three classes of histone marks showed Neuron specific gene, Grin2b, has sharp activatory H3K4me3 peak at TSS, with no H3K27me3 peaks, and therefore, high expression levels in NeuN+ nuclei. But in NeuN-, it is not expressed at all as it contained higher level of inhibitory H3K27me3 levels that repressed the expression and no H3K4me3 peaks. Similar pattern was observed for non-neuronal, specifically oligodendrocyte related gene, Mag, which had higher activatory H3K4me3 peak at TSS with no H3K27me3 enrichment in NeuN-, which resulted into high level of expression only in NeuN- nuclei. NeuN+ did not express this gene as it had high enrichment of H3K27me3 and no H3K4me3 that resulted into suppressed gene expression. As an example of bivalent regions, Hox gene cluster were visualized and found both NeuN+ and NeuN- bears high levels of H3K4me3 and H3K27me3 marks that resulted into no gene expression in both cell types.

Aging related cognitive decline corresponds to massive epigenome changes at the promoter of both neuronal and non-neuronal cells in hippocampal CA1
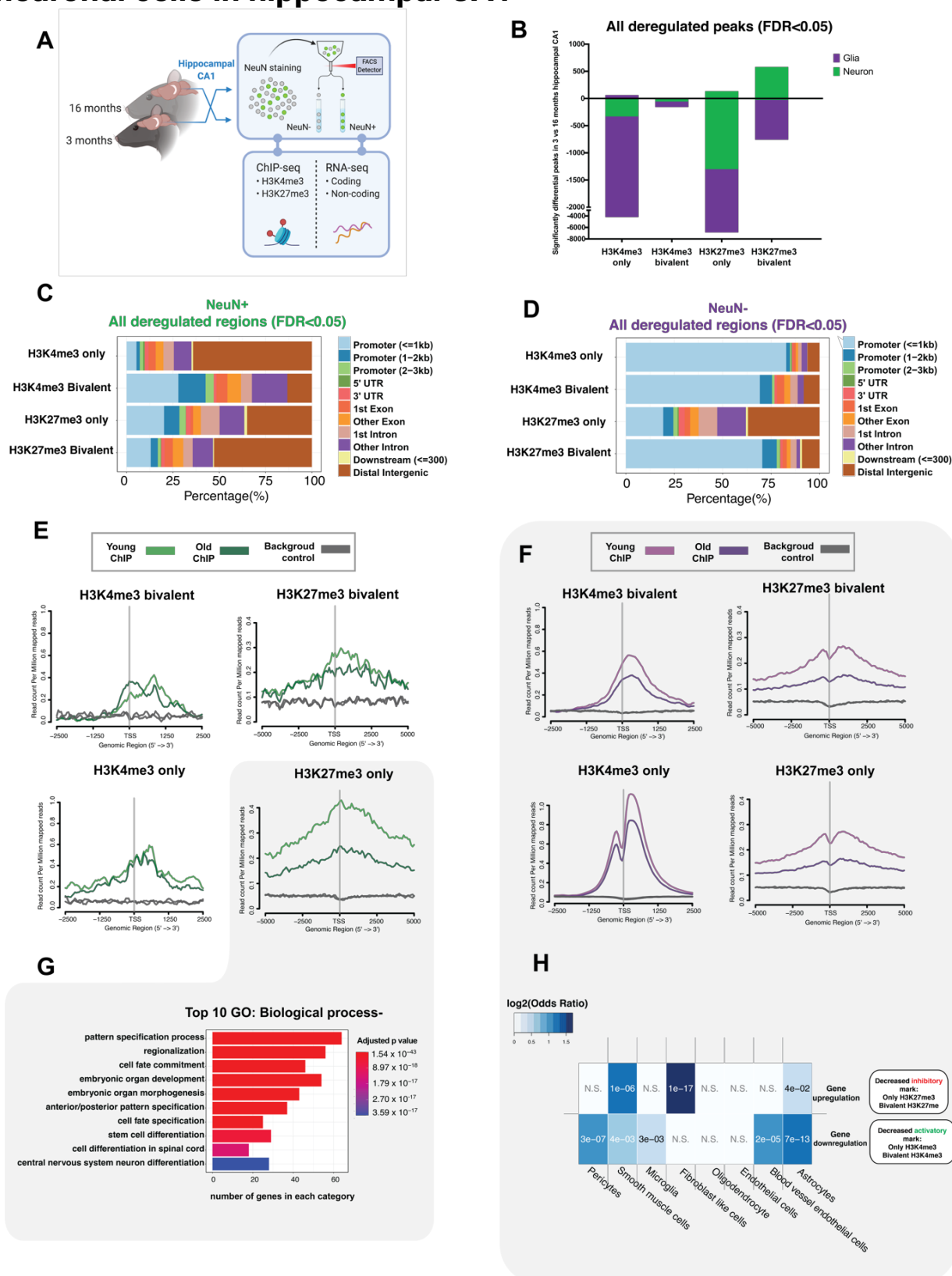

Figure 3: *Cell-type specific epigenome and transcriptome analysis in 16 months aged mouse hippocampal CA1.* (A) Experimental scheme for characterizing both epigenome

and transcriptome in 3 vs 16 months old mice hippocampal CA1 (see methods for replicate numbers). (B) Number of significantly increased or decreased histone marks in NeuN+ and NeuN- at FDR < 0.05 (C) NeuN+ deregulated histone modification peaks are distributed all over different genomic elements. (D) NeuN- deregulated histone modification peaks are distributed mostly at the promoter regions, with exception to H3K27me3 only. (E) H3K27me3-only peaks at the promoters of NeuN+ cells were found to be decreased due to aging. (F) All analyzed histone modifications at around TSS of NeuN- were decreased due to aging. (G) Gene ontology: biological processes for age-related H3K27me3 only decrease at the gene promoters of NeuN+. (H)Odds ratio analysis to find out which non-neuronal cell types are affected in aspect of possible gene up or downregulation according to deregulated histone code in aged hippocampal CA1 NeuN-cells; N.S. : not significant, adjusted p values for overlap are in each box.

In order to characterize cell type specific epigenome and transcriptome changes due to aging associated cognitive decline, NeuN+ and NeuN- nuclei were sorted from 3 and 16 months old mouse hippocampal CA1 tissues for ChIP-seq of H3K4me3 and H3K27me3 marks, in parallel with total RNA-seq for measuring coding and non-coding RNAs (Figure 3**A**). Along with H3K4me3 and H3K27me3 regions of the genome, bivalent promoter regions were also analyzed with its H3K4me3 or H3K27me3 contents separately. Among these four groups of histone marks, differential binding between old vs young revealed higher degree of changes in NeuN-, compared to NeuN+ samples, with overall decrease of histone marks, except increase in bivalent H3K27me3 marks of NeuN+ (Figure 3**B**). Since H3K4me3 and H3K27me3 are opposing histone marks in terms of their regulatory functions, their decrease would possibly exert both up and down regulatory gene expression effects in NeuN+ and NeuN- cells. Annotation of deregulated peaks showed NeuN- cells promoters to be highly affected due to aging, compared to NeuN+ (Figure 3**C, D**). To visualize the magnitude of changes at the transcription start sites, ChIP-seq data of deregulated genes were plotted. H3K27me3-only peaks at the promoter of aged NeuN+ was decreased, while other marks did not show significant changes (Figure 3**E)**. On the contrary, all marks showed decreased enrichment at the TSS regions of aged NeuN- (Figure 3**F)**. Decreased H3K27me3-only genes in aged NeuN+ were all related to cell fate commitment, differentiation or cell development related genes, which are usually suppressed in mature neurons (Figure 3**G)**. Recent study has implicated that removing inhibitory H3K27me3 mark from neurons by knocking out PRC2, it's methyltransferase enzyme, resulted in gene de-repression and aberrant expression of non-neuron or development related genes that led to neurodegeneration at later months (von

Schimmelmann, Feinberg et al. 2016). This implies, 16 months old neurons lost their insulating H3K27me3 from promoter, which would make them vulnerable to age related neuronal dysfunctions by probable aberrant non-neuron related gene expression later in life, leading to cognitive defect.

As for NeuN-, since all epigenetic marks decreased at the TSS region (Figure 3**F)**, to understand it's probable gene regulation, we grouped these changes into two categories according to canonical histone codes, for decreased H3K27me3-only and bivalent-H3K27me3: gene upregulation and for decreased H3K4me3-only and bivalent-H3K4me3: gene downregulation. Since NeuN- is combination of different glial cells, we wanted to find which glial type is mostly affected. From a recent mouse brain single cell RNA sequencing dataset focusing mostly on major brain glial cell types as well as brain vasculature cells including pericytes, fibroblast like cells, smooth muscle cells etc (Vanlandewijck, He et al. 2018), we obtained top 500 genes expressed into those cell clusters, and took the unique genes in each of those cell types. Odds ratio gene overlap analysis between our age deregulated genes due to decreased NeuN- epigenome and glial cell types was done to find out genes of which glial types are up- or down- regulated. This analysis showed pericytes, astrocytes, smooth muscle cells and blood vessel endothelial cell related genes lost the activatory H3K4me3 marks, which might lead to loss of function for these important glial types, that gives support to the blood brain barrier (Figure 3**H)**. Microglia related genes also lost activatory marks, which were lysosome related genes (Supplementary figure 5). This might pave the way for lysosomal dysfunction in later aging in microglia. Microglia typically has deteriorated functions, including hypo-motility, burdened lysosomes, in aging and neurodegeneration, leading to cognitive decline (Pluvinage, Haney et al. 2019). For epigenetic changes leading to possible gene upregulation, it was Fibroblast like cells, smooth muscle cells and astrocyte related genes (Figure 3**F)**. Microglial gene upregulation due to aging or inflammation is regulated by enhancer activation (Wendeln, Degenhardt et al. 2018), therefore it was expected that we do not see our promoter-epigenome regulated gene upregulation do not overlap with microglia. To conclude, our data shows evidence of both H3K4me3 and H3K27me3 epigenetic deregulation specifically in non-neuronal cells, while only inhibitory H3K27me3 mark was deregulated in neurons.

Transcriptional changes in neurons and non-neuronal cells in 16 months old mouse hippocampal CA1 region
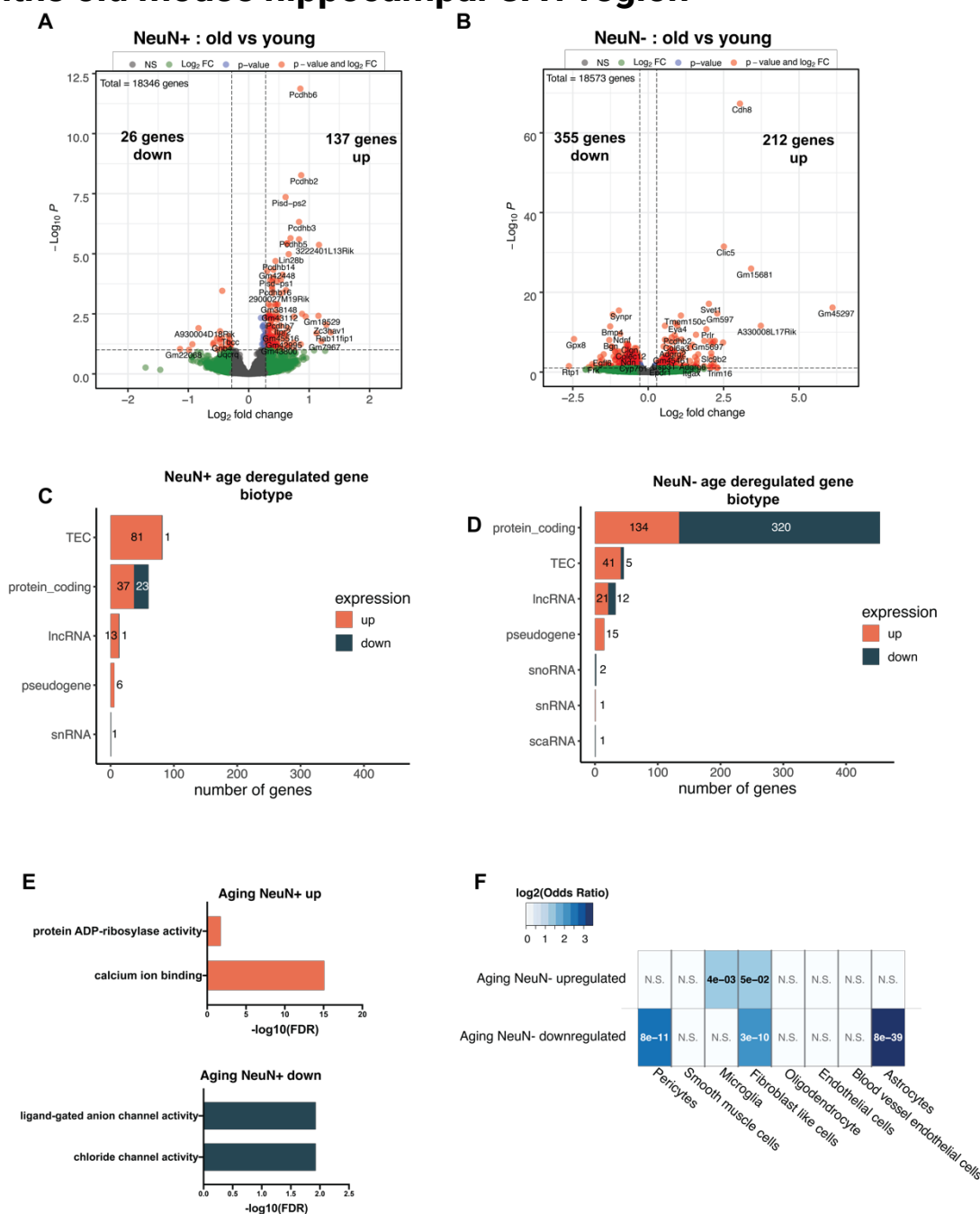
Figure 4: ***Gene expression changes in aging neurons and non-neuronal cells*** (A) Volcano plot for old vs young NeuN+ nuclei show mostly upregulated genes. Significance cutoff: FDR <0.1, fold change 1.2. (B) Volcano plot for old vs young NeuN- nuclei showing bidirectional gene expression changes. Significance cutoff: FDR <0.1, fold change 1.2 (C) RNA biotypes for up and downregulated NeuN+ genes in aging. (D) RNA biotypes for up and downregulated NeuN- genes in aging. (E) GO: molecular function analysis for up (top) and down(bottom) regulated genes in aged NeuN+ nuclei. (F) Odds ratio analysis to find out which non-neuronal cell types are affected due to gene expression changes in aging NeuN- nuclei; N.S.: not significant, adjusted p values for overlap are in each box.

Since we observed massive epigenome changes in thousands of H3K4me3 and H3K27me3 peaks at 16 months of age in the hippocampal CA1, we were wondering whether these epigenome changes conferred to gene expression changes as well. Therefore, we also generated NeuN+/NeuN- nuclei specific nuclear RNAseq data. Differential expression analysis between old vs young NeuN+ nuclei showed 137 genes was upregulated and 26 genes was downregulated (
Figure *4A*). But only few of them were protein coding genes (37 up, 23 down), while majority of them were found to be termed as TEC (To be Experimentally Confirmed), according to Ensembl annotation (
Figure *4C*). These transcripts are expressed sequence tags (EST) having poly A features that putatively code for protein, but have not been studied yet or assigned with any specific gene names. Gene ontology analysis in aged deregulated NeuN+ genes showed that the upregulated protein coding genes are related to calcium ion signaling and Poly (ADP-ribose) polymerase (PARP) activity (
Figure *4E,* top), both of which have been implicated to drive age related neuronal dysfunctions via NAD+ overuse and depletion (Thibault, Gant et al. 2007, Salech, Ponce et al. 2020). Downregulated protein coding genes were related to anion/chloride channel activity (
Figure *4E,* bottom), associated with downregulation of gamma-aminobutyric acid type A receptor subunit gamma1 (Gabrg1) and glycine receptor alpha 2 (Glra2). In contrast to NeuN+, NeuN- nuclei were affected more with higher number of significantly deregulated genes (
Figure *4B*), as we have seen before in terms of epigenome changes (Figure 3). Majority of those deregulated genes were protein coding genes, with more than twice the number of genes significantly downregulated (320 genes), than upregulated (134 genes). Similar to Figure 3H, we wanted to know which glia types were functionally most affected due to up or downregulated genes. For upregulated genes, microglia and fibroblast like cell related genes were significantly overlapped. But for downregulated genes, the highest significant overlap was for Astrocytes and pericytes, important regulator of neuronal support function in the brain. Mechanistically, along with inflammatory cell related genes going up, supporting non-neuronal cell related genes are going down in mouse hippocampal CA1 aging, while neuronal genes related to calcium signaling and PARP activity.

Decreased H3K4me3 and gene expression changes in aging leads to decreased glial support functions through extra-cellular matrix related pathways.
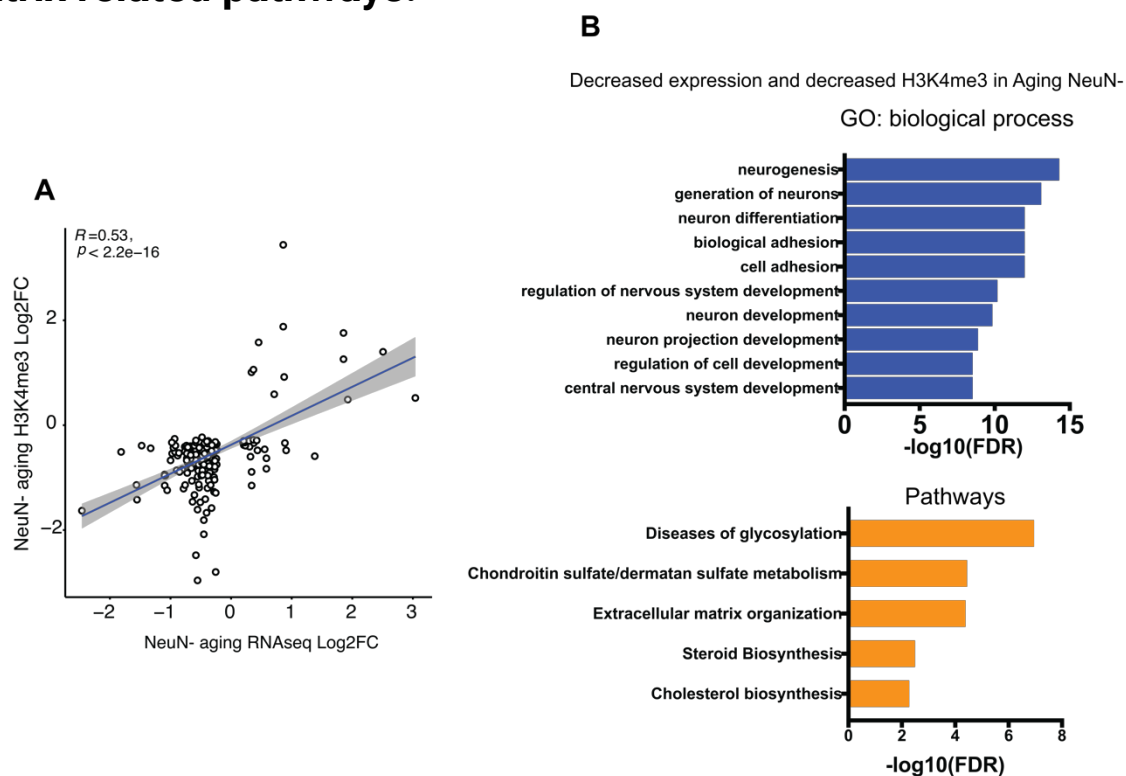

Figure 5: **Decreased H3K4me3 epigenome and downregulated transcriptome correlation in aged NeuN-**. (A) Pearson correlation for genes with log2 fold changes of H3K4me3 and RNAseq in aged NeuN- nuclei. Only FDR < 0.1 were takes for correlation in both datasets. Correlation coefficient and p value are given inside the figure. (B)GO biological processes and Reactome pathways for genes with decreased expression and decreased H3K4me3 mark.

As both epigenome and transcriptome of aged NeuN- showed high level of changes (mostly downregulation), we wanted to correlate them to understand whether epigenome changes were instructive of the transcriptome changes, leading to cellular dysfunctions. As decreased H3K4me3-only and H3K27me3-only class of epigenetic marks were the highest changes in magnitude (Figure 3**B**), we asked whether there was any positive correlation between the degree of log2 fold changes of histone marks with RNA expression changes in aged NeuN-. Although only 61 genes were common between changes in H3K27me3 and RNA expression, we found 260 genes with both H3K4me3 and RNA expression changes, out of 593 genes having deregulated RNA expression in aged NeuN- (Supplementary figure 6). Pearson correlation shows significant positive correlation (R= 0.53, p = 2.2 x 10^{-16}) meaning, majority of the downregulated RNAs in aged

NeuN- had decreased H3K4me3 mark (Figure 5**A**). Gene ontology of those decreased RNA and H3K4me3 genes in aged NeuN- showed neuron related pathways to be significant, although these genes were deregulated in NeuN- nuclei (Figure 5**B,** top). This would indicate, aged NeuN- have decreased gene expression and epigenome which leads to loss of important support function towards neurons. Pathway analysis showed these genes are related to critical glia-related metabolic processes, like glycosylation, chondroitine sulfate and cholesterol biosynthesis, that would affect the extra cellular matrix (Figure 5**B,** bottom), thereby leading to dysfunctional cellular environment and cognitive dysfunction. In summary, our data supports that age related H3K4me3 changes at TSS region lead to RNA expression changes in aged non-neuronal cells, without any H3K4me3 changes in neuronal cells, which might be a key mechanism for age related cognitive decline in mice aging.

DISCUSSION

Aging of the brain is a complex, unavoidable process that can lead to memory decline and eventually dementia. Until now, several studies have taken place to understand the overall gene expression changes in brain via the analysis of the aging hippocampus(Stilling, Benito et al. 2014) via RNA-seq or whole brain analysis using single cell RNA sequencing(Ximerakis, Lipnick et al. 2019). However, both of those studies have focused mostly on young vs. very old aged brains. It is already known that age-induced cognitive decline manifest already in middle-aged individuals(Schneider, Mangialasche et al. 2014) and in terms of mouse, they lose their hippocampus dependent spatial memory starting only at 16 months of age with very few gene expression changes detected via bulk tissue RNA sequencing(Peleg, Sananbenesi et al. 2010) .In order to better understand cell type specific transcriptome and epigenome changes in the hippocampal CA1 at the onset of age-associated memory decline, we have utilized nuclei FACS sorting methods to analyze hippocampal neuronal (NeuN+) and non-neuronal (NeuN-) epigenome and transcriptome of 16 months old mice.

A recent study form our lab employed the nuclei FACS sorting method to study NeuN+ and NeuN- DNA methylation and histone modifications during memory formation in young mice, more specifically, the authors asked how contextual fear conditioning changed the epigenome and transcriptome of plasticity related genes in the hippocampus and cortex(Halder, Hennion et al. 2016). However, this study lacked NeuN+ and NeuN- nuclear RNAseq data making it difficult to correlate gene-expression with epigenomic data. After testing several different nuclei isolation and sorting protocol, we have optimized the nuclei isolation and sorting protocol for this study to yield high quality NeuN+ and NeuN- nuclei for RNA sequencing. At first, we characterized the NeuN+ and NeuN- epigenetic marks by total RNA sequencing and also ChIP sequencing for H3K4me3 and H3K27me3 mark from neuronal and non-neuronal cell types. Comparison of the NeuN+ and NeuN- transcriptome revealed similar number of neuronal and non-neuronal genes (both protein coding and non-coding). Key marker genes for different neuronal and non-neuronal cells showed clear separation indicating our nuclear RNA sequencing data to be highly specific for those individual cell types. This gene list containing coding and non-coding RNAs can serve as a reference data set to identify neuronal and non-neuronal genes from bulk tissue RNAseq of the brain for future studies. The activatory H3K4me3 and inhibitory H3K27me3 data allowed us to study the

epigenome, along with bivalent chromatin regions of NeuN+ and NeuN- individually, which correlated with gene expression (see Figure 2). These epigenetic marks were occupied at the promoters of important pathways necessary (by activatory H3K4me3) or detrimental if expressed (repressed by H3K27me3) for both neuronal or non-neuronal cells. One crucial aspect was that NeuN+ overall had higher number of histone modifications compared to NeuN-, but the promoter occupancy was similar in number for both cell types. Further studies are required to investigate the function of these additional histone modifications in the NeuN+ nuclei specifically.

So far, many studies have been published looking at aging brain over all, but the specific transcriptional and epigenetic alteration on the verge of cognitive decline at 16 months of age in mice has not been studies. To find cure against age induced cognitive decline and age-related neurodegenerative diseases, it is of paramount importance to elucidate the underlying molecular mechanism. Therefore, with our cell type specific methods established, we investigated epigenome and transcriptome changes in 16 months mouse hippocampal CA1, compared to 3 months. Although 16 months NeuN+ neuronal cells showed modest changes in H3K27me3 both at its individual and bivalent marks, without any effect on H3K4me3, NeuN- non-neuronal nuclei showed drastic decrease in terms of both H3K4m3 and H3K27me3 marks. ChIP-seq profile plot at the TSS regions showed only H3k27me3 was decreased for NeuN+, where all the NeuN- marks were decreased. These data indicate, that in the aging brain neuronal and non-neuronal cells are differentially affected at the epigenome level. For decrease of H3K27me3 in NeuN+ at genes involved in cell fate commitment and developmental related genes, it indicates that neurons were losing their inhibitory epigenetic mark that would potentially de-repress non-neuronal genes and can further lead to neurodegeneration, similar phenomenon also being observed using PRC2 (Polycomb repressive complex 2, methyltransferase for H3K27me3) conditional neuron specific knockout mouse model in the brain(von Schimmelmann, Feinberg et al. 2016). As for NeuN- non-neuronal cells, since both the H3K4me3 and H3K27me3 marks decreased, we wanted to find out how this decrease in epigenetic code could potentially affect which glial cell types by turning on or off their genes (See Figure 3). By using a public data set of single cell sequencing of brain vasculature and other major glia types(Vanlandewijck, He et al. 2018), we found astrocytes and pericytes were the most significantly affected glial types to be losing their activatory histone marks, thereby potentially decrease in gene expression in later life.

In parallel, we also performed NeuN+ and NeuN- nuclear total RNA sequencing and it revealed very minor protein coding gene expression changes in NeuN+. Most of the genes deregulated in NeuN+ were unknown transcripts (TEC, see Figure 4) of which no function is known yet. It is tempting to speculate that these unknown transcripts are a result of to age induced somatic recombination(Verheijen, Vermulst et al. 2018) or increased transposon activity(Li, Prazak et al. 2013) that have been observed in the aging brain. Further studies will be required to test this hypothesis. In contrast to NeuN+, NeuN- had 454 deregulated protein coding genes and few non coding genes. For NeuN-, since majority of deregulated genes were protein coding, like our previous approach for ChIP-seq data, we wanted to find out which glial cell types were mostly affected. We found, in line with epigenetic data, that astrocytes and pericytes related genes were significantly downregulated, which are important cells related to neuronal support function. Microglia related genes were the most significantly upregulated, which have already been shown to mediate inflammatory signaling in the aging brain(von Bernhardi, Eugenin-von Bernhardi et al. 2015). Interestingly, we found several upregulated genes in 16 months NeuN- such as CD22, Neat1, which have been also been shown to be still upregulated at later time points, 20(Pluvinage, Haney et al. 2019) and 24(Butler, Johnston et al. 2019) months respectively. Knockdown approach for both of those genes have shown to reverse age-related cognitive decline. Interestingly, we found Neat1 to be upregulated only in NeuN- cells, not in neurons although the authors of the previous study suggested that the Neat1 aging related affects exclusively for neurons(Butler, Johnston et al. 2019). It will be very important to further characterize the source of this gene and potential non-neuronal cell related mechanisms in aspect of aging.

Finally, we correlated our epigenome and transcriptome datasets of aging brain and found age related H3K4me3 changes were significantly correlated with gene expression changes in NeuN- nuclei. This epigenome led transcriptome changes were mostly related to neuronal support function and extra cellular matrix related pathways. Therefore, we have discovered for the first time, age induced epigenetic changes leading to non-neuronal cell related dysfunctions, while neurons are moderately unchanged in terms of both histone modifications and transcriptome in 16 months mouse hippocampal CA1. This study paves the way to further study the cell-type specific data and provides a useful tool and valuable resource in terms of finding an effective cure against aging related cognitive dysfunction.

MATERIALS AND METHODS

Animals

C57BL/6J mice from Janvier were used for this study, according to their age group (3 or 16 months). The mice were ordered 2 weeks before reaching their target age and kept in standard home cages with food and water provided *ad libitum*. Mice were sacrificed by cervical dislocation, hippocampal CA1 tissue were dissected under microscope in ice cold PBS, flash frozen immediately in liquid nitrogen and stored at -80 °C until further processing. This was done due to avoiding any aberrant transcriptional changes happening due to sample processing. All experiments were performed according to the animal protection law of the state of Lower Saxony.

NeuN+/NeuN- nuclei specific RNA isolation

Frozen CA1 tissues from left hemispheres of each mouse were used for each biological replicate (n = 6 per age group) and processed on ice to maintain high RNA integrity. Frozen tissue was homogenized in 500 uL EZ prep lysis buffer (part of Sigma NUC101-1KT kit) using a plastic pestle in a 1.5mL Eppendorf tube with 45 strokes. Lysates were transferred into 2 mL DNA low binding tubes (Eppendorf), EZ prep lysis buffer was added up to 2 mL and incubated on ice for 7 minutes. It was then centrifuged for 5 minutes at 500xg, supernatant was removed and crude nuclear pellet was resuspended again into 2 mL lysis buffer and incubated on ice for 7 minutes. After centrifuging for 5 minutes at 500xg, supernatant was removed and nuclei pellet was resuspended into 1800ul nuclei storage buffer (NSB: RNAse-free PBS; Invitrogen, 0.5% RNase free BSA; Serva, 1:200 RNaseIN plus inhibitor; Promega, 1x EDTA-free protease inhibitor; Roche) and centrifuged again. Then the nuclear pellet was resuspended into 500µl NSB and filtered through 40 µm filter (BD falcon) with additional 100 µL NSB to collect residual nuclei from the filter. Nuclei were stained with anti-NeuN-Alexa488 conjugated antibody (1:1000) for 45 minutes at 4°C in dark and washed once with 1000µl NSB by centrifugation. Finally, the nuclei pellet was resuspended into 500µl NSB and FACS-sorted with FACSaria III using 85 µm nozzle. Nuclei were gated by their size, excluding doublets and neuronal nuclei were separated from non-neuronal nuclei by their NeuN-Alexa488 fluorescence signal. Sorted nuclei were collected into a 15 mL falcon tube precoated with 1mL NSB, spun down and RNA was isolated using Trizol LS. After addition of chloroform according to the

Trizol LS protocol, aqueous phase was collected and RNA was isolated by using Zymo RNA clean & concentrator-5 kit with DNAse treatment.

Total RNA sequencing

Since RNA amounts in NeuN- was very low, RNA 6000 Pico chip (Agilent) was used to measure concentrations of both NeuN+ and NeuN- nuclear RNA. 3ng RNA was used to prepare total RNAseq libraries using SMARTer® Stranded Total RNA-Seq Kit v2 -Pico Input kit (Takara), which incorporates ribosomal cDNA removal and illumina ready sequencing libraries. Libraries were sequenced using single-end 50 bp in HiSeq 2000.

NeuN+/NeuN- nuclei specific chromatin isolation

Frozen CA1 tissues from right hemispheres of three mice were used for each biological replicate (n = 4). Tissues were homogenized in low sucrose buffer (320mM Sucrose, 5mM CaCl2, 5mM MgAc2, 0.1mM EDTA, 10mM HEPES, 0.1% Triton X-100, 1mM DTT and ROche protease inhibitor) with plastic pestles, formaldehyde (1%) fixed for 10 minutes and quenched with 125mM glycine for 5 minutes. Lysates were further homogenized using T 10 basic ULTRA-TURRAX (with S10N-5G Dispersing tool). Debris were removed by sucrose gradient centrifugation for 3200xg for 10 minutes using high sucrose buffer (1000mM Sucrose, 3mM MgAc2, 10mM HEPES, 1mM DTT and Roche protease inhibitor). The resulting nuclear pellet was resuspended into PBTB (PBS+Tween20+BSA; 1% BSA, 0.2% Tween-20, protease inhibitor, in PBS buffer) and stained with anti-NeuN-Alexa488 conjugated antibody (1:1000) for 25 minutes and washed 3 times with PBS. Stained nuclei were then FACS sorted with FACSaria III using 85 μm nozzle. Nuclei were gated similarly as described previously(Halder, Hennion et al. 2016). Sorted nuclei were collected into a 15mL falcon tube and transferred into 1.5mL tubes. The nuclear pellet was flash frozen in liquid nitrogen and saved at -80°C for further processing.

Chromatin immunoprecipitation (ChIP)

For chromatin shearing, the pellet was resuspended into 100uL RIPA buffer (140mM NaCl, 1mM EDTA, 0.1% Sodium Deoxycholate, 1% Triton X-100, 10mM Tris pH 8, 1% SDS, protease inhibitor) and sonicated for 25 cycles in Diagenode bioruptor plus with high power (30 cycles on/ 30 cycles off). Chromatin shearing was checked by taking a small aliquot and decrosslinking the DNA by a speed decrosslinking method. In a single 1.5 mL tube, 2 ul chromatin, 93 ul water, 4 ul NaCl (5M), 8 ul, Proteinase K (20 mg/ml), 4 ul RNase (1 mg/ml) was taken and incubated at 65°C for 1 hour. DNA was isolated using Zymo ChIP DNA Clean & Concentrator kit. To make sure the Sheared chromatin size was

between 100-300bp, Agilent DNA high sensitivity kit was used in Bioanalyzer 2100 machine. Chromatin amount was measured with Qubit 2.0 fluorometer (DNA high sensitivity kit). 0.2µg chromatin with 1 µg of H3K4me3 antibody (Abcam ab8580) and 0.3ng chromatin with 2 µg H3K27me3 antibody (Millipore 17-622) were used to perform ChIP as previously described(Halder, Hennion et al. 2016). 2% input chromatin for ChIP without antibody was used as background control alongside the experiment.

Of note, to account for global chromatin mark changes, exogenous drosophila chromatin and drosophila specific antibody used in the samples as spike in controls (ChIP-seq spike in normalization kit, Active Motif). Since drosophila spike in reads obtained were lower than required, it was disregarded, however, the sequencing raw data will contain them. In downstream processing, those reads were disregarded.

ChIP-seq library preperation

The resulting ChIP DNA was subjected to library preparation using NEBNext Ultra II DNA library preparation kit. After end repair and adaptor ligation, an aliquot of libraries was used to perform test qPCR to determine optimal PCR cycles needed to avoid over amplification. Libraries were sequenced for single end 50 base pair reads into illumina HiSeq 2000 machine.

RNA-Seq Analysis

Base calling, fastq conversion, quality control, mapping of reads to mouse reference genome (mm10) were performed as described before (Kerimoglu, Sakib et al. 2017). Sequencing reads were counted using FeaturesCount (http://bioinf.wehi.edu.au/featureCounts/). Differential expression was analyzed with DESeq2 package of Bioconductor. Only genes that are expressed above basemean cutoff 5 and expressed in 50% of the samples between compared groups are kept for downstream analysis. RUVSeq package was used to remove unwanted variations from the datasets. FDR < 0.1 and fold change more or less than 20% were used for significantly differential genes. BiomartR package was used for finding out RNA biotypes from Ensembl annotation database. ToppFun (https://toppgene.cchmc.org/enrichment.jsp) was used for gene ontology and pathway analysis.

Calculating RPKM values for young samples:

For either NeuN+ or NeuN- raw RNAseq counts from young groups, it was converted to RPKM (Reads per kilo base per million mapped reads) counts which takes into account the gene lengths and library sizes (total number of raw reads), using following equation:

$$RPKM = \frac{Raw\ counts}{Total\ number\ of\ raw\ reads\ \times\ gene\ length} \times 10^6$$

Pseudo count of 1 was added to all RPKM counts to avoid conflicts with log2 transformation. This was used for subsequent ranking and characterization of promoter peaks in young NeuN+/NeuN- samples.

ChIP-Seq Analysis

Base calling and fastq conversion were performed using Illumina pipeline. Quality control was performed using fastqc. Reads were mapped to mm10 mouse reference genome with bowtie2. PCR duplicates were removed by "samtools *rmdup -s*" function.

Peak calling:

Peak calling was performed using MACS2. For a particular group, all the samples were pooled and peaks were called against corresponding input control, using q value < 0.1. For H3K4me3, narrow peaks and for H3K27me3, broad peaks were called.

Combining peaks associated with gene expression in young samples:

For both NeuN+ and NeuN- H3K4me3 and H3K27me3 ChIPseq pooled peak sets, they were individually annotated to find out nearby genes using "annotatePeaks.pl" function from HOMER tools. Genes within TSS +/- 2000bp of either H3K4me3 or H3K27me3 peaks were taken and overlapped with their corresponding cell-type specific genes with RPKM expression values. This gave 4 sets of genes: (A) genes with H3K4me3, (B) genes with H3K27me3, (C) genes with both H3K4me3 and H3K27me3, (D) genes with no marks. RPKM values for these four groups are plotted as box plots to find out gene expression distributions.

Determining H3K4me3 only, H3K27me3 only and Bivalent peaks for differential binding analysis:

Using the pooled peaksets from H3K4me3 and H3K27me3, "bedtools subtract -A " command was used to find unique peaks for H3K4me3 and H3K27me3, which is termed as "H3K4me3 only" or "H3K27me3 only" peaks. For getting bivalent regions, "bedtools intersect" command was used between H3K4m3 and H3K27me3 peaks for each

individual groups in young or old samples, (NeuN+ or NeuN- individually). These peaksets were used for differential binding analysis using "DiffBind" package, between young and old samples in four different groups in both NeuN+ and NeuN- samples seperately:

- H3K4me3 only Old vs. Young

- H3K27me3 only Old vs. Young

- Bivalent H3K4me3 Old vs Young

- Bivalent H3K27me3 Old vs Young

Differential peaks were annotated using HOMER package. A peak within TSS +/- 2500bp of a gene was associated being promoter peak. IGV was used to create genome browser plots for both ChIPseq and RNAseq. "ngsplot" was used to create ChIP seq profile plot and heatmap around TSS regions.
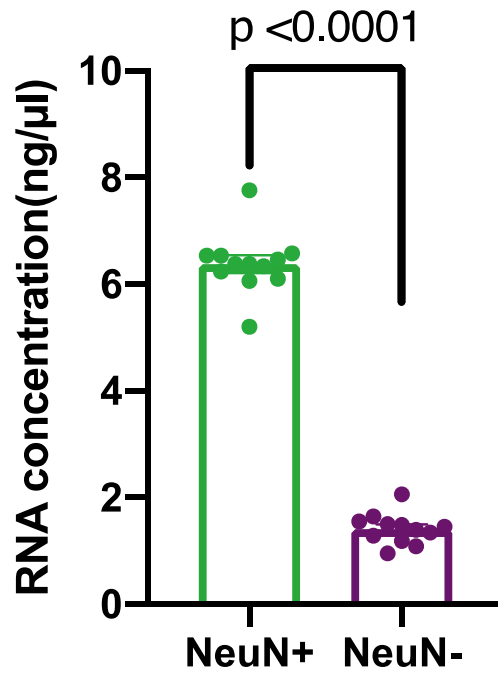
Statistical tests

All statistical tests were performed using either Graphpad prism 8 or in R.

Data availability

All RNA and ChIP-seq datasets will be made available via GEO database.

SUPPLEMENTAL DATA

Supplementary figure 1: FACS sorted NeuN+ and NeuN- RNA concentrations. Student's t test, p value <0.0001.
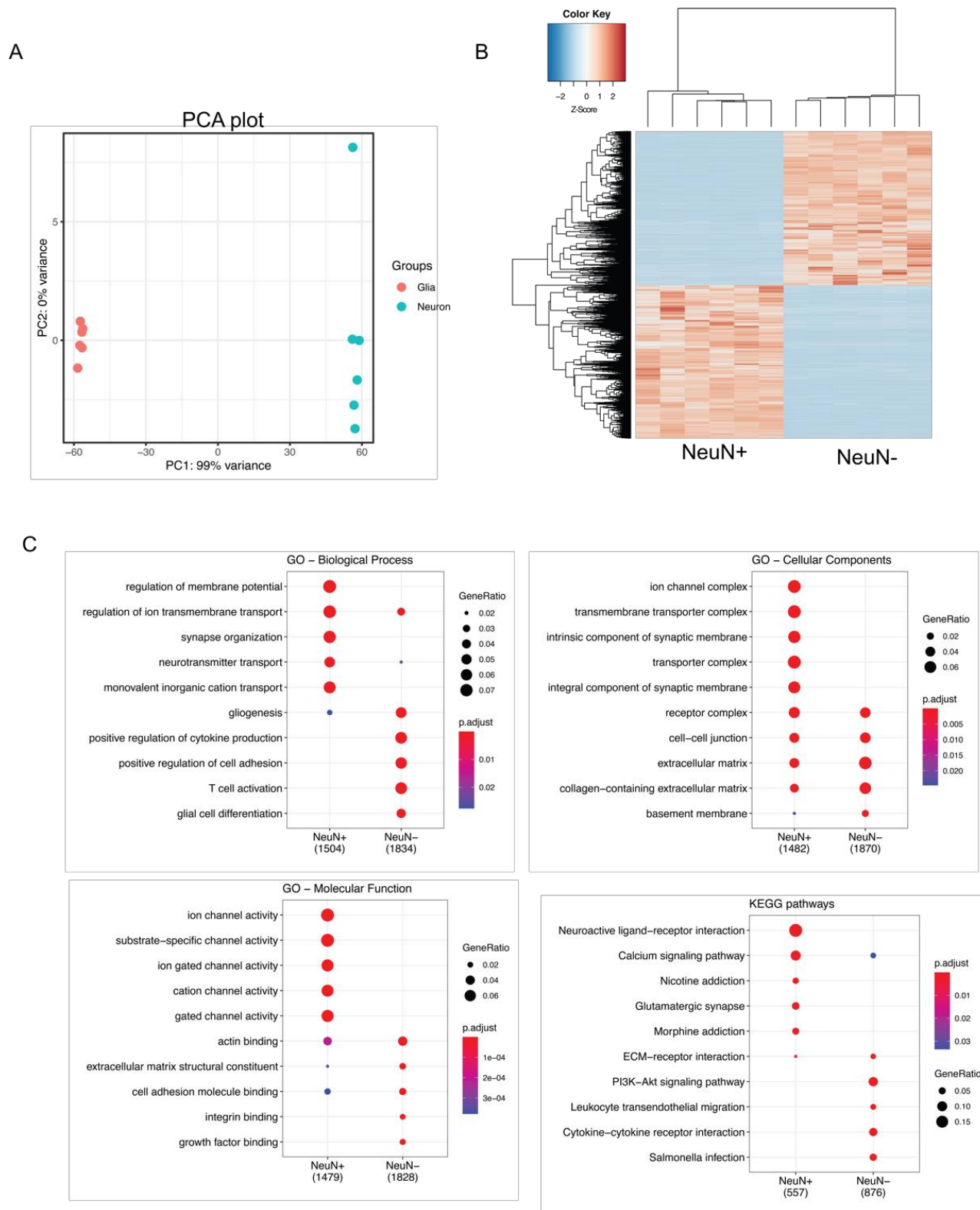
Supplementary figure 2: (A) PCA plot for NeuN+ an NeuN- RNAseq data. (B) heatmap of significant differentially expressed genes between NeuN+ and NeuN- (FDR<0.05, fold change 5). (C) Gene ontology and KEGG pathway analysis for NeuN+ and NeuN- enriched genes (FDR<0.05, fold change 5).

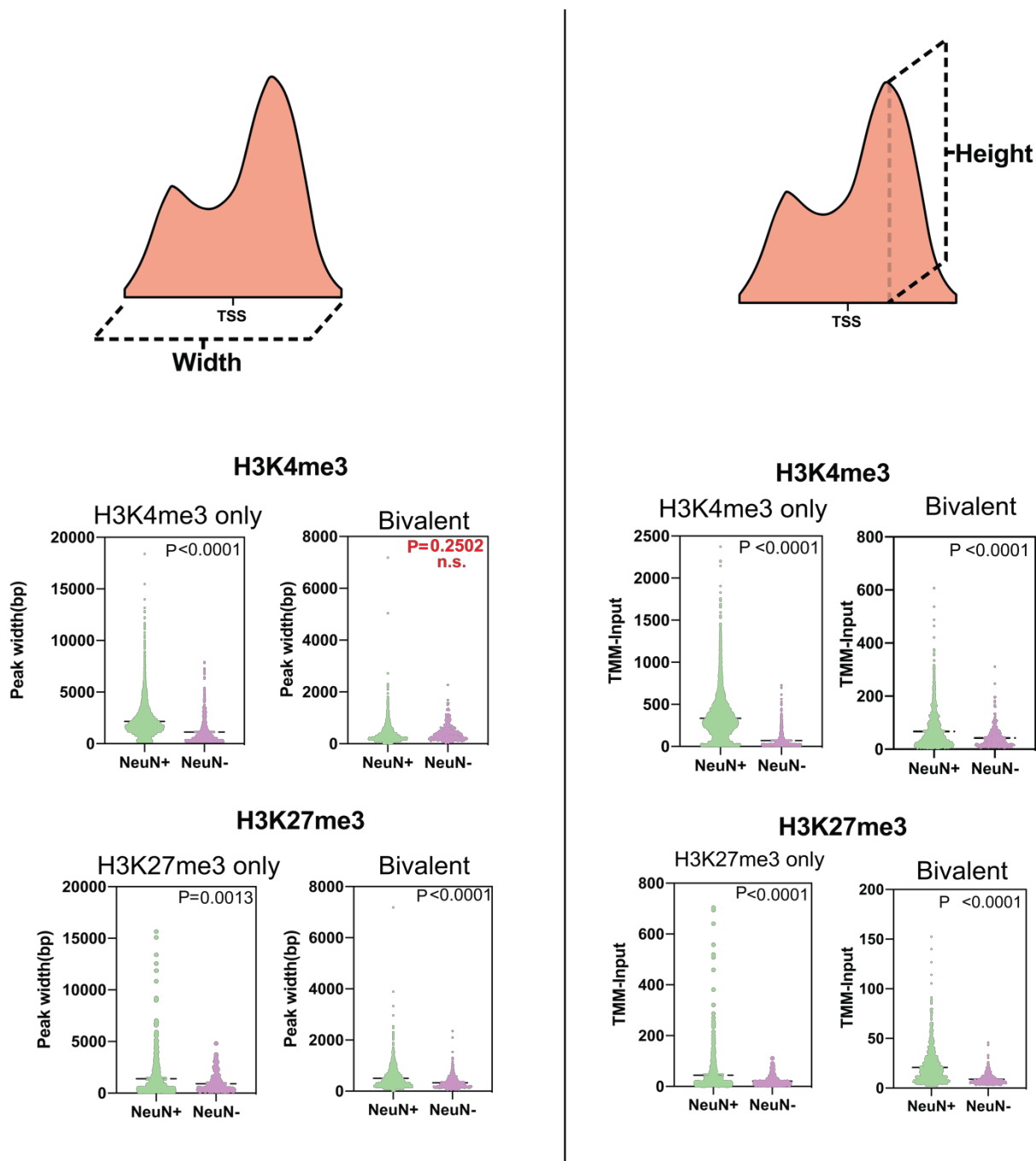**Supplementary figure 3**: Peak height and width differences between NeuN+ and NeuN-epigenetic marks. Left panel shows the differences between width and right panel shows the differences between height of peaks between NeuN+ and NeuN-. For H3K4me3 or H3K27me3, peaks were subdivided into occupancy for only that mark or in bivalent regions. All p values were obtained by student's t test, significance threshold p < 0.05. Only H3K4me3 peaks in bivalent regions showed no difference between NeuN+ and NeuN-.
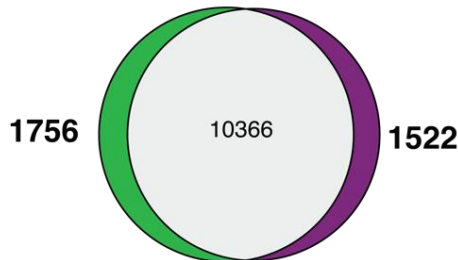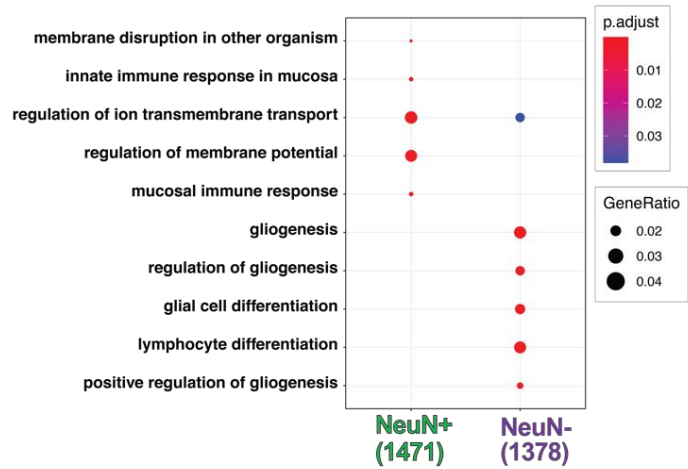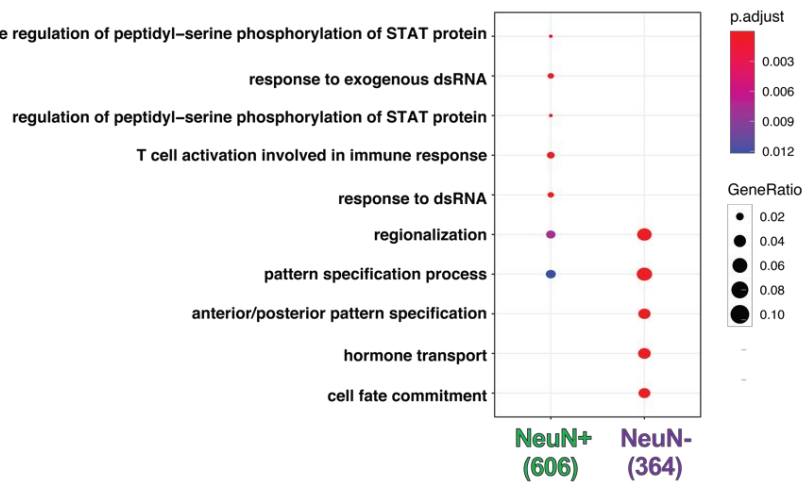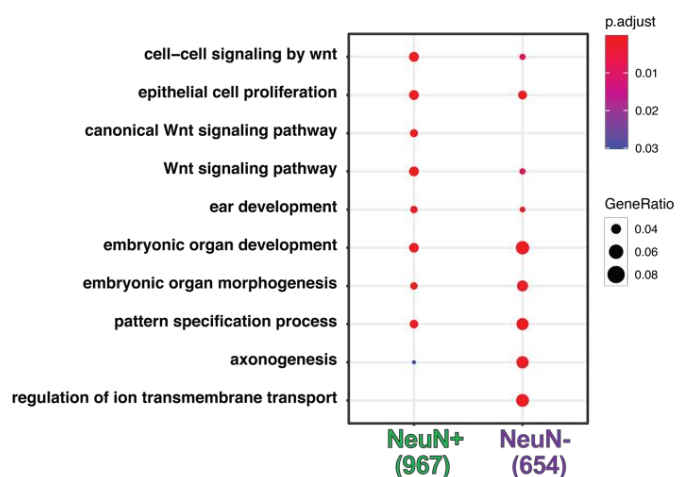
Supplementary figure 4: Gene ontology analysis for genes bearing unique H3K4me3, unique H3K27me3 or unique bivalent regions in NeuN+ and NeuN- nuclei.

-log10(FDR)
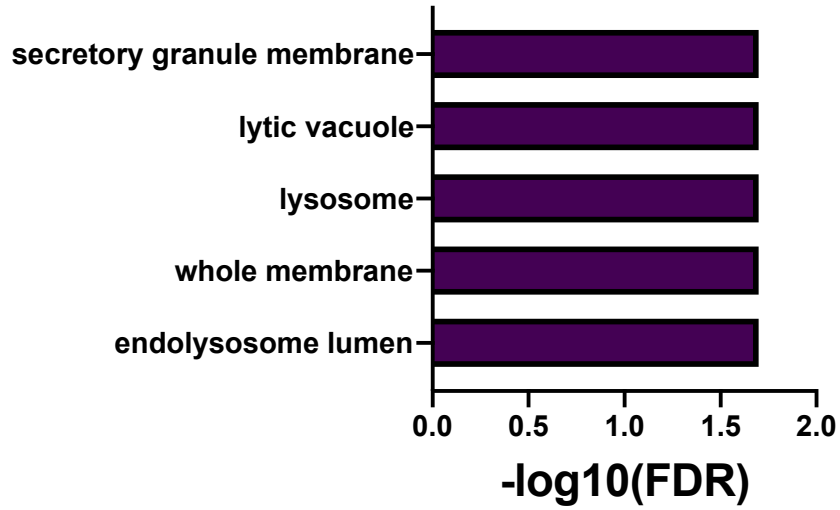
Supplementary figure 5: GO: cellular components analysis for microglia related genes with decreased H3K4me3 marks
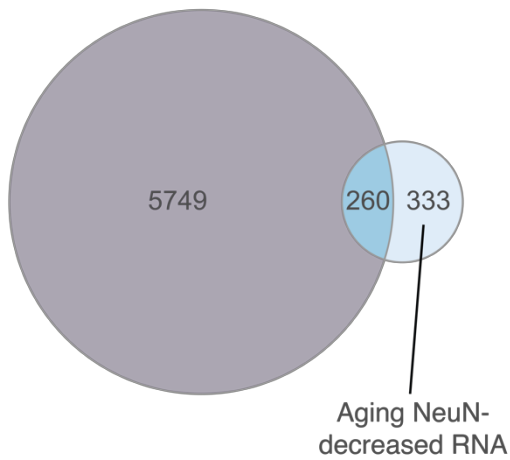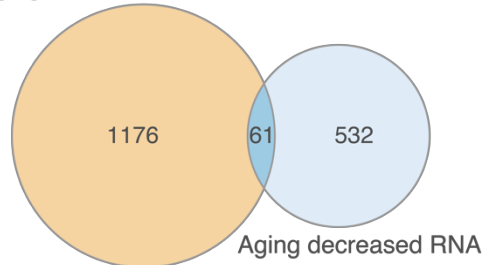
Supplementary figure 6: Venn diagram overlap for decreased NeuN- H3K4me3 and H3K27me3 and decreased NeuN- RNA expression in aging.

REFERENCES

Akkers, R. C., S. J. van Heeringen, U. G. Jacobi, E. M. Janssen-Megens, K. J. Francoijs, H. G. Stunnenberg and G. J. Veenstra (2009). "A hierarchy of H3K4me3 and H3K27me3 acquisition in spatial gene regulation in Xenopus embryos." Dev Cell **17**(3): 425-434.

Benayoun, B. A., E. A. Pollina, P. P. Singh, S. Mahmoudi, I. Harel, K. M. Casey, B. W. Dulken, A. Kundaje and A. Brunet (2019). "Remodeling of epigenome and transcriptome landscapes with aging in mice reveals widespread induction of inflammatory responses." Genome Res **29**(4): 697-709.

Blalock, E. M., K. C. Chen, K. Sharrow, J. P. Herman, N. M. Porter, T. C. Foster and P. W. Landfield (2003). "Gene microarrays in hippocampal aging: statistical profiling identifies novel processes correlated with cognitive impairment." J Neurosci **23**(9): 3807-3819.

Butler, A. A., D. R. Johnston, S. Kaur and F. D. Lubin (2019). "Long noncoding RNA NEAT1 mediates neuronal histone methylation and age-related memory impairment." Sci Signal **12**(588).

Cao, T., S. Rajasingh, S. Samanta, B. Dawn, D. C. Bittel and J. Rajasingh (2018). "Biology and clinical relevance of noncoding sno/scaRNAs." Trends Cardiovasc Med **28**(2): 81-90.

Cruickshanks, H. A., T. McBryan, D. M. Nelson, N. D. Vanderkraats, P. P. Shah, J. van Tuyn, T. Singh Rai, C. Brock, G. Donahue, D. S. Dunican, M. E. Drotar, R. R. Meehan, J. R. Edwards, S. L. Berger and P. D. Adams (2013). "Senescent cells harbour features of the cancer epigenome." Nat Cell Biol **15**(12): 1495-1506.

Ecker, J. R., D. H. Geschwind, A. R. Kriegstein, J. Ngai, P. Osten, D. Polioudakis, A. Regev, N. Sestan, I. R. Wickersham and H. Zeng (2017). "The BRAIN Initiative Cell Census Consortium: Lessons Learned toward Generating a Comprehensive Brain Cell Atlas." Neuron **96**(3): 542-557.

Gong, H., H. Qian, R. Ertl, C. M. Astle, G. G. Wang, D. E. Harrison and X. Xu (2015). "Histone modifications change with age, dietary restriction and rapamycin treatment in mouse brain." Oncotarget **6**(18): 15882-15890.

Halder, R., M. Hennion, R. O. Vidal, O. Shomroni, R. U. Rahman, A. Rajput, T. P. Centeno, F. van Bebber, V. Capece, J. C. Garcia Vizcaino, A. L. Schuetz, S. Burkhardt, E. Benito, M. Navarro Sala, S. B. Javan, C. Haass, B. Schmid, A. Fischer and S. Bonn (2016). "DNA methylation changes in plasticity genes accompany the formation and maintenance of memory." Nat Neurosci **19**(1): 102-110.

Ianov, L., M. De Both, M. K. Chawla, A. Rani, A. J. Kennedy, I. Piras, J. J. Day, A. Siniard, A. Kumar, J. D. Sweatt, C. A. Barnes, M. J. Huentelman and T. C. Foster (2017). "Hippocampal Transcriptomic Profiles: Subfield Vulnerability to Age and Cognitive Impairment." Front Aging Neurosci **9**: 383.

Ianov, L., A. Rani, B. S. Beas, A. Kumar and T. C. Foster (2016). "Transcription Profile of Aging and Cognition-Related Genes in the Medial Prefrontal Cortex." Front Aging Neurosci **8**: 113.

Kerimoglu, C., M. S. Sakib, G. Jain, E. Benito, S. Burkhardt, V. Capece, L. Kaurani, R. Halder, R. C. Agís-Balboa, R. Stilling, H. Urbanke, A. Kranz, A. F. Stewart and A. Fischer (2017). "KMT2A and KMT2B Mediate Memory Function by Affecting Distinct Genomic Regions." <u>Cell reports</u> **20**(3): 538-548.

Kozlenkov, A., P. Roussos, A. Timashpolsky, M. Barbu, S. Rudchenko, M. Bibikova, B. Klotzle, W. Byne, R. Lyddon, A. F. Di Narzo, Y. L. Hurd, E. V. Koonin and S. Dracheva (2014). "Differences in DNA methylation between human neuronal and glial cells are concentrated in enhancers and non-CpG sites." <u>Nucleic Acids Res</u> **42**(1): 109-127.

Lake, B. B., R. Ai, G. E. Kaeser, N. S. Salathia, Y. C. Yung, R. Liu, A. Wildberg, D. Gao, H. L. Fung, S. Chen, R. Vijayaraghavan, J. Wong, A. Chen, X. Sheng, F. Kaper, R. Shen, M. Ronaghi, J. B. Fan, W. Wang, J. Chun and K. Zhang (2016). "Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain." <u>Science</u> **352**(6293): 1586-1590.

Li, W., L. Prazak, N. Chatterjee, S. Gruninger, L. Krug, D. Theodorou and J. Dubnau (2013). "Activation of transposable elements during aging and neuronal decline in Drosophila." <u>Nat Neurosci</u> **16**(5): 529-531.

Murman, D. L. (2015). "The Impact of Age on Cognition." <u>Semin Hear</u> **36**(3): 111-121.

Nativio, R., G. Donahue, A. Berson, Y. Lan, A. Amlie-Wolf, F. Tuzer, J. B. Toledo, S. J. Gosai, B. D. Gregory, C. Torres, J. Q. Trojanowski, L. S. Wang, F. B. Johnson, N. M. Bonini and S. L. Berger (2018). "Dysregulation of the epigenetic landscape of normal aging in Alzheimer's disease." <u>Nat Neurosci</u> **21**(4): 497-505.

Nativio, R., Y. Lan, G. Donahue, S. Sidoli, A. Berson, A. R. Srinivasan, O. Shcherbakova, A. Amlie-Wolf, J. Nie, X. Cui, C. He, L. S. Wang, B. A. Garcia, J. Q. Trojanowski, N. M. Bonini and S. L. Berger (2020). "An integrated multi-omics approach identifies epigenetic alterations associated with Alzheimer's disease." <u>Nat Genet</u> **52**(10): 1024-1035.

O'Sullivan, R. J., S. Kubicek, S. L. Schreiber and J. Karlseder (2010). "Reduced histone biosynthesis and chromatin changes arising from a damage signal at telomeres." <u>Nat Struct Mol Biol</u> **17**(10): 1218-1225.

Peleg, S., F. Sananbenesi, A. Zovoilis, S. Burkhardt, S. Bahari-Javan, R. C. Agis-Balboa, P. Cota, J. L. Wittnam, A. Gogol-Doering, L. Opitz, G. Salinas-Riester, M. Dettenhofer, H. Kang, L. Farinelli, W. Chen and A. Fischer (2010). "Altered histone acetylation is associated with age-dependent memory impairment in mice." <u>Science</u> **328**(5979): 753-756.

Pluvinage, J. V., M. S. Haney, B. A. H. Smith, J. Sun, T. Iram, L. Bonanno, L. Li, D. P. Lee, D. W. Morgens, A. C. Yang, S. R. Shuken, D. Gate, M. Scott, P. Khatri, J. Luo, C. R. Bertozzi, M. C. Bassik and T. Wyss-Coray (2019). "CD22 blockade restores homeostatic microglial phagocytosis in ageing brains." <u>Nature</u> **568**(7751): 187-192.

Prolla, T. A. (2002). "DNA microarray analysis of the aging brain." <u>Chem Senses</u> **27**(3): 299-306.

Salech, F., D. P. Ponce, A. C. Paula-Lima, C. D. SanMartin and M. I. Behrens (2020). "Nicotinamide, a Poly [ADP-Ribose] Polymerase 1 (PARP-1) Inhibitor, as an Adjunctive Therapy for the Treatment of Alzheimer's Disease." Front Aging Neurosci **12**: 255.

Schneider, L. S., F. Mangialasche, N. Andreasen, H. Feldman, E. Giacobini, R. Jones, V. Mantua, P. Mecocci, L. Pani, B. Winblad and M. Kivipelto (2014). "Clinical trials and late-stage drug development for Alzheimer's disease: an appraisal from 1984 to 2014." J Intern Med **275**(3): 251-283.

Shah, P. P., G. Donahue, G. L. Otte, B. C. Capell, D. M. Nelson, K. Cao, V. Aggarwala, H. A. Cruickshanks, T. S. Rai, T. McBryan, B. D. Gregory, P. D. Adams and S. L. Berger (2013). "Lamin B1 depletion in senescent cells triggers large-scale changes in gene expression and the chromatin landscape." Genes Dev **27**(16): 1787-1799.

Simen, A. A., K. A. Bordner, M. P. Martin, L. A. Moy and L. C. Barry (2011). "Cognitive dysfunction with aging and the role of inflammation." Ther Adv Chronic Dis **2**(3): 175-195.

Stilling, R. M., E. Benito, M. Gertig, J. Barth, V. Capece, S. Burkhardt, S. Bonn and A. Fischer (2014). "De-regulation of gene expression and alternative splicing affects distinct cellular pathways in the aging hippocampus." Front Cell Neurosci **8**: 373.

Thibault, O., J. C. Gant and P. W. Landfield (2007). "Expansion of the calcium hypothesis of brain aging and Alzheimer's disease: minding the store." Aging Cell **6**(3): 307-317.

Vanlandewijck, M., L. He, M. A. Mae, J. Andrae, K. Ando, F. Del Gaudio, K. Nahar, T. Lebouvier, B. Lavina, L. Gouveia, Y. Sun, E. Raschperger, M. Rasanen, Y. Zarb, N. Mochizuki, A. Keller, U. Lendahl and C. Betsholtz (2018). "A molecular atlas of cell types and zonation in the brain vasculature." Nature **554**(7693): 475-480.

Verheijen, B. M., M. Vermulst and F. W. van Leeuwen (2018). "Somatic mutations in neurons during aging and neurodegeneration." Acta Neuropathol **135**(6): 811-826.

von Bernhardi, R., L. Eugenin-von Bernhardi and J. Eugenin (2015). "Microglial cell dysregulation in brain aging and neurodegeneration." Front Aging Neurosci **7**: 124.

von Schimmelmann, M., P. A. Feinberg, J. M. Sullivan, S. M. Ku, A. Badimon, M. K. Duff, Z. Wang, A. Lachmann, S. Dewell, A. Ma'ayan, M. H. Han, A. Tarakhovsky and A. Schaefer (2016). "Polycomb repressive complex 2 (PRC2) silences genes responsible for neurodegeneration." Nat Neurosci **19**(10): 1321-1330.

Wendeln, A. C., K. Degenhardt, L. Kaurani, M. Gertig, T. Ulas, G. Jain, J. Wagner, L. M. Hasler, K. Wild, A. Skodras, T. Blank, O. Staszewski, M. Datta, T. P. Centeno, V. Capece, M. R. Islam, C. Kerimoglu, M. Staufenbiel, J. L. Schultze, M. Beyer, M. Prinz, M. Jucker, A. Fischer and J. J. Neher (2018). "Innate immune memory in the brain shapes neurological disease hallmarks." Nature **556**(7701): 332-338.

Ximerakis, M., S. L. Lipnick, B. T. Innes, S. K. Simmons, X. Adiconis, D. Dionne, B. A. Mayweather, L. Nguyen, Z. Niziolek, C. Ozek, V. L. Butty, R. Isserlin, S. M. Buchanan, S. S. Levine, A. Regev, G. D. Bader, J. Z. Levin and L. L. Rubin (2019). "Single-cell transcriptomic profiling of the aging mouse brain." Nat Neurosci **22**(10): 1696-1708.

Yankner, B. A., T. Lu and P. Loerch (2008). "The aging brain." <u>Annu Rev Pathol</u> **3**: 41-66.

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTION

MSS and AF designed the experiments. MSS performed cell type specific epigenome and transcriptome raw data generation, including protocol optimization, sample preparation and library preparation. MSS performed ChIP-seq and RNA-seq data analysis. CK, MRI, LK gave valuable suggestions regarding data analysis. AM, TB helped with mouse brain dissection. SB, DMK, TPC performed high throughput sequencing and raw data generation. MSS and AF wrote the manuscript.

COMPETING INTERESTS
The authors declare no competing interests

General Discussion

6. General Discussion

Why did I perform nuclei based genomic analysis from the brain?

The brain is one of the most complex organs, composed of distinct cell types that are functionally different based on their specific patterns of gene expression or epigenomic signature(Graff, Kim et al. 2011). Capturing transcriptome or epigenome from brain tissues is becoming common practice to discover the molecular underpinning of gene regulatory networks in brain development and diseases, but this comes with a caveat. Given that not all cell types response equally in healthy or in a diseased situation, changes in epigenetic marks or gene expression in a small subset or rare cell types might be undetected as it contributes to a small fraction of the whole tissue. Moreover, if a given gene is regulated in opposing direction in different cells, the cumulative expression pattern might appear static in whole brain tissue RNA sequencing. Therefore, direct measurement and analysis of epigenome or transcriptome form particular cell types of the brain are critical to obtain thorough insights about the systems biology of a disease or homeostasis. While several methods exist for cell type specific genome wide analysis from the brain(Kim, Lim et al. 2015), i.e. Laser Capture Microdissection (LCM) or more of a transgene approach such as Translating Ribosome Affinity Purification (TRAP also known as RiboTag), these have several limitations. As for LCM, it is a low throughput method, and only selected cells of interest can be harvested. As for the TRAP method, it involves raising mice lines with particular genotypes and later, purifying RNA by an affinity-based approach that only captures a certain species of RNA molecules being translated. For brain tissue, it is very difficult to isolate whole cells given the intricate nature of neurons, its dendrites and axons. Although proteolytic dispersion and FACS based cell sorting has been utilized to generate cell type specific data from the brain (Darmanis, Sloan et al. 2015, Zeisel, Munoz-Manchado et al. 2015), this method has been shown to put cells under stress, leading to altered gene expression (Huang, Hsing et al. 2010, Lacar, Linker et al. 2016). Moreover, it is impossible to perform cell dissociation from archived frozen brain tissues from the patients which limits the possibility of generating an enormous amount of disease associated cell type specific epigenomic or transcriptomic data. To avoid harsh treatments of cell preparation from the brain tissue and for utilizing frozen brain tissues for genome wide epigenomic and transcriptomic

analysis, nuclei isolation methods are becoming the norm for cell type specific analysis of the brain (Habib, Li et al. 2016, Habib, Avraham-Davidi et al. 2017, Bakken, Hodge et al. 2018, Del-Aguila, Li et al. 2019, Hodge, Bakken et al. 2019).

In this thesis, I have further developed and utilized brain nuclei isolation for subsequent staining methods for high throughput brain cell type isolation via FACS and next generation sequencing from ultra-low input RNA or chromatin materials. While chromatin modifications have been shown to be crucial for neurodevelopment(D'Mello 2019, Salinas, Connolly et al. 2020), learning-memory formation(Graff and Tsai 2013, Sweatt 2013, Halder, Hennion et al. 2016), and diseased brain(Coneys and Wood 2020, Lee, Lee et al. 2020, Myrum, Kittleson et al. 2020, Nativio, Lan et al. 2020, Zhang, Qu et al. 2020), cell type specific genomic analysis is undoubtedly important to deepen our understanding and discover novel mechanisms. Developing this nuclei based technique has enabled me to study diverse topics in the field of neuroepigenetics, starting from finding the epigenetic basic of human brain folding, to discover mechanistic details via profiling transcriptome and epigenome of a neuron specific knockout mice as a model for memory impairments, and finally to describe cell type specific epigenetic and transcriptomic deregulation of aging brain which could potentially be useful to design novel therapy against dementia.

Epigenetic mechanism of brain folding

The unique cognitive and behavioral abilities of mammals is governed by having 6 layered neocortex, which is most of the cerebral cortex of the brain (Geschwind and Rakic 2013, Kaas 2013). While many rodents contain lissencephalic brains with no gyration, the human brain is gyrated which occurred from an incredible number of neuronal populations arising from neural stem cells (NSC) and neural progenitor cells (NPC) during embryogenesis(Lui, Hansen et al. 2011, Sun and Hevner 2014). Among the two main types of NPCs (the apical or basal progenitors), the basal progenitor cells (BP) derives the basal intermediate progenitors (bIPs) that ultimately leads to cortical neuron formation. Apart from this indirect pathway, cortical neurons are produced directly via apical progenitor cells (Radial glia cells) as well (Lui, Hansen et al. 2011, Borrell and Gotz 2014, Taverna, Gotz et al. 2014).

In the first study, by utilizing BP specific antibody (Tbr2), we determined that human BPs have a higher amount of histone acetylation, compared to mouse BPs. Now, what cellular and molecular changes would occur if the acetylation levels were increased in mouse brain? To test the hypothesis, E12.5 mice were injected TSA (HDAC inhibitor) to promote histone acetylation and their brain tested after 4-6 days. Since BAF155 cKO mice were shown to have a higher number of BP(Narayanan, Pham et al. 2018), the effect of TSA was also tested alongside WT mice. Firstly, immunohistochemistry experiments from the HDAC inhibitor treated mice brain showed that BPs were increasing in numbers (manuscript 1, *Figure 4*) and mouse brain started to have cortical folding (manuscript 1, *Figure 6*). To find out the exact molecular mechanism, I performed FACS sorting of Tbr2+ and Tbr2- BP nuclei from control and HDAC inhibitor treated mice and did RNAseq and ChIPseq of H3K9ac. Data analysis revealed that this increased H3 acetylation was causative for upregulation of key transcription factor gene, Trnp1, which is regulating BP proliferation (Stahl, Walcher et al. 2013, Martinez-Martinez, De Juan Romero et al. 2016). In contrast, this was not observed for Tbr2- nuclei. As a validation experiment, epigenome editing by increasing acetylation using inactivated Cas9, fused with a histone acetyltransferase (Kat2a) led to specific increase in acetylation at Trnp1 promoter, thereby increasing it's gene expression has also shown similar increase in BPs. Therefore, our cell type specific analysis of the BPs from HDAC inhibitor treated embryonic mouse brain has identified the exact epigenetic and transcriptomic mechanism of how increasing histone acetylation in BPs led to up regulation of key transcription factor responsible for gyrification in lissencephalic rodent brain.

Setd1B regulates neuronal gene transcription, thereby memory formation

While in the previous study, I focused on studying cell type specific histone acetylation, in the second manuscript, I investigated histone methylation and its role in memory formation. As introduced before, histone 3 lysine 4 trimethylation (H3K4me3) is generated by SET-domain containing lysine methyltransferases and Setd1B is one of them. In humans, mutations or loss of function has been linked with intellectual disability, and autism (Shulha, Cheung et al. 2012, Labonne, Lee et al. 2016, Hiraide, Nakashima et al. 2018). While those might indicate developmental related roles of Setd1B, in this

161

present study, we used conditional knockout model mice where Setd1B was removed only from mature excitatory neurons at the forebrain, including hippocampus. As expected, loss of Setd1B from the neurons led to hippocampus dependent spatial memory of those transgenic mice. Since it was neuron specific knockout, I isolated neuron specific nuclei and investigated four epigenetic marks including H3K4me3 and transcriptome between WT and cKO mice. Moreover, I performed single nuclei RNA sequencing from WT mouse brain to measure the expression patterns of Setd1B, along with other histone methyltransferases in WT situation and implicated its controlling neuron specific gene expression.

While Setd1B cKO neurons showed significantly decreased H3K4me3 at TSS (transcription start sites), surprisingly, we observed these deregulated peaks were unique in comparison to other two histone methyltransferases cKO (Kmt2a and Kmt2b) studied by our group. Interestingly, Setd1B regulated H3K4me3 peaks were the broadest peaks and genes related to those peaks are related to memory formation and neuronal functions, compared to Kmt2a and Kmt2b cKO mice. This is in line with the recent publication stating that learning and memory formation induces hippocampal genes with broadest H3K4me3 peaks(Collins, Sweatt et al. 2019). Our findings also corroborated with the water maze performance among Setd1B, Kmt2a and Kmt2b mice, where Setd1B showed the most drastic changes. I then performed single nuclei RNA sequencing of sorted neurons from hippocampal CA region in wildtype mice to characterize basal expression of Setd1b, Kmt2a and Kmt2b in neurons. Overall comparison stated that even though Setd1b is expressed among a subset of neurons, it mostly harbored neuronal functions and memory process related gene expression programs, compared to neurons expressing the other two histone methyltransferases. All together, we propose that Setd1b regulates key neuronal networks specific for learning and memory formation by controlling H3K4me3 peak width at key neuronal genes and their expression.

Aging leads to massive deregulation of histone methylation and concomitant glial dysfunction higher than neurons

While in the previous study I investigated the role of a histone methyltransferase in learning and memory, in the 3rd manuscript, I explored histone methylation in detail in aspect of age associated cognitive decline and corresponding gene expression. While previous attempts to map the epigenome and transcriptome of aging mouse brain has been taken(Stefanelli, Azam et al. 2018, Benayoun, Pollina et al. 2019), it has to be noted that none of the studies takes into consideration of either cell-type or hippocampal subregion specificity. Moreover, they profiled combination of different epigenetic marks than the histone marks studied in this manuscript. As discussed in manuscript 3, age related cognitive decline in mice starts around 16 months of age, with little gene expression changes seen before(Peleg, Sananbenesi et al. 2010). In an attempt to specifically decipher the promoter epigenome and transcriptomic dysfunction of aging brain related cognitive decline in a cell type specific manner, I characterized activatory H3K4me3 mark, inhibitory H3K27me3 mark and gene expression profiles in neuronal and non-neuronal nuclei in 16-month-old mice hippocampal CA1, compared to 3-month-old mice.

In this study, I first characterized neuronal and non-neuronal RNA species and provide important genes specifically expressed in those cell types. This will serve as a very important resource to further characterize all these coding and non-coding genes enriched in neurons or non-neuronal cells. As neuronal and non-neuronal gene promoter epigenomic mark and their corresponding gene expression has not been correlated yet, I performed this in young samples first and showed in both cell types, genes bearing activatory H3K4me3, bivalent region consisting both marks and inhibitory H3K27me3 showed stepwise increasing expression levels, with H3K4me3 to be the highest and H3K27me3 to be the lowest. It is also crucial to note that bivalent or H3K27me3 marks were enriched in pathways that are repressed in both cell types, in agreement with the previous publication stating H3K27me3 insulates unnecessary gene transcription programs to maintain normal function of a given cell(von Schimmelmann, Feinberg et al. 2016).

In 16 months old hippocampal CA1, we observed massive decrease in both H3K4me3 and H3K27me3 marks at the promoters of genes related to diverse glial types. As for neurons, only H3K27me3 was observed to be downregulated at the promoters. These cell type specific changes corroborated previous findings where authors profiled DNA methylation in neuronal and non-neuronal cells upon aging and Alzheimer's disease of the human brain (Gasparoni, Bultmann et al. 2018). Depletion of activatory H3K4me3 mark in non-neuronal cells overlapped significantly with Astrocytes and pericytes, two of the main cell types important for neuronal support function, while neuronal inhibitory H3K27me3 depletion showed pathways related to neuronal cell fate and differentiation, which shows some agreement with the previous studies showing neuronal aneuploidy in aging and age related diseases (Rosenkrantz and Carbone 2017, Caneus, Granic et al. 2018, Shepherd, Yang et al. 2018, Yurov, Vorsanova et al. 2018, Potter, Chial et al. 2019), that in turn leads to neurodegeneration. Transcriptomic analysis of neuronal and non-neuronal nuclei shows non-neuronal cells to be more affected in expressing protein coding genes, compared to neurons. This detection was only possible as we profiled neuronal and non-neuronal nuclei specifically from aging hippocampal CA1. Altogether, this data set provides novel gene expression and epigenetic changes in aging hippocampus, that can be utilized to find novel drug against dementia

.

Appendix

General References

7. Appendix

A password protected folder containing all the supplementary files for the three manuscripts in this thesis is accessible at the following link:

Password: **sakibphdthesis2020**

[Click here to visit the folder online](#)

Alternatively, please scan the QR code below to access the folder as well.

8. General References

Abruzzi, K., X. Chen, E. Nagoshi, A. Zadina and M. Rosbash (2015). "RNA-seq profiling of small numbers of Drosophila neurons." Methods Enzymol **551**: 369-386.

Aguilar-Valles, A., T. Vaissiere, E. M. Griggs, M. A. Mikaelsson, I. F. Takacs, E. J. Young, G. Rumbaugh and C. A. Miller (2014). "Methamphetamine-associated memory is regulated by a writer and an eraser of permissive histone methylation." Biol Psychiatry **76**(1): 57-65.

Albert, M. S. and M. B. Moss (1988). Geriatric neuropsychology, The Guilford Press.

Allfrey, V. G., R. Faulkner and A. E. Mirsky (1964). "Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis." Proc Natl Acad Sci U S A **51**: 786-794.

Allis, C. D., S. L. Berger, J. Cote, S. Dent, T. Jenuwien, T. Kouzarides, L. Pillus, D. Reinberg, Y. Shi, R. Shiekhattar, A. Shilatifard, J. Workman and Y. Zhang (2007). "New nomenclature for chromatin-modifying enzymes." Cell **131**(4): 633-636.

Bakken, T. E., R. D. Hodge, J. A. Miller, Z. Yao, T. N. Nguyen, B. Aevermann, E. Barkan, D. Bertagnolli, T. Casper, N. Dee, E. Garren, J. Goldy, L. T. Graybuck, M. Kroll, R. S. Lasken, K. Lathia, S. Parry, C. Rimorin, R. H. Scheuermann, N. J. Schork, S. I. Shehata, M. Tieu, J. W. Phillips, A. Bernard, K. A. Smith, H. Zeng, E. S. Lein and B. Tasic (2018). "Single-nucleus and single-cell transcriptomes compared in matched cortical cell types." PLoS One **13**(12): e0209648.

Bannister, A. J. and T. Kouzarides (2011). "Regulation of chromatin by histone modifications." Cell Res **21**(3): 381-395.

Benayoun, B. A., E. A. Pollina, P. P. Singh, S. Mahmoudi, I. Harel, K. M. Casey, B. W. Dulken, A. Kundaje and A. Brunet (2019). "Remodeling of epigenome and transcriptome landscapes with aging in mice reveals widespread induction of inflammatory responses." Genome Res **29**(4): 697-709.

Benito, E., H. Urbanke, B. Ramachandran, J. Barth, R. Halder, A. Awasthi, G. Jain, V. Capece, S. Burkhardt, M. Navarro-Sala, S. Nagarajan, A. L. Schutz, S. A. Johnsen, S. Bonn, R. Luhrmann, C. Dean and A. Fischer (2015). "HDAC inhibitor-dependent transcriptome and memory reinstatement in cognitive decline models." J Clin Invest **125**(9): 3572-3584.

Bernstein, B. E., T. S. Mikkelsen, X. Xie, M. Kamal, D. J. Huebert, J. Cuff, B. Fry, A. Meissner, M. Wernig, K. Plath, R. Jaenisch, A. Wagschal, R. Feil, S. L. Schreiber and E. S. Lander (2006). "A bivalent chromatin structure marks key developmental genes in embryonic stem cells." Cell **125**(2): 315-326.

Bhasin, M., E. L. Reinherz and P. A. Reche (2006). "Recognition and classification of histones using support vector machine." J Comput Biol **13**(1): 102-112.

Blalock, E. M., K. C. Chen, K. Sharrow, J. P. Herman, N. M. Porter, T. C. Foster and P. W. Landfield (2003). "Gene microarrays in hippocampal aging: statistical profiling identifies novel processes correlated with cognitive impairment." J Neurosci **23**(9): 3807-3819.

Blanco, E., M. Gonzalez-Ramirez, A. Alcaine-Colet, S. Aranda and L. Di Croce (2020). "The Bivalent Genome: Characterization, Structure, and Regulation." Trends Genet **36**(2): 118-131.

Borrell, V. and M. Gotz (2014). "Role of radial glial cells in cerebral cortex folding." <u>Current Opinion in Neurobiology</u> **27**: 39-46.

Campbell, G. L., M. Schanchner and S. O. Sharrow (1977). "Isolation of glial cell-enriched and -depleted populations from mouse cerebellum by density gradient centrifugation and electronic cell sorting." <u>Brain Res</u> **127**(1): 69-86.

Caneus, J., A. Granic, R. Rademakers, D. W. Dickson, C. M. Coughlan, H. J. Chial and H. Potter (2018). "Mitotic defects lead to neuronal aneuploidy and apoptosis in frontotemporal lobar degeneration caused by MAPT mutations." <u>Mol Biol Cell</u> **29**(5): 575-586.

Cappello, S., A. Attardo, X. Wu, T. Iwasato, S. Itohara, M. Wilsch-Brauninger, H. M. Eilken, M. A. Rieger, T. T. Schroeder, W. B. Huttner, C. Brakebusch and M. Gotz (2006). "The Rho-GTPase cdc42 regulates neural progenitor fate at the apical surface." <u>Nat Neurosci</u> **9**(9): 1099-1107.

Chen, L., J. Melendez, K. Campbell, C. Y. Kuan and Y. Zheng (2009). "Rac1 deficiency in the forebrain results in neural progenitor reduction and microcephaly." <u>Dev Biol</u> **325**(1): 162-170.

Chen, Z., S. Li, S. Subramaniam, J. Y. Shyy and S. Chien (2017). "Epigenetic Regulation: A New Frontier for Biomedical Engineers." <u>Annu Rev Biomed Eng</u> **19**: 195-219.

Collins, B. E., C. B. Greer, B. C. Coleman and J. D. Sweatt (2019). "Histone H3 lysine K4 methylation and its role in learning and memory." <u>Epigenetics Chromatin</u> **12**(1): 7.

Collins, B. E., J. D. Sweatt and C. B. Greer (2019). "Broad domains of histone 3 lysine 4 trimethylation are associated with transcriptional activation in CA1 neurons of the hippocampus during memory formation." <u>Neurobiol Learn Mem.</u> **5**(161): 149-157.

Coneys, R. and I. C. Wood (2020). "Alzheimer's disease: the potential of epigenetic treatments and current clinical candidates." <u>Neurodegener Dis Manag</u> **10**(3): 543-558.

D'Mello, S. R. (2019). "Regulation of Central Nervous System Development by Class I Histone Deacetylases." <u>Dev Neurosci</u> **41**(3-4): 149-165.

Darmanis, S., S. A. Sloan, Y. Zhang, M. Enge, C. Caneda, L. M. Shuer, M. G. Hayden Gephart, B. A. Barres and S. R. Quake (2015). "A survey of human brain transcriptome diversity at the single cell level." <u>Proc Natl Acad Sci U S A</u> **112**(23): 7285-7290.

De Pietri Tonelli, D., J. N. Pulvers, C. Haffner, E. P. Murchison, G. J. Hannon and W. B. Huttner (2008). "miRNAs are essential for survival and differentiation of newborn neurons but not for expansion of neural progenitors during early neurogenesis in the mouse embryonic neocortex." <u>Development</u> **135**(23): 3911-3921.

Del-Aguila, J. L., Z. Li, U. Dube, K. A. Mihindukulasuriya, J. P. Budde, M. V. Fernandez, L. Ibanez, J. Bradley, F. Wang, K. Bergmann, R. Davenport, J. C. Morris, D. M. Holtzman, R. J. Perrin, B. A. Benitez, J. Dougherty, C. Cruchaga and O. Harari (2019). "A single-nuclei RNA sequencing study of Mendelian and sporadic AD in the human brain." <u>Alzheimers Res Ther</u> **11**(1): 71.

Dincer, A., D. P. Gavin, K. Xu, B. Zhang, J. T. Dudley, E. E. Schadt and S. Akbarian (2015). "Deciphering H3K4me3 broad domains associated with gene-regulatory networks and conserved epigenomic landscapes in the human brain." <u>Transl Psychiatry</u> **5**: e679.

Dong, S., M. F. Walker, N. J. Carriero, M. DiCola, A. J. Willsey, A. Y. Ye, Z. Waqar, L. E. Gonzalez, J. D. Overton, S. Frahm, J. F. Keaney, 3rd, N. A. Teran, J. Dea, J. D. Mandell, V. Hus Bal, C. A. Sullivan, N. M. DiLullo, R. O. Khalil, J. Gockley, Z. Yuksel, S. M. Sertel, A. G. Ercan-Sencicek, A. R. Gupta, S. M. Mane, M. Sheldon, A. I. Brooks, K. Roeder, B. Devlin, M. W. State, L. Wei and S. J. Sanders (2014). "De novo insertions and deletions of predominantly paternal origin are associated with autism spectrum disorder." Cell Rep **9**(1): 16-23.

Estivill-Torrus, G., H. Pearson, V. van Heyningen, D. J. Price and P. Rashbass (2002). "Pax6 is required to regulate the cell cycle and the rate of progression from symmetrical to asymmetrical division in mammalian cortical progenitors." Development **129**(2): 455-466.

Farrelly, L. A., R. E. Thompson, S. Zhao, A. E. Lepack, Y. Lyu, N. V. Bhanu, B. Zhang, Y. E. Loh, A. Ramakrishnan, K. C. Vadodaria, K. J. Heard, G. Erikson, T. Nakadai, R. M. Bastle, B. J. Lukasak, H. Zebroski, 3rd, N. Alenina, M. Bader, O. Berton, R. G. Roeder, H. Molina, F. H. Gage, L. Shen, B. A. Garcia, H. Li, T. W. Muir and I. Maze (2019). "Histone serotonylation is a permissive modification that enhances TFIID binding to H3K4me3." Nature **567**(7749): 535-539.

Fischer, A. (2014). "Epigenetic memory: the Lamarckian brain." EMBO J **33**(9): 945-967.

Fornasiero, E. F., S. Mandad, H. Wildhagen, M. Alevra, B. Rammner, S. Keihani, F. Opazo, I. Urban, T. Ischebeck, M. S. Sakib, M. K. Fard, K. Kirli, T. P. Centeno, R. O. Vidal, R. U. Rahman, E. Benito, A. Fischer, S. Dennerlein, P. Rehling, I. Feussner, S. Bonn, M. Simons, H. Urlaub and S. O. Rizzoli (2018). "Precisely measured protein lifetimes in the mouse brain reveal differences across tissues and subcellular fractions." Nat Commun **9**(1): 4230.

Fullard, J. F., C. Giambartolomei, M. E. Hauberg, K. Xu, G. Voloudakis, Z. Shao, C. Bare, J. T. Dudley, M. Mattheisen, N. K. Robakis, V. Haroutunian and P. Roussos (2017). "Open chromatin profiling of human postmortem brain infers functional roles for non-coding schizophrenia loci." Hum Mol Genet **26**(10): 1942-1951.

Gasparoni, G., S. Bultmann, P. Lutsik, T. F. J. Kraus, S. Sordon, J. Vlcek, V. Dietinger, M. Steinmaurer, M. Haider, C. B. Mulholland, T. Arzberger, S. Roeber, M. Riemenschneider, H. A. Kretzschmar, A. Giese, H. Leonhardt and J. Walter (2018). "DNA methylation analysis on purified neurons and glia dissects age and Alzheimer's disease-specific changes in the human cortex." Epigenetics Chromatin **11**(1): 41.

Geisler, S. J. and R. Paro (2015). "Trithorax and Polycomb group-dependent regulation: a tale of opposing activities." Development **142**(17): 2876-2887.

Geschwind, D. H. and P. Rakic (2013). "Cortical evolution: judge the brain by its cover." Neuron **80**(3): 633-647.

Glickstein, S. B., J. A. Monaghan, H. B. Koeller, T. K. Jones and M. E. Ross (2009). "Cyclin D2 is critical for intermediate progenitor cell proliferation in the embryonic cortex." J Neurosci **29**(30): 9614-9624.

Gong, H., H. Qian, R. Ertl, C. M. Astle, G. G. Wang, D. E. Harrison and X. Xu (2015). "Histone modifications change with age, dietary restriction and rapamycin treatment in mouse brain." Oncotarget **6**(18): 15882-15890.

Graff, J., D. Kim, M. M. Dobbin and L. H. Tsai (2011). "Epigenetic regulation of gene expression in physiological and pathological brain processes." Physiol Rev **91**(2): 603-649.

Graff, J. and L. H. Tsai (2013). "Histone acetylation: molecular mnemonics on the chromatin." Nat Rev Neurosci **14**(2): 97-111.

Grindberg, R. V., J. L. Yee-Greenbaum, M. J. McConnell, M. Novotny, A. L. O'Shaughnessy, G. M. Lambert, M. J. Arauzo-Bravo, J. Lee, M. Fishman, G. E. Robbins, X. Lin, P. Venepally, J. H. Badger, D. W. Galbraith, F. H. Gage and R. S. Lasken (2013). "RNA-sequencing from single nuclei." Proc Natl Acad Sci U S A **110**(49): 19802-19807.

Gruber, R., Z. Zhou, M. Sukchev, T. Joerss, P. O. Frappart and Z. Q. Wang (2011). "MCPH1 regulates the neuroprogenitor division mode by coupling the centrosomal cycle with mitotic entry through the Chk1-Cdc25 pathway." Nat Cell Biol **13**(11): 1325-1334.

Gupta, S., S. Y. Kim, S. Artis, D. L. Molfese, A. Schumacher, J. D. Sweatt, R. E. Paylor and F. D. Lubin (2010). "Histone methylation regulates memory formation." J Neurosci **30**(10): 3589-3599.

Habib, N., I. Avraham-Davidi, A. Basu, T. Burks, K. Shekhar, M. Hofree, S. R. Choudhury, F. Aguet, E. Gelfand, K. Ardlie, D. A. Weitz, O. Rozenblatt-Rosen, F. Zhang and A. Regev (2017). "Massively parallel single-nucleus RNA-seq with DroNc-seq." Nat Methods **14**(10): 955-958.

Habib, N., Y. Li, M. Heidenreich, L. Swiech, I. Avraham-Davidi, J. J. Trombetta, C. Hession, F. Zhang and A. Regev (2016). "Div-Seq: Single-nucleus RNA-Seq reveals dynamics of rare adult newborn neurons." Science **353**(6302): 925-928.

Halder, R., M. Hennion, R. O. Vidal, O. Shomroni, R. U. Rahman, A. Rajput, T. P. Centeno, F. van Bebber, V. Capece, J. C. Garcia Vizcaino, A. L. Schuetz, S. Burkhardt, E. Benito, M. Navarro Sala, S. B. Javan, C. Haass, B. Schmid, A. Fischer and S. Bonn (2016). "DNA methylation changes in plasticity genes accompany the formation and maintenance of memory." Nat Neurosci **19**(1): 102-110.

Halder, R., M. Hennion, R. O. Vidal, O. Shomroni, R. U. Rahman, A. Rajput, T. P. Centeno, F. van Bebber, V. Capece, J. C. Vizcaino, A. L. Schuetz, S. Burkhardt, E. Benito, M. N. Sala, S. B. Javan, C. Haass, B. Schmid, A. Fischer and S. Bonn (2016). "DNA methylation changes in plasticity genes accompany the formation and maintenance of memory." Nat Neurosci **19**(1): 102-110.

Harikumar, A. and E. Meshorer (2015). "Chromatin remodeling and bivalent histone modifications in embryonic stem cells." EMBO Rep **16**(12): 1609-1619.

Haydar, T. F., C. Y. Kuan, R. A. Flavell and P. Rakic (1999). "The role of cell death in regulating the size and shape of the mammalian forebrain." Cereb Cortex **9**(6): 621-626.

Hedden, T. and J. D. Gabrieli (2004). "Insights into the ageing mind: a view from cognitive neuroscience." Nat Rev Neurosci **5**(2): 87-96.

Hiraide, T., M. Nakashima, K. Yamoto, T. Fukuda, M. Kato, H. Ikeda, Y. Sugie, K. Aoto, T. Kaname, K. Nakabayashi, T. Ogata, N. Matsumoto and H. Saitsu (2018). "De novo variants in SETD1B are associated with intellectual disability, epilepsy and autism." Hum Genet. **137**(1): 95-104.

Hodge, R. D., T. E. Bakken, J. A. Miller, K. A. Smith, E. R. Barkan, L. T. Graybuck, J. L. Close, B. Long, N. Johansen, O. Penn, Z. Yao, J. Eggermont, T. Hollt, B. P. Levi, S. I. Shehata, B.

Aevermann, A. Beller, D. Bertagnolli, K. Brouner, T. Casper, C. Cobbs, R. Dalley, N. Dee, S. L. Ding, R. G. Ellenbogen, O. Fong, E. Garren, J. Goldy, R. P. Gwinn, D. Hirschstein, C. D. Keene, M. Keshk, A. L. Ko, K. Lathia, A. Mahfouz, Z. Maltzer, M. McGraw, T. N. Nguyen, J. Nyhus, J. G. Ojemann, A. Oldre, S. Parry, S. Reynolds, C. Rimorin, N. V. Shapovalova, S. Somasundaram, A. Szafer, E. R. Thomsen, M. Tieu, G. Quon, R. H. Scheuermann, R. Yuste, S. M. Sunkin, B. Lelieveldt, D. Feng, L. Ng, A. Bernard, M. Hawrylycz, J. W. Phillips, B. Tasic, H. Zeng, A. R. Jones, C. Koch and E. S. Lein (2019). "Conserved cell types with divergent features in human versus mouse cortex." Nature **573**(7772): 61-68.

Hong, J., H. Zhang, Y. Kawase-Koga and T. Sun (2013). "MicroRNA function is required for neurite outgrowth of mature neurons in the mouse postnatal cerebral cortex." Front Cell Neurosci **7**: 151.

Huang, H. L., H. W. Hsing, T. C. Lai, Y. W. Chen, T. R. Lee, H. T. Chan, P. C. Lyu, C. L. Wu, Y. C. Lu, S. T. Lin, C. W. Lin, C. H. Lai, H. T. Chang, H. C. Chou and H. L. Chan (2010). "Trypsin-induced proteome alteration during cell subculture in mammalian cells." J Biomed Sci **17**: 36.

Hyun, K., J. Jeon, K. Park and J. Kim (2017). "Writing, erasing and reading histone lysine methylations." Exp Mol Med **49**(4): e324.

Ianov, L., M. De Both, M. K. Chawla, A. Rani, A. J. Kennedy, I. Piras, J. J. Day, A. Siniard, A. Kumar, J. D. Sweatt, C. A. Barnes, M. J. Huentelman and T. C. Foster (2017). "Hippocampal Transcriptomic Profiles: Subfield Vulnerability to Age and Cognitive Impairment." Front Aging Neurosci **9**: 383.

Ianov, L., A. Rani, B. S. Beas, A. Kumar and T. C. Foster (2016). "Transcription Profile of Aging and Cognition-Related Genes in the Medial Prefrontal Cortex." Front Aging Neurosci **8**: 113.

Jakovcevski, M., H. Ruan, E. Y. Shen, A. Dincer, B. Javidfar, Q. Ma, C. J. Peter, I. Cheung, A. C. Mitchell, Y. Jiang, C. L. Lin, V. Pothula, A. F. Stewart, P. Ernst, W. D. Yao and S. Akbarian (2015). "Neuronal Kmt2a/Mll1 histone methyltransferase is essential for prefrontal synaptic plasticity and working memory." J Neurosci **35**(13): 5097-5108.

Jensen, L. R., M. Amende, U. Gurok, B. Moser, V. Gimmel, A. Tzschach, A. R. Janecke, G. Tariverdian, J. Chelly, J. P. Fryns, H. Van Esch, T. Kleefstra, B. Hamel, C. Moraine, J. Gecz, G. Turner, R. Reinhardt, V. M. Kalscheuer, H. H. Ropers and S. Lenzner (2005). "Mutations in the JARID1C gene, which is involved in transcriptional regulation and chromatin remodeling, cause X-linked mental retardation." Am J Hum Genet **76**(2): 227-236.

Jiang, Y., A. Matevossian, H. S. Huang, J. Straubhaar and S. Akbarian (2008). "Isolation of neuronal chromatin from brain tissue." BMC Neurosci **9**: 42.

Jones, M. J., S. J. Goodman and M. S. Kobor (2015). "DNA methylation and healthy human aging." Aging Cell **14**(6): 924-932.

Jones, P. A. (2012). "Functions of DNA methylation: islands, start sites, gene bodies and beyond." Nat Rev Genet **13**(7): 484-492.

Jones, W. D., D. Dafou, M. McEntagart, W. J. Woollard, F. V. Elmslie, M. Holder-Espinasse, M. Irving, A. K. Saggar, S. Smithson, R. C. Trembath, C. Deshpande and M. A. Simpson (2012). "De novo mutations in MLL cause Wiedemann-Steiner syndrome." Am J Hum Genet **91**(2): 358-364.

Kaas, J. H. (2013). "The evolution of brains from early mammals to humans." Wiley Interdiscip Rev Cogn Sci **4**(1): 33-45.

Karlic, R., H. R. Chung, J. Lasserre, K. Vlahovicek and M. Vingron (2010). "Histone modification levels are predictive for gene expression." Proc Natl Acad Sci U S A **107**(7): 2926-2931.

Kelava, I., E. Lewitus and W. B. Huttner (2013). "The secondary loss of gyrencephaly as an example of evolutionary phenotypical reversal." Front Neuroanat **7**: 16.

Kerimoglu, C., R. C. Agis-Balboa, A. Kranz, R. Stilling, S. Bahari-Javan, E. Benito-Garagorri, R. Halder, S. Burkhardt, A. F. Stewart and A. Fischer (2013). "Histone-methyltransferase MLL2 (KMT2B) is required for memory formation in mice." J Neurosci **33**(8): 3452-3464.

Kerimoglu, C., M. S. Sakib, G. Jain, E. Benito, S. Burkhardt, V. Capece, L. Kaurani, R. Halder, R. C. Agis-Balboa, R. Stilling, H. Urbanke, A. Kranz, A. F. Stewart and A. Fischer (2017). "KMT2A and KMT2B Mediate Memory Function by Affecting Distinct Genomic Regions." Cell Rep **20**(3): 538-548.

Kim, M. and J. Costello (2017). "DNA methylation: an epigenetic mark of cellular memory." Exp Mol Med **49**(4): e322.

Kim, T., C. S. Lim and B. K. Kaang (2015). "Cell type-specific gene expression profiling in brain tissue: comparison between TRAP, LCM and RNA-seq." BMB Rep **48**(7): 388-394.

Kriegstein, A., S. Noctor and V. Martinez-Cerdeno (2006). "Patterns of neural stem and progenitor cell division may underlie evolutionary cortical expansion." Nat Rev Neurosci **7**(11): 883-890.

Krishnaswami, S. R., R. V. Grindberg, M. Novotny, P. Venepally, B. Lacar, K. Bhutani, S. B. Linker, S. Pham, J. A. Erwin, J. A. Miller, R. Hodge, J. K. McCarthy, M. Kelder, J. McCorrison, B. D. Aevermann, F. D. Fuertes, R. H. Scheuermann, J. Lee, E. S. Lein, N. Schork, M. J. McConnell, F. H. Gage and R. S. Lasken (2016). "Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons." Nat Protoc **11**(3): 499-524.

Labonne, J. D., K. H. Lee, S. Iwase, I. K. Kong, M. P. Diamond, L. C. Layman, C. H. Kim and H. G. Kim (2016). "An atypical 12q24.31 microdeletion implicates six genes including a histone demethylase KDM2B and a histone methyltransferase SETD1B in syndromic intellectual disability." Hum Genet. **135**(7): 757-771.

Lacar, B., S. B. Linker, B. N. Jaeger, S. R. Krishnaswami, J. J. Barron, M. J. E. Kelder, S. L. Parylak, A. C. M. Paquola, P. Venepally, M. Novotny, C. O'Connor, C. Fitzpatrick, J. A. Erwin, J. Y. Hsu, D. Husband, M. J. McConnell, R. Lasken and F. H. Gage (2016). "Nuclear RNA-seq of single neurons reveals molecular signatures of activation." Nat Commun **7**: 11022.

Lange, C., W. B. Huttner and F. Calegari (2009). "Cdk4/cyclinD1 overexpression in neural stem cells shortens G1, delays neurogenesis, and promotes the generation and expansion of basal progenitors." Cell Stem Cell **5**(3): 320-331.

Lee, M. Y., J. Lee, S. J. Hyeon, H. Cho, Y. J. Hwang, J. Y. Shin, A. C. McKee, N. W. Kowall, J. I. Kim, T. D. Stein, D. Hwang and H. Ryu (2020). "Epigenome signatures landscaped by histone H3K9me3 are associated with the synaptic dysfunction in Alzheimer's disease." Aging Cell **19**(6): e13153.

Lepack, A. E., C. T. Werner, A. F. Stewart, S. L. Fulton, P. Zhong, L. A. Farrelly, A. C. W. Smith, A. Ramakrishnan, Y. Lyu, R. M. Bastle, J. A. Martin, S. Mitra, R. M. O'Connor, Z. J. Wang, H. Molina, G. Turecki, L. Shen, Z. Yan, E. S. Calipari, D. M. Dietz, P. J. Kenny and I.

Maze (2020). "Dopaminylation of histone H3 in ventral tegmental area regulates cocaine seeking." Science **368**(6487): 197-201.

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

Marsit, C. J. (2015). "Influence of environmental exposure on human epigenetic regulation." J Exp Biol **218**(Pt 1): 71-79.

Martinez-Martinez, M. A., C. De Juan Romero, V. Fernandez, A. Cardenas, M. Gotz and V. Borrell (2016). "A restricted period for formation of outer subventricular zone defined by Cdh1 and Trnp1 levels." Nature communications **7**: 11812.

Meng, S., H. Zhou, Z. Feng, Z. Xu, Y. Tang and M. Wu (2019). "Epigenetics in Neurodevelopment: Emerging Role of Circular RNA." Front Cell Neurosci **13**: 327.

Mi, D., C. B. Carr, P. A. Georgala, Y. T. Huang, M. N. Manuel, E. Jeanes, E. Niisato, S. N. Sansom, F. J. Livesey, T. Theil, K. Hasenpusch-Theil, T. I. Simpson, J. O. Mason and D. J. Price (2013). "Pax6 exerts regional control of cortical progenitor proliferation via direct repression of Cdk6 and hypophosphorylation of pRb." Neuron **78**(2): 269-284.

Mongan, N. P., R. D. Emes and N. Archer (2019). "Detection and analysis of RNA methylation." F1000Res **8**.

Mullen, R. J., C. R. Buck and A. M. Smith (1992). "NeuN, a neuronal specific nuclear protein in vertebrates." Development **116**(1): 201-211.

Mutch, C. A., J. D. Schulte, E. Olson and A. Chenn (2010). "Beta-catenin signaling negatively regulates intermediate progenitor population numbers in the developing cortex." PLoS One **5**(8): e12376.

Myrum, C., J. Kittleson, S. De, B. R. Fletcher, J. Castellano, G. Kundu, K. G. Becker and P. R. Rapp (2020). "Survey of the Arc Epigenetic Landscape in Normal Cognitive Aging." Mol Neurobiol **57**(6): 2727-2740.

Narayanan, R., L. Pham, C. Kerimoglu, T. Watanabe, R. Castro Hernandez, G. Sokpor, P. A. Ulmke, K. A. Kiszka, A. B. Tonchev, J. Rosenbusch, R. H. Seong, U. Teichmann, J. Frahm, A. Fischer, S. Bonn, A. Stoykova, J. F. Staiger and T. Tuoc (2018). "Chromatin Remodeling BAF155 Subunit Regulates the Genesis of Basal Progenitors in Developing Cortex." iScience **4**: 109-126.

Nativio, R., Y. Lan, G. Donahue, S. Sidoli, A. Berson, A. R. Srinivasan, O. Shcherbakova, A. Amlie-Wolf, J. Nie, X. Cui, C. He, L. S. Wang, B. A. Garcia, J. Q. Trojanowski, N. M. Bonini and S. L. Berger (2020). "An integrated multi-omics approach identifies epigenetic alterations associated with Alzheimer's disease." Nat Genet **52**(10): 1024-1035.

Ng, S. B., A. W. Bigham, K. J. Buckingham, M. C. Hannibal, M. J. McMillin, H. I. Gildersleeve, A. E. Beck, H. K. Tabor, G. M. Cooper, H. C. Mefford, C. Lee, E. H. Turner, J. D. Smith, M. J. Rieder, K. Yoshiura, N. Matsumoto, T. Ohta, N. Niikawa, D. A. Nickerson, M. J. Bamshad and J. Shendure (2010). "Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome." Nat Genet **42**(9): 790-793.

Nishioka, M., T. Shimada, M. Bundo, W. Ukai, E. Hashimoto, T. Saito, Y. Kano, T. Sasaki, K. Kasai, T. Kato and K. Iwamoto (2013). "Neuronal cell-type specific DNA methylation patterns of the Cacna1c gene." Int J Dev Neurosci **31**(2): 89-95.

Okaty, B. W., M. E. Freret, B. D. Rood, R. D. Brust, M. L. Hennessy, D. deBairos, J. C. Kim, M. N. Cook and S. M. Dymecki (2015). "Multi-Scale Molecular Deconstruction of the Serotonin Neuron System." Neuron **88**(4): 774-791.

Olney, A. H. (2007). "Macrocephaly syndromes." Semin Pediatr Neurol **14**(3): 128-135.

Peleg, S., F. Sananbenesi, A. Zovoilis, S. Burkhardt, S. Bahari-Javan, R. C. Agis-Balboa, P. Cota, J. L. Wittnam, A. Gogol-Doering, L. Opitz, G. Salinas-Riester, M. Dettenhofer, H. Kang, L. Farinelli, W. Chen and A. Fischer (2010). "Altered histone acetylation is associated with age-dependent memory impairment in mice." Science **328**(5979): 753-756.

Potter, H., H. J. Chial, J. Caneus, M. Elos, N. Elder, S. Borysov and A. Granic (2019). "Chromosome Instability and Mosaic Aneuploidy in Neurodegenerative and Neurodevelopmental Disorders." Front Genet **10**: 1092.

Prolla, T. A. (2002). "DNA microarray analysis of the aging brain." Chem Senses **27**(3): 299-306.

Reddy, A. S., D. O'Brien, N. Pisat, C. T. Weichselbaum, K. Sakers, M. Lisci, J. S. Dalal and J. D. Dougherty (2017). "A Comprehensive Analysis of Cell Type-Specific Nuclear RNA From Neurons and Glia of the Brain." Biol Psychiatry **81**(3): 252-264.

Rosenkrantz, J. L. and L. Carbone (2017). "Investigating somatic aneuploidy in the brain: why we need a new model." Chromosoma **126**(3): 337-350.

Salinas, R. D., D. R. Connolly and H. Song (2020). "Invited Review: Epigenetics in neurodevelopment." Neuropathol Appl Neurobiol **46**(1): 6-27.

Saurat, N., T. Andersson, N. A. Vasistha, Z. Molnar and F. J. Livesey (2013). "Dicer is required for neural stem cell multipotency and lineage progression during cerebral cortex development." Neural Dev **8**: 14.

Saxena, A., A. Wagatsuma, Y. Noro, T. Kuji, A. Asaka-Oba, A. Watahiki, C. Gurnot, M. Fagiolini, T. K. Hensch and P. Carninci (2012). "Trehalose-enhanced isolation of neuronal sub-types from adult mouse brain." Biotechniques **52**(6): 381-385.

Scandaglia, M., J. P. Lopez-Atalaya, A. Medrano-Fernandez, M. T. Lopez-Cascales, B. Del Blanco, M. Lipinski, E. Benito, R. Olivares, S. Iwase, Y. Shi and A. Barco (2017). "Loss of Kdm5c Causes Spurious Transcription and Prevents the Fine-Tuning of Activity-Regulated Enhancers in Neurons." Cell Rep **21**(1): 47-59.

Schaum, N., B. Lehallier, O. Hahn, R. Palovics, S. Hosseinzadeh, S. E. Lee, R. Sit, D. P. Lee, P. M. Losada, M. E. Zardeneta, T. Fehlmann, J. T. Webber, A. McGeever, K. Calcuttawala, H. Zhang, D. Berdnik, V. Mathur, W. Tan, A. Zee, M. Tan, C. Tabula Muris, A. O. Pisco, J. Karkanias, N. F. Neff, A. Keller, S. Darmanis, S. R. Quake and T. Wyss-Coray (2020). "Ageing hallmarks exhibit organ-specific temporal signatures." Nature **583**(7817): 596-602.

Sen, P., P. P. Shah, R. Nativio and S. L. Berger (2016). "Epigenetic Mechanisms of Longevity and Aging." Cell **166**(4): 822-839.

Shen, E. Y., Y. Jiang, B. Javidfar, B. Kassim, Y. E. Loh, Q. Ma, A. C. Mitchell, V. Pothula, A. F. Stewart, P. Ernst, W. D. Yao, G. Martin, L. Shen, M. Jakovcevski and S. Akbarian (2016). "Neuronal Deletion of Kmt2a/Mll1 Histone Methyltransferase in Ventral Striatum is Associated with Defective Spike-Timing-Dependent Striatal Synaptic Plasticity, Altered Response to Dopaminergic Drugs, and Increased Anxiety." Neuropsychopharmacology **41**(13): 3103-3113.

Shepherd, C. E., Y. Yang and G. M. Halliday (2018). "Region- and Cell-specific Aneuploidy in Brain Aging and Neurodegeneration." Neuroscience **374**: 326-334.

Shilatifard, A. (2012). "The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis." Annu Rev Biochem **81**: 65-95.

Shulha, H. P., I. Cheung, C. Whittle, J. Wang, D. Virgil, C. L. Lin, Y. Guo, A. Lessard, S. Akbarian and Z. Weng (2012). "Epigenetic signatures of autism: trimethylated H3K4 landscapes in prefrontal neurons." Arch Gen Psychiatry **69**(3): 314-324.

Siegmund, K. D., C. M. Connor, M. Campan, T. I. Long, D. J. Weisenberger, D. Biniszkiewicz, R. Jaenisch, P. W. Laird and S. Akbarian (2007). "DNA methylation in the human cerebral cortex is dynamically regulated throughout the life span and involves differentiated neurons." PLoS One **2**(9): e895.

Stahl, R., T. Walcher, C. De Juan Romero, G. A. Pilz, S. Cappello, M. Irmler, J. M. Sanz-Aquela, J. Beckers, R. Blum, V. Borrell and M. Gotz (2013). "Trnp1 regulates expansion and folding of the mammalian cerebral cortex by control of radial glial fate." Cell **153**(3): 535-549.

Stefanelli, G., A. B. Azam, B. J. Walters, M. A. Brimble, C. P. Gettens, P. Bouchard-Cannon, H. M. Cheng, A. M. Davidoff, K. Narkaj, J. J. Day, A. J. Kennedy and I. B. Zovkic (2018). "Learning and Age-Related Changes in Genome-wide H2A.Z Binding in the Mouse Hippocampus." Cell Rep **22**(5): 1124-1131.

Stilling, R. M., E. Benito, M. Gertig, J. Barth, V. Capece, S. Burkhardt, S. Bonn and A. Fischer (2014). "De-regulation of gene expression and alternative splicing affects distinct cellular pathways in the aging hippocampus." Front Cell Neurosci **8**: 373.

Striedter, G. F. (2005). Principles of brain evolution, Sinauer associates.

Sun, T. and R. F. Hevner (2014). "Growth and folding of the mammalian cerebral cortex: from molecules to malformations." Nat Rev Neurosci **15**(4): 217-232.

Sweatt, J. D. (2013). "The emerging field of neuroepigenetics." Neuron **80**(3): 624-632.

Tabula Muris, C. (2020). "A single-cell transcriptomic atlas characterizes ageing tissues in the mouse." Nature **583**(7817): 590-595.

Takata, A., B. Xu, I. Ionita-Laza, J. L. Roos, J. A. Gogos and M. Karayiorgou (2014). "Loss-of-function variants in schizophrenia risk and SETD1A as a candidate susceptibility gene." Neuron **82**(4): 773-780.

Taverna, E., M. Gotz and W. B. Huttner (2014). "The cell biology of neurogenesis: toward an understanding of the development and evolution of the neocortex." Annual review of cell and developmental biology **30**: 465-502.

Toyama, B. H., J. N. Savas, S. K. Park, M. S. Harris, N. T. Ingolia, J. R. Yates, 3rd and M. W. Hetzer (2013). "Identification of long-lived proteins reveals exceptional stability of essential cellular structures." Cell **154**(5): 971-982.

Tunovic, S., J. Barkovich, E. H. Sherr and A. M. Slavotinek (2014). "De novo ANKRD11 and KDM1A gene mutations in a male with features of KBG syndrome and Kabuki syndrome." Am J Med Genet A **164A**(7): 1744-1749.

von Schimmelmann, M., P. A. Feinberg, J. M. Sullivan, S. M. Ku, A. Badimon, M. K. Duff, Z. Wang, A. Lachmann, S. Dewell, A. Ma'ayan, M. H. Han, A. Tarakhovsky and A. Schaefer

(2016). "Polycomb repressive complex 2 (PRC2) silences genes responsible for neurodegeneration." Nat Neurosci **19**(10): 1321-1330.

Waddington, C. H. (1942). "The epigenotype. 1942." Int J Epidemiol **41**(1): 10-13.

Wang, C. M., S. N. Tsai, T. W. Yew, Y. W. Kwan and S. M. Ngai (2010). "Identification of histone methylation multiplicities patterns in the brain of senescence-accelerated prone mouse 8." Biogerontology **11**(1): 87-102.

Webb, W. M., R. G. Sanchez, G. Perez, A. A. Butler, R. M. Hauser, M. C. Rich, A. L. O'Bierne, T. J. Jarome and F. D. Lubin (2017). "Dynamic association of epigenetic H3K4me3 and DNA 5hmC marks in the dorsal hippocampus and anterior cingulate cortex following reactivation of a fear memory." Neurobiol Learn Mem **142**(Pt A): 66-78.

Wei, J. W., K. Huang, C. Yang and C. S. Kang (2017). "Non-coding RNAs as regulators in epigenetics (Review)." Oncol Rep **37**(1): 3-9.

Woods, C. G., J. Bond and W. Enard (2005). "Autosomal recessive primary microcephaly (MCPH): a review of clinical, molecular, and evolutionary findings." Am J Hum Genet **76**(5): 717-728.

Xu, Y., S. Zhang, S. Lin, Y. Guo, W. Deng, Y. Zhang and Y. Xue (2017). "WERAM: a database of writers, erasers and readers of histone acetylation and methylation in eukaryotes." Nucleic Acids Res **45**(D1): D264-D270.

Yurov, Y. B., S. G. Vorsanova, I. A. Demidova, A. D. Kolotii, I. V. Soloviev and I. Y. Iourov (2018). "Mosaic Brain Aneuploidy in Mental Illnesses: An Association of Low-level Post-zygotic Aneuploidy with Schizophrenia and Comorbid Psychiatric Disorders." Curr Genomics **19**(3): 163-172.

Zech, M., S. Boesch, E. M. Maier, I. Borggraefe, K. Vill, F. Laccone, V. Pilshofer, A. Ceballos-Baumann, B. Alhaddad, R. Berutti, W. Poewe, T. B. Haack, B. Haslinger, T. M. Strom and J. Winkelmann (2016). "Haploinsufficiency of KMT2B, Encoding the Lysine-Specific Histone Methyltransferase 2B, Results in Early-Onset Generalized Dystonia." Am J Hum Genet **99**(6): 1377-1387.

Zeisel, A., A. B. Munoz-Manchado, S. Codeluppi, P. Lonnerberg, G. La Manno, A. Jureus, S. Marques, H. Munguba, L. He, C. Betsholtz, C. Rolny, G. Castelo-Branco, J. Hjerling-Leffler and S. Linnarsson (2015). "Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq." Science **347**(6226): 1138-1142.

Zhang, W., J. Qu, G. H. Liu and J. C. I. Belmonte (2020). "The ageing epigenome and its rejuvenation." Nat Rev Mol Cell Biol **21**(3): 137-150.

Zilles, K., N. Palomero-Gallagher and K. Amunts (2013). "Development of cortical folding during evolution and ontogeny." Trends Neurosci **36**(5): 275-284.

Acknowledgements

9. Acknowledgements

First and foremost, my utmost and sincere gratitude goes to Prof. Dr. Andre Fischer, for not only giving me the opportunity carry out my PhD thesis work in his lab, but also to trust me with my ideas, allowing me to work independently in the lab, giving me enormous guidance, inspiration and motivation during ups and downs, which have supported greatly to shape my personal and professional development. You endowed me with your precious time, money and trust, for which, I am forever grateful! No matter what questions I had, you always kept your door open for me and gave your valuable and helpful comments. Thank you very much, Andre! My sincere gratitude goes to Dr. Fahranaz Sananbenesi as well for all those cheerful moments and conversations we had in the lab that really went a long way to keep me motivated!

This thesis work would have been incomplete without the supervision and fruitful suggestions of Prof. Dr. Tiago Outeiro and Dr. Camin Dean! Thank you for giving your precious time to help me and guide me through the complex maze of many different research directions to reach the end of this thesis. I truly cherish your selfless help for reviewing my thesis! Furthermore, I would like to express my sincere gratitude towards the extended members of the examination committee, Prof. Dr. Alexander Flügel, Dr. Dr. Oliver Schlüter and Prof. Dr. Argyris Papantonis. Thank you very much in advance for agreeing to evaluate my PhD thesis work and attending my defense!

I would also like to cordially thank Dr. Tran Tuoc for all the fruitful collaboration works regarding mouse brain folding project and agreeing to incorporate the manuscript in this thesis. Along with my studies regarding learning, memory and neurodegeneration, I am grateful to you for introducing me to the beautiful world of neurodevelopment! My sincere appreciation goes to Dr. Francesca Odoardi from Flügel lab for the enormous help with setting up the FACS machine.

Now, my "second family" here in Germany, the Fischer lab peeps! You guys have shaped me who I am today, therefore, I want to personally thank all of you individually, regardless being the former and current members of the lab! Since my first exposer to Fischer lab back in 2015 during my lab rotation, then my master's thesis in 2016 and till my last 4 years of PhD, I cherish every single memory! I can write a mini thesis about all the help, support and friendly gestures I got from all of you, seriously! My sincere gratitude goes to Dr. Cemil "amazing" Kerimoglu, who has been my supervisor, colleague and friend! Theoretically, I was a toddler when I started working in the fields of science and your

"working hard, enthusiastically" motivated me immensely to push boundaries and work hard. Because of your unique persona, I always felt at home around you! Thank you, Cemil! And Sasha!!! How can I ever thank you enough for being the best lab mate and office mate at the same time! First of all, I really appreciate your trust on me to be a part of your Setd1b project and sharing the workloads together! Apart from working, you have been truly a dear friend of mine, with whom I could trust and share all about my good and challenging times in the lab! I learned a lot from you and will cherish all of our memories for years to come! Thank you Sasha! My sincere gratitude goes to Naim vai (Reza), who has been a big brother of mine from another mother! I have been around you for almost 10 years now and wish to do the same in future as well! Thanks for all the support, vai! I am grateful to Lalit being such an awesome mentor and being a symbol of strength and resilience to drive scientific research further. I learned a lot from you! My special thanks goes to Susi for being the fairy godmother, with whom I could rely upon every single time for all the sequencing works and non-work related stuffs as well! I want to thank Ricardo for all the jokes, scientific conversations, cleaning up enrichment cages, specially your help to do my experiments for long non coding RNA project while I was sick! Thanks for being such a nice soul! Same goes for Masha, thank you so much for taking care of my experiments while I was at the hospital! Means a lot to me! My gratitude goes to Henning for all the fun and sarcasms you keep on doing, hilarious, man! Keep doing that! And my birthday mate, Robert!! I so wished we could have submitted the PhD thesis on the same time, like we were born and started our doctoral journey (sad face :-/) together on this earth! But your confidence and way of delivering speeches have always fascinated me! Thanks for always inspiring! My sincere thanks goes to Tona and Dennis, who have been incredibly helpful and instrumental to solve all bioinformatic and IT related problems I faced. My special thanks goes to Ulrike and Daniel! Both of you have been my source of refuge, whenever I had any administrative questions or tasks to be fulfilled. You have solved all my problems and supported me greatly! I would also like to thank Jiayin, Anna-Lena, Aditi, Jonathan, Joy, Masha M., and Sophie for all the good moments and chats. Even though the following persons are not currently present at the lab, I would like to send my best sincere gratitude to Carmelo, Melanie, Tea, Hendrik, Eva, Christian, Rashi, Magda, Vincenzo, Gaurav, Vakho and Miki, for all the advices and supervisions, I wholeheartedly cherish them! Lastly, my PhD experience would have been incomplete if I didn't have my

mentoring experience with my students, Agnes, Iga, Arda, Linda and Ranjit! Thanks for your contributions.

During this thesis, along with my thesis projects, fruitful collaborations have been made with Dr. Eugenio Fornasiero (Neuro- & Sensory Physiology, UMG), Pauline Ulmke (Tran lab, Neuroanatomy), Prof. Dr. Alexander Dityatev (DZNE Magdeburg), Prof. Dr. Gregor Eichele (MPI-BPC), Dr. med. Malte Tiburcy (UMG), Dr. Patapia zafeiriou (UMG) and Dr. med. Wolfram-Hubertus Zimmermann (UMG). Even though those were not part of the thesis, I am sure those projects will go a long way and yield great discoveries in future.

The IMPRS Neuroscience and GGNB office were incredibly helpful during my PhD times in Göttingen. Thanks to Jonas, Sandra, Franziska and Kirsten for all the support! My sincere appreciations goes to Prof. Dr. Agnès Nadjar (University of Bordeaux, France) and Florina Camarasu of the Neurasmus scholarship program, for which, I came to Göttingen to do my master's degree and eventually join as a PhD in Fischer lab. In this regard, I would like to remember late Prof. Dr. Michael Hörner, who I miss so dearly. 6 years ago, in February 2014, I did my skype interview with him and got selected to join the IMPRS neuroscience program and become a "Neuro". I wish you were here now.

The acknowledgements will be incomplete if I do not thank my "first family", firstly my dear wife, Samantha! From the moment of my accident and surgery 3 months ago, you have been selflessly taking care of me while doing your own PhD! Words cannot express how lucky I am to grow a family with you! Without your support, I would not have been able to write this whole thesis, let alone complete my PhD! Thank you, Samantha! While I am away from my home in Bangladesh, I want to thank my father and mother, to constantly give me moral support through video calls, loving so much and keeping trust on me. Thank you Abbu & Ammu! My gratitude goes to my little sister, Retee for all her lovely gestures and support. Can't wait to see you all in person again! Lastly, the Bangladeshi community in Göttingen have been an absolute treasure during my time here in Göttingen. Thanks goes to all my relatives and friends in Bangladesh as well. Last but not the least, all praise and thanks goes to the Almighty, the most merciful and compassionate, for giving me the ability to complete this PhD thesis in full! Alhamdulillah!

THE END